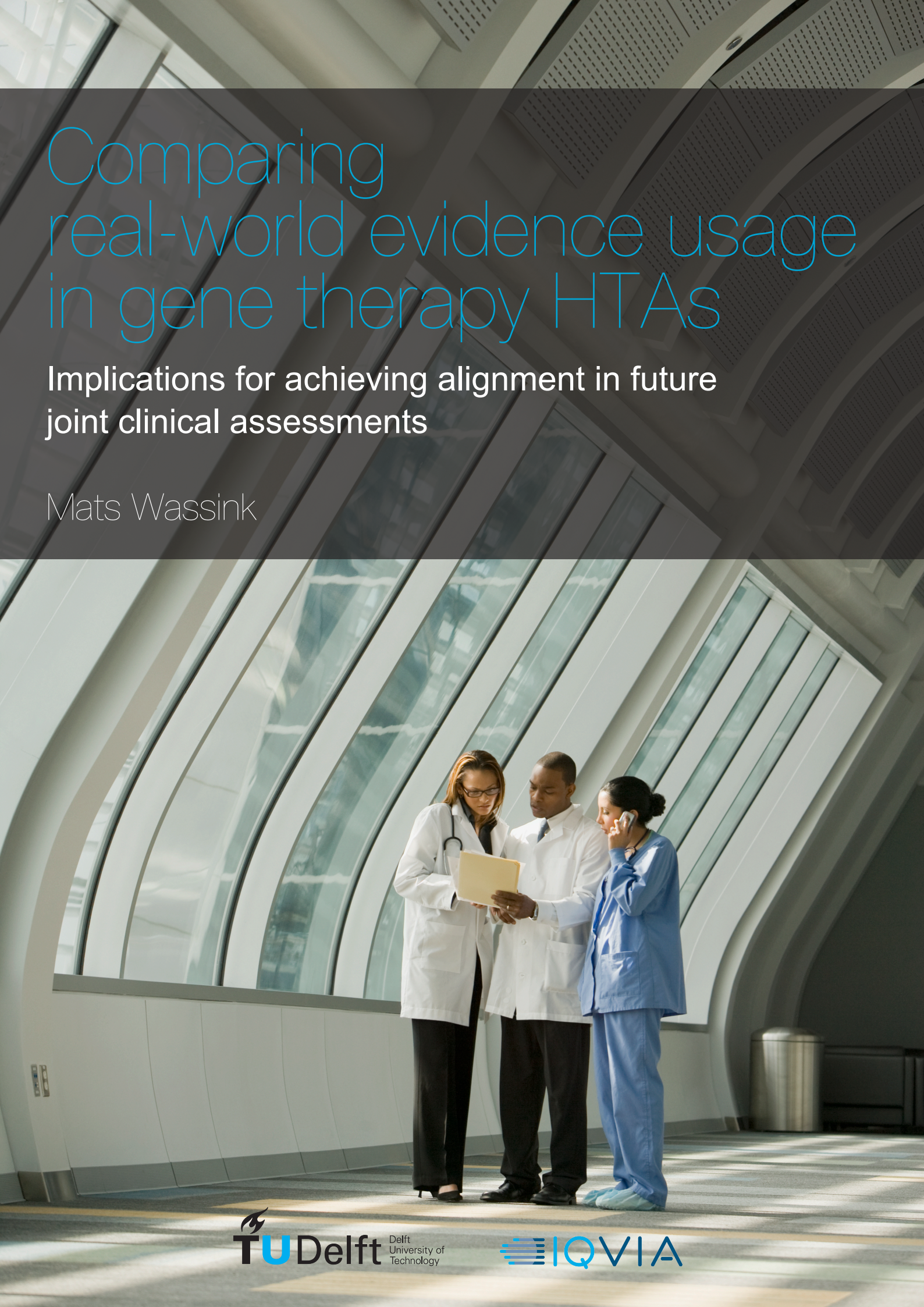
Comparing real-world evidence usage in gene therapy HTAs

Implications for achieving alignment in future joint clinical assessments

Mats Wassink



TU Delft Delft University of Technology

IQVIA

This page was intentionally left blank

Comparing real-world evidence usage in gene therapy health technology assessments

Implications for achieving alignment in future joint clinical assessments

by

ir. Mats Wassink

in partial fulfillment of the requirements for the degree of

Master of Science

in **Management of Technology**

at the Delft University of Technology,

to be defended publicly on Tuesday March 4, 2022 at 14:00.

Student number:	4377613	
Programme:	MSc Management of Technology – MSc Thesis Project (MOT2910 - 30 ECTS)	
Location:	TU Delft, Faculty of Technology, Policy & Management	
First supervisor:	dr. ir. Bert Enserink	TU Delft, Multi-Actor Systems - Section of Policy Analysis
Second supervisor:	ir. Asli Boru	TU Delft, Delft Centre for Entrepreneurship
Advisor:	dr. Saba Hinrichs-Krapels	TU Delft, Multi-Actor Systems - Section of Policy Analysis
External supervisor:	Tom Nijhuis, MSc	IQVIA, Global Consulting Services
External supervisor:	Tom Constandse, MSc	IQVIA, Global Consulting Services

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft Delft University of Technology

IQVIA

Executive Summary

Background & Objective The unprecedented value and long-term uncertainties of gene therapies have challenged established health technology assessment (HTA) methods. Real-world data (RWD) and real-world evidence (RWE) have gained traction for their potential role in filling the evidence gap that HTA bodies encounter in appraising gene therapies. Yet, the existing body of literature fails to specify what role real-world data currently plays and could potentially play in future gene therapy HTAs.
Substantial differences in amenability could pose feasibility challenges in aligning HTA bodies for future joint clinical assessments. As such, the objective of this research is to *"identify feasibility challenges in alignment for EU-wide gene therapy joint clinical assessments, based on the current and future role that real-world data and real-world evidence play in gene therapy HTAs"*.

Methods An initial literature review laid the theoretical foundation for the research. It unveiled a scarcity of literature that delivers empirical evidence on gene therapy HTA practices and the role that RWD/RWE plays in HTAs. A multi-faceted retrospective comparative analysis of EMA-approved gene therapy HTAs delivered this empirical evidence. Preliminary findings were probed in three use cases and verified in semi-structured interviews.

Results Nineteen HTA reports published by the HTA bodies G-BA (Germany) and NICE (England) were identified for the ten in-scope gene therapies.
Most challenges and considerations in gene therapy HTAs were similar to that of other therapy types. Similarly, HTA bodies have no frameworks or payments schemes tailored explicitly to gene therapies.

Both the absolute volume of RWD/RWE usage and the RWD/RWE acceptance rates of G-BA and NICE differed substantially; Whereas NICE had an average inclusion of 14 sources per HTA report (with an acceptance rate of 56%), G-BA had 8 (with 32% acceptance rate).
RWD/RWE was found to have the lowest acceptance rate if it supports evidence on an external comparator. On the other hand, RWD/RWE supporting the effectiveness of the intervention is relatively often accepted by both NICE and G-BA.
Based on the exclusion rationales, two main factors for not accepting RWD/RWE were identified. One was insufficient information to substantiate the choice of RWD/RWE; the other was an inappropriate RWE study design, which does not reflect the standard of care practices.

Conclusion The retrospective comparative analysis unveiled differences in RWD/RWE usage in gene therapy HTAs, which may lead to feasibility challenges in future joint clincal assessments. To optimise the transferability of these outcomes to national HTA bodies, alignment on assessment elements and evidentiary requirements is necessary. This research proposes that, based on differences in RWD/RWE usage, particular aspects of the PICO framework may be more difficult to align on than others. Achieving alignment in a multi-stakeholder environment may be challenging, as differences in available resources and existing knowledge result in differences in absorptive capacity among HTA bodies. Such differences should be considered in future alignment strategies, as they may impede collaboration efforts, which play a key role in facilitating alignments and the corresponding knowledge transfer. To this point, the research provides 'stepping stones' for future research to implement knowledge diffusion models and formulate strategies to increase adoption of research output
Moreover, this research proposes that when manufacturers consider submitting RWD/RWE for gene therapy HTAs, they should be critical towards the data that they submit and towards the appropriateness of the RWE study design. To guide such reflections and prevent the potential additional administrative burden from joint clinical assessments, the work proposes that manufacturers should take an increasingly active role in the innovation system.

Acknowledgements

For what it's worth, writing a thesis does not get easier the second time around. Yet, thanks to the many people supporting me, it was an enjoyable experience nonetheless.

Formulating a thesis topic that allowed me to combine my background in life sciences with insights from the 'Management of Technology' study programme took me a while. **Saba**, thank you for providing a perfect balance of guidance and freedom in this period of finding a suitable topic, that continued along the way as being my informal first supervisor. I am grateful for our weekly meetings, where your contagious energy and enthusiasm drove me to "think outside the box" and "top it up a notch."
Asli, thank you for your input and feedback during my research. You challenged me to be critical of myself and guided me to interpret my research from different perspectives. It truly enhanced my research.
I also want to express my gratitude to **Bert**; thank you for agreeing to be my formal first supervisor, despite a thesis topic that may be outside of your usual research interests.

Moreover, I would like to thank the **interviewees** who participated in this research. Your input helped me to validate and interpret the findings presented in this research.

This thesis was written as part of my internship with the **IQVIA** life science consulting services team, an opportunity that I am grateful for. During my internship, I got acquainted with strategy consulting in the life sciences industry and I learned new ways to present data. I have applied these newly developed skills in this thesis.
A special thanks within IQVIA go to Tom and Tom. I want to thank **Tom C.** for your support, guidance and enthusiasm during our weekly catch up. Your expertise and critical view provided much-needed nuances in this thesis. **Tom N.**, thank you for giving me the opportunity me to write my thesis with the consulting services team. My thanks and appreciations also go to my **colleagues at IQVIA** for creating a welcoming environment and the helpful discussions along the way.

This thesis concludes my time at the TU Delft. I want to thank my **family and friends** for their support throughout my seven and a half years of studying. A special shout out to my **housemates**, for the fruitful 'thuis UB' sessions in our lounge in the last period of writing this thesis. Finally, I want to express my gratitude towards **Heike, Jos, Rebecca, Imke** and **Anne** for your unconditional love and encouragements, allowing me to be my best self.

Mats Wassink
Rotterdam, February 2022

Contents

List of Figures

List of Tables

List of Abbreviations

ATMP	Advanced therapy medicinal product
CEA	Cost effectiveness assessment
EC	European Committee
EMA	European Medicines Agency
EU	European Union
EU4	Germany, France, Spain, Italy
FDA	US Food and Drug Administration
G-BA	Gemeinsamer Bundesausschuss
HST	Highly specialised technology
HTA	Health technology assessment
NICE	National Institute for Health and Care Excellence
QALY	Quality adjusted life year
RCT	Randomised controlled trial
REA	Relative effectiveness assessment
RWD	Real-world data
RWE	Real-world evidence
UK	United Kingdom
US	United States

1 Introduction

Biomedical advances and innovations have considerably improved healthcare outcomes across the world. This is especially true for the early 21st century, following multiple synergistic scientific discoveries, including the finalisation of the human genome project and the development of the genetic modification tool CRISPR/Cas9 (Green *et al*., 2015). In fact, the recent record-breaking speed of the COVID-19 vaccine development is considered to symbolise a "renaissance of scientific innovation" (Ernst & Young LLP, 2021). Indeed, while revolutionary technologies like cell & gene therapies, mRNA vaccines, and personalised medicine once appeared fiction, these health technologies have now started to become a reality.

1.1. Market approval and market access of health technologies

Health technologies refer to an "intervention developed to prevent, diagnose, or treat medical conditions; promote health; provide rehabilitation; or organise healthcare delivery" (HTA Glossary, 2021). As such, they play a crucial role in solving global challenges that healthcare systems face (Farid, 2019). Nevertheless, biomedical innovation is merely one aspect of getting novel health technologies to the patients. Subsequent steps generally include obtaining marketing authorisation, health technology assessments, and price negotiations.

The first step, marketing authorisation, is granted by regulatory agencies, i.e. the Food and Drug Administration (FDA) in the United States or the European Medicines Agency (EMA) in the European Union (EU). These agencies assess the evidence on the quality and safety of a new health technology for a specific patient population. After receiving marketing authorisation, pharmaceutical companies may need to obtain reimbursement for their product. In most countries, reimbursement, also referred to as market access, can be vital for patient access because the prices are often too high for individuals to pay for themselves (Zaprutko *et al*., 2017). If a health technology is reimbursed, it is considered a national health expenditure cost that will be covered by social health insurance or national health services. Reimbursement decision-making and corresponding price negotiations with the health technology manufacturer are often based on health technology assessments (HTAs).

Health technology assessments

An HTA is defined as "a multidisciplinary process that uses explicit methods to determine the value of a health technology at different points in its lifecycle" (O'Rourke *et al*., 2020). Its purpose is to inform decision-makers, thereby promoting an equitable, efficient, and high-quality health system (O'Rourke *et al*., 2020). HTAs have been compared to a bridge between science and policy that allows for the transfer of knowledge derived from scientific research to the decision making process (Battista, 1996). The output of an HTA is a recommendation regarding the reimbursement of a particular health technology that could be classified as positive, restricted or negative. This recommendation is then taken into account by public and private payers in pricing and reimbursement decisions (Trosman *et al*., 2011).

Health technology assessment bodies

Contrary to the centralised marketing authorisation process in the EU, these reimbursement decisions are made on a national level (Kleijnen *et al*., 2012). Each HTA body assesses the added value of a new health technology in the context of its local standard of care (Van Nooten *et al*., 2012). Therefore,

a particular health technology may receive reimbursement in one country but not in the other.
Despite the differences between HTA bodies, four archetypes can generally be defined (IQVIA, internal research and Jean *et al*. (2018)). The (former) EU5 countries (e.g. France, Germany, Italy, Spain, and the UK) tend to give a good overview of the different archetypes (Figure 1.1). It should be noted that these archetypes are not stringent and that HTA bodies may consider aspects that are also relevant for other HTA body archetypes.

Archetypes	Primary Goal	Country Examples *(not exhaustive)*
Relative Effectiveness Assessments	Comparing clinical evidence of a new product to an existing comparator	
Cost-Effectiveness	Understand value for money of a new treatment vs. an existing treatment	
Budget Optimization	Efficiently allocating the limited budget/resources	
Competitive Rationalizing Free Market	Commercial viability – attractive plan design that generates business and profit	 *(MCOs & PBMs)*

Figure 1.1. HTA Archetypes in the (former) EU5 + US. Modified from IQVIA internal research. **MCO**: Managed care organisation. **PBM**: Pharmacy benefit managements.

HTA bodies within the comparative clinical effectiveness archetype (i.e. France & Germany) are particularly interested in the additional clinical benefit that a new therapy may bring to the healthcare system compared to existing alternatives. Therefore, this approach is often referred to as a relative effectiveness assessment (REA).
HTA bodies in the cost-effectiveness archetype (i.e. UK) are generally more concerned with the value for money that a new treatment will bring. This value is often evaluated using cost-effectiveness assessment (CEA) models. In such assessments, the incremental cost per quality-adjusted life-year (QALY) is modelled against a willingness-to-pay threshold.
For markets such as Spain and Italy, the HTA outcome is primarily determined by the budget impact of a new health technology, given the inevitable resource constraints. The financial consequences of a new health technology within a specific healthcare setting are estimated using a budget impact model at either a national, regional, or local level.
The remaining group of HTA bodies can be categorised as free-market payers. The primary objective of these private agencies is to be profitable. HTA outcomes are based on the formulary design of a health technology and negotiations with the manufacturer.

Joint clinical assessments

While HTA bodies focus on different aspects of a drug's value, it appears that nationally performed REAs may result in considerable duplication of work and inefficient use of resources (Garattini and Padula, 2020).
To facilitate knowledge transfer between different agencies on HTA methodologies, the European Network for Health Technology Assessment (EUnetHTA) was established (EunetHTA, 2021).

EunetHTA aims to "increase quality and efficiency of joint HTA work at the European level" (EunetHTA, 2021) through joint clinical assessments. Similar collaborative approaches to HTAs are explored by different EU-member states (BeNeLuxA, 2021).
Recently, the European Commission adopted a new regulation on HTAs that formalises a centralised, supranational approach (European Commission, 2021). The prospective EU-wide HTA process aims to harmonise the national HTA processes to generate a single, joint clinical assessment, focusing on

REAs of innovative medicines and medical devices. However, further clarity and alignment on aspects like evidence requirements is needed (Kanavos *et al*., 2019). Substantial differences in evidence requirements could pose feasibility challenges to the transferability to national HTA bodies, potentially leading to additional administrative and regulatory barriers (Kanavos *et al*., 2019).

Real-world data and real-world evidence

Varying evidence requirements may therefore pose a considerable challenge to the joint clinical assessment concept. In traditional evidence-based medicine, particular clinical study methodologies are placed in a hierarchy based on the relative strength of evidence they deliver (Figure 1.2).
Double-blind, randomised controlled trials have long been considered as the golden standard for evidence generation on clinical efficacy and safety (Velasco-Garrido and Buss, 2005). However, the inherent uncertainty regarding the (long-term) benefits of gene therapies may often not be captured within the conventional time-frame of such trials.

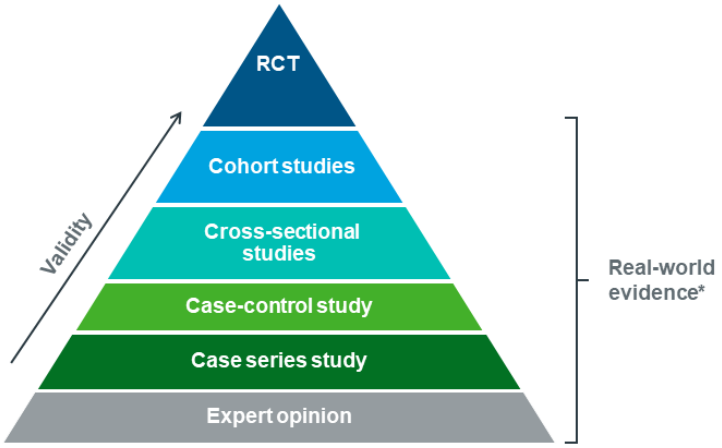


Figure 1.2. General hierarchy of evidence. Based on Katkade *et al*. (2018) and Murad *et al*. (2016). *: Not exhaustive. RCT: Randomised controlled trail.

Alternative study designs may provide supplementary evidence on the studied health technology. Such studies have increasingly been embraced in healthcare systems as real-world data (RWD) and real-world evidence (RWE). Especially in the medical device industry, it has been widely adopted (Sherman *et al*., 2016).

Contrary to randomised controlled trials, RWE generally has a low internal validity and high external validity. In other words, while the randomised controlled trials may be more valuable in demonstrating causality, RWD and evidence may provide helpful information on the outcomes of the health technology in a setting that is representative of routine clinical practice. Therefore, real-world studies may supplement randomised controlled trials to fill evidentiary gaps on the relative effectiveness of a health technology.
However, HTA bodies appear reluctant to wide-scale adoption of RWD/RWE. In part, its adoption is impeded by the association of RWD/RWE with confounding bias and the lack of quality and transparency in data (Bowrin *et al*., 2019). Another major hurdle impeding widespread use of RWD and RWE in HTAs is the lack of guidance for HTA bodies on what RWD and -evidence entails, how to appraise it and what it can be used for (Makady *et al*., 2017b). At the same time, lacking guidance from HTA bodies for the industry on the RWD quality requirements has also impeded adoption.

Gene therapy medicinal products

A therapeutic area where RWD/RWE may play an increasingly important role, is that of gene therapy medicinal products (from here on referred to as 'gene therapies'). Following the definition provided by the EMA, gene therapies work by "inserting 'recombinant' genes into the body, usually to treat a variety of diseases, including genetic disorders, cancer or long-term diseases" (European Medicines Agency, 2021). This translates to these therapies having the potential to target the underlying cause of genetic conditions and acquired diseases and potentially prevent, treat or cure genetic conditions and hereditary diseases. However, this potential value appears to be accompanied by extremely high pricing. Although many factors may contribute to this price, two, in particular, are recurrently mentioned in literature. On the one hand, this price is derived from the high costs and risks associated with developing and manufacturing these therapy types (Angelis *et al.*, 2020). The paradigm used in the price-setting strategy, on the other hand, also appears to play a role. Advocates justify the price by the considerable savings from curing chronic conditions that would otherwise require more costly lifelong medical interventions (Ylä-Herttuala, 2015).

Nevertheless, much like the potential value, the potential long-term patient benefits are uncertain. Evidence on long-term effectiveness is scarce during reimbursement decision-making, as the claimed effectiveness may exceed the time horizon of clinical trials that support the HTA dossiers. Although extrapolation of costs and effects is not uncommon in economic evaluations, little evidence and experience substantiate treatment durability assumptions. Moreover, gene therapies commonly target rare diseases, where it may not be feasible to conduct double-blind, randomised clinical trials due to practical or ethical reasons (Hettle *et al.*, 2017; Coyle *et al.*, 2020). Considering these inherent challenges, RWD/RWE may play an important role in filling in the encountered evidence gaps.

1.2. Problem statement

HTA assessment methods originate from the 1970s and have been incrementally adapted to new health technologies and changing healthcare systems (Banta, 2009). However, the recent introduction of gene therapies has demonstrated that established HTA methods may no longer suffice. The unprecedented potential value of these therapies, combined with the long-term uncertainty of health outcomes, challenge these methods that are commonly designed to capture short-term and direct impacts (Leyens and Brand, 2016). As a result, gene therapies "push against the boundaries of the methodological and budgetary capacity available" (Angelis *et al.*, 2020).

RWD/RWE may be essential in filling the evidence gap that HTA bodies encounter in appraising these therapy types. The potential of RWE usage is increasingly recognised in the literature. However, the lack of consensus on the definition appears to be a factor that impedes adoption in practice (Jaksa *et al.*, 2021). The authors note that HTA bodies may not align on questions that could be answered by RWD/RWE and on how its quality could be assessed. This misalignment may be further confounded by the varying quality of evidence within a particular study design in the evidence hierarchy (Figure 1.2) (Murad *et al.*, 2016). Similarly, RWE guidance for and from HTA bodies appears limited, forming another barrier for adoption from HTAs and the industry.

Despite the increased attention for RWD/RWE, empirical evidence on how gene therapies are currently assessed in HTAs appears to be scarce. The literature also fails to specify what role RWD/RWE currently plays and could potentially play in the future in gene therapy HTAs. As noted earlier, substantial differences in evidentiary requirements could potentially pose feasibility challenges joint clinical assessment concept. As such, varying amenability of RWD/RWE usage by HTA bodies could be indicative of potential challenges in aligning HTA bodies for future joint clinical assessments, thereby impeding transferability of its outcomes.

1.3. Research objectives & questions

The need to tailor HTAs to gene therapies is becoming more urgent, given the prospect of an increasing number of cell & gene therapies seeking market access in the coming years. Indeed, FDA (2019b) predict that from 2025, 10–20 cell and gene therapy products will be approved annually. In a similar vein, Eder and Wild (2019) recently identified 141 advanced development stage clinical trials (Phase III and IV) investigating cell & gene therapies.

Joint clinical assessments of these technologies may present opportunities to allow for faster and more uniform assessments, as well as improved patient access to innovative health technologies (European Commission, 2021; Kanavos *et al*., 2019). However, Allen *et al*. (2017) note that "in order to move forward to a more harmonised HTA environment within Europe, it is first necessary to understand the variation in HTA practices within Europe". Substantial variations in how HTA bodies assess the added clinical benefit of a new health technology may pose feasibility challenges to joint clinical assessments. Alignment on this aspect is hypothesised to form a boundary condition for future EU-wide clinical assessments of gene therapies.
Therefore, the objective of this thesis is to "identify feasibility challenges for alignment in EU-wide gene therapy HTAs, based on how they are currently assessed and the current and future role that RWE plays in the decision-making process".

Following the research objective, the main question to be addressed in this research is:

" What are the implications of RWD/RWE usage by HTA bodies for achieving alignment in future joint clinical assessments of gene therapies? "

To conduct this research in a structured way, several sub-questions are defined:

1. *How are gene therapies currently assessed in HTAs?*
   This sub-question provides insight into the recent gene therapy HTA landscape developments. The objective is twofold; the first objective is to gather empirical evidence and understand the current situation. In addition, relevant findings from the literature will be compared to the gathered empirical evidence.

2. (a) *What role does RWD/RWE play in the HTA appraisal process of gene therapies?*

   (b) *How has RWD/RWE usage in gene therapy HTAs evolved?*

   Complementary to sub-question 1, this question aims to develop an overview of the RWE usage in gene therapy HTAs. More specifically, both a quantitative and qualitative understanding of the added value of RWE in these HTAs will be developed.

3. (a) *What factors impact RWD/RWE usage in gene therapy HTA appraisals?*

   (b) *What are steps to be taken to extend RWD/RWE usage in HTA appraisals of gene therapies?*

   Following a practical understanding of the field, this sub-question aims to identify factors such as guidelines or attitudes that may impact RWD/RWE usage. In addition, the findings of the previous sub-questions are integrated to identify potential misfits between factors that may impact RWD/RWE usage and practice. Such misfits may present opportunities for which enablers should be defined.

4. *What are the implications of RWD/RWE usage by HTA bodies for their alignment in the assessment of gene therapies?*
   The final sub-question aims to integrate the findings of the previous sub-questions and identify potential feasibility challenges for alignment between HTA agencies. The outcome is then placed

in the context of the recently accepted legislation that enables EU-wide joint clinical assessment of gene therapies.

5. *What are the implications of RWD/RWE usage by HTA bodies for gene therapy manufacturers?* In addition to the previous question, the findings will also be interpreted from the perspective of gene therapy manufacturers.

1.4. Scientific contribution & societal relevance

From an academic perspective, this research is relevant because the apparent misfit between established HTA methods, gene therapies and the potential role of RWE has only sporadically been linked in literature. This research aims to deliver empirical evidence to address this knowledge gap. By doing so, it aims to contribute to the driving force behind innovation in HTA methodology, as "the improvement of evaluation methods will be driven by academics and not HTA agencies, as the latter tend to be conservative, asking for increasingly large volumes of evidence, without an appetite for innovative methodology" (Pochopień *et al*., 2021).
Moreover, insights derived from varying RWD/RWE usage and amenability can be used in strategy development to achieve alignment of HTA bodies in future European joint clinical assessments.

The research is also considered relevant from a practical perspective. While industry experts appear to have a sense of recent developments, empirical evidence is lacking. Therefore, the generated insights may be used to understand better how RWD/RWE is being used in practice. Moreover, these insights may help gene therapy manufacturers understand relevant considerations when submitting RWD/RWE in future HTAs.

1.5. Scope

The research focuses on RWD and RWE usage in HTAs of gene therapies approved by the EMA between December 2015 and December 2020. The scope is limited to HTA reports published by the German HTA body Gemeinsamer Bundesausschuss (G-BA) England's National Institute for Health and Care Excellence (NICE) between December 2015 and November 2021.

1.6. Thesis layout

This chapter provided an introduction to the research area and the problem that the presented research aims to solve. The thesis proceeds as follows. First, additional background on key terminology is provided, after which a literature review on the intersection of these key terms is presented. The literature review will lay the foundation for the research presented in this thesis. The subsequent chapter elaborates on the research methodology and discusses data collection and analysis methods. Next, in chapter 4, a retrospective analysis of EMA-approved gene therapy HTAs is presented to deliver empirical evidence on this aspect. Three illustrative use-cases provide context to the identified RWD/RWE usage. Preliminary findings and the underlying assumptions will be probed in semi-structured interviews as described in chapter 5. Chapter 6 addresses the limitations of this research and discusses the scientific and managerial contributions. Finally, the thesis concludes with key findings, opportunities for future research, and links to the Management of Technology study programme. The bibliography and appendices will complement this thesis.

2

Background & Literature Review

This chapter provides additional background information on the topics mentioned in the introduction. In addition, a literature review is presented to obtain a better understanding of relevant concepts and the current state of the literature. By analysing and synthesising relevant scientific literature, a knowledge gap may be identified to serve as a basis for the presented research.

2.1. Background

First, additional information will be provided on the concept of joint clinical assessments, as well as the two HTA bodies that are in-scope of this thesis (G-BA (Germany) and NICE (England)). The subsequent section will define real-world evidence and compare it to the golden standard of medical evidence generation. Finally, information is provided on recent developments regarding real-world evidence usage.

HTA bodies

G-BA

In Germany, EMA marketing authorisation grants automatic reimbursement to most medicinal products. These new drugs are priced freely for the first 12 months after launch, pending completion of the G-BA early benefit assessment (locally referred to as Arzneimittelmarkt-Neuordnungsgesetz (AMNOG), which translates to Pharmaceuticals Market Reorganisation Act)).
The manufacturer must submit a benefit dossier to the G-BA at the launch time. The G-BA then commissions the independent Institute for Quality and Efficiency in Health Care (IQWiG) to evaluate the health technology at hand. However, for therapies that target rare diseases (orphan drugs), G-BA assesses the therapy, and IQWiG only assesses patient numbers and costs. Based on this assessment, the G-BA decides on the level of additional benefit compared to an appropriate comparator therapy.
An additional benefit is defined as “a patient-relevant therapeutic effect in mortality (extension of survival), morbidity (shortening of the illness duration/improvement or delayed deterioration in the state of health), quality of life and/or safety/tolerability (reduction of side effects)” (IQVIA, internal research). There are six gradations to indicate the extent of additional benefit (Schulz *et al.*, 2020):

- Quantifiable additional benefit, categorised as
  - Major additional benefit: Sustainable and not-yet achieved large improvement of the therapy-relevant benefit (e.g. healed or considerable improvement in overall survival, long-term absence of serious symptoms or avoidance severe side effects)
  - Considerable added benefit: not-yet achieved significant improvements of therapy-relevant benefits (e.g. reduction of serious symptoms, moderate increase of overall survival time or relevant avoidance of (severe) side effects)
  - Minor added benefit: not-yet achieved moderate improvements of the therapy-relevant benefits (e.g. reduction of non-serious symptoms or avoidance of adverse effects)
- Non-quantifiable additional benefit: scientific data do not allow any quantification
- No additional benefit: no additional benefits are proved

- Less additional benefit than the comparative therapy: benefits of the therapy are less than benefits of the appropriate comparator

Schulz *et al*. (2020) note that German legislation states that EMA approval sufficiently indicates additional benefit for orphan therapies. Consequently, the latter two benefit categories ('no additional benefit' and 'less additional benefit') are not applicable in these cases. Moreover, the evidence requirements for orphan therapies are less stringent, as they do not have to be compared against a comparator. However, if the annual sales of an orphan therapy exceed €50 million, the manufacturer has to re-submit a full dossier, including an appropriate comparator.
Based on the extent of additional benefit in the AMNOG report, the reimbursement price is negotiated to be effective from the 13th month.

NICE

Rather than focusing on the relative effectiveness of a therapy, NICE focuses more on the cost-effectiveness of a drug.
The English HTA body uses one of three appraisal routes: single technology appraisals (a single therapy), multiple technology appraisals (several therapies used for one condition) or highly specialised technology (HST) appraisals (drugs for very rare conditions). Before the NICE review, the manufacturer submits an evidence report to an independent evidence review group. This evidence submission package includes data on cost-effectiveness, clinical efficacy and safety.
NICE evaluates the incremental cost-effectiveness ratio based on the estimated costs per QALY. QALYs reflect the state of health of an individual, expressed both in quality and length of life, so that one QALY is equal to one year of life in perfect health (NICE, 2021b). The number of QALYs is calculated by estimating the number of years of life remaining for a patient following a particular treatment, weighted by the utility value that is associated with a given health state (expressed on a scale from 0 (death) to 1 (full health)).
Based on the clinical efficacy and cost-effectiveness, NICE issues an appraisal recommendation:

- Recommended for use in line with marketing authorisation from EMA or in line with expected usage in clinical practice
- Recommended for a subset of patient populations
- Recommended for use in the cancer drug fund (applicable when there is uncertainty concerning clinical efficacy and cost-effectiveness of a cancer drug)
- Only in research
- Not recommended if the drug is ineffective or not cost-effective in comparison to current treatment practices

While there is no formal threshold below which a health technology is considered to be cost-effective, therapies with an incremental cost-effectiveness ratio above £30,000 per QALY are generally unlikely to receive a positive appraisal (IQVIA, internal research and NICE (2013)). However, the threshold for ultra-orphan diseases in highly specialised technology programs is more generous, varying from £100,000 to £300,000 (NICE, 2021a).

Joint clinical assessments

From the descriptions provided above, the inherent differences between HTA bodies become apparent. For the different HTA archetypes, value may constitute different aspects: for G-BA added value comprises the added benefit that a new therapy may bring, based scientific evidence that may or may not prove this benefit. For NICE, added value is quantified in the incremental cost-effectiveness ratio, based on the estimated costs per QALY. These inherent differences of value perceptions may illustrate

potential challenges in European collaboration efforts.
Yet, over the last years, EUnetHTA has laid a strong foundation for sustainable EU-wide cooperation by providing methodological frameworks and guidelines, as well as harmonised databases (Erdös *et al*., 2019).

In addition to the above, EunetHTA aims to "increase quality and efficiency of joint HTA work at the European level" (EunetHTA, 2021) through joint clinical assessments. Similar collaborative approaches to HTAs are explored by different EU-member states (BeNeLuxA, 2021).
Recently, the European Commission adopted a new regulation on HTAs that formalises a centralised, supranational approach (European Commission, 2021). The prospective EU-wide HTA process aims to harmonise the national HTA processes to generate a single, joint clinical assessment, focusing on REAs of innovative medicines and medical devices. While the idea may sound appealing, the legislation has been controversial. For one, member states have expressed their concerns on their sovereignty in the HTA process (Garattini and Padula, 2020). To this point, the legislation states that member states remain ultimately responsible for concluding the REA outcome. They may complement the outcome with additional analyses needed to assess added value in their national healthcare context. However, the European Federation of Pharmaceutical Industries and Associations (EFPIA) argues that if member states can decide if and how they use the joint work on a case-by-case basis, this could result in an arbitrary and unpredictable system (APM Health Europe, 2021).
Kanavos *et al*. (2019) describe the legislation as a step in the right direction. However, the authors note that further clarity and alignment on aspects like evidence requirements across therapeutic areas is needed to prevent unnecessary administrative and regulatory barriers.

Real-world data and real-world evidence

In 2017, Makady and colleagues gathered and reviewed publicly available definitions of RWD (used for the synthesis of RWE) to clarify the similarities and differences between them (Makady *et al*., 2017a). The authors identified 38 definitions of RWD and divided these into four categories:

1. Data collected in a non-randomised controlled trial setting
2. Data collected in a non-interventional/ non-controlled setting
3. Data collected in a non-experimental setting
4. Other (i.e., data that do not fit into the other three categories)

They found that the majority of RWD definitions fit the first category. In line with this finding, RWD is henceforth defined as an umbrella term for data collected outside the setting of randomised controlled trials (IMI GetReal, 2016). RWE is hereafter defined as the evidence derived from the analysis and/or synthesis of RWD (IMI GetReal, 2016).
The following illustrative example may be considered to distinguish RWD and RWE: whereas RWD could comprise 'raw' data (i.e. epidemiological data from patient registries to substantiate assumptions), RWE would comprise a retrospective analysis of such registries to draw conclusions that can be submitted as evidence.

RWD/RWE types

Both RWD and the derived evidence can be generated from multiple sources and study types. Makady *et al*. (2017a) found that registries, electronic health records and claims databases are most often mentioned as RWD sources in literature documents and interviews. The most cited study designs to derive RWE were found to be observational studies and pragmatic clinical trials (Makady *et al*., 2017a). However, there are multiple data sources available beyond the ones mentioned above (RWE Navigator, 2021b). An overview of RWD/RWE sources is provided in appendix A4.

Randomised controlled trials vs RWE

In randomised controlled trials, aspects of the studied health technology (i.e. efficacy or safety) are measured against a comparator in a homogeneous patient population (Katkade *et al*., 2018). These trials are performed under controlled and standardised conditions to minimise bias and potential confounders. As such, they have high internal validity (e.g. high confidence that any observed difference between patient groups can indeed be attributed to the health technology under investigation) (Sekaran and Bougie, 2016). Randomised controlled trials, therefore, remain the gold standard in evidence-based medicine.

However, the controlled (ideal) conditions do not necessarily represent the 'real-world' setting with heterogeneous patient populations with comorbidities and more complex care needs (Katkade *et al*., 2018). An inherent flaw in this study design is, therefore, the limited external validity (e.g. the extent of generalisability of the observed results to the general practice) (Sekaran and Bougie, 2016).

Especially in describing aspects such as patient characteristics, the burden of illness and existing treatment pathways, RWD can provide additional value (IQVIA, internal expertise). However, its role in HTA outcomes in the absence of a randomised controlled trial due to practical and ethical infeasibility is unclear.

In part, HTA bodies appear reluctant to wide-scale adoption of RWE due to its association with confounding bias and the lack of quality and transparency in data (Bowrin *et al*., 2019). In a similar vein, Murad *et al*. (2016) argues that in the evidence hierarchy (Figure 1.2), the straight lines that separate study designs should be changed to wavy lines to reflect the varying quality of evidence.

Another major hurdle impeding widespread use of RWD and RWE in HTAs is the lack of guidance for HTA bodies on what RWD and -evidence entails, how to appraise it and what it can be used for (Makady *et al*., 2017b). At the same time, lacking guidance from HTA bodies for the industry on the RWD quality requirements has also impeded adoption.

RWE guidance

The need for guidance appears to be increasingly recognised—several ongoing initiatives to streamline evidence generation and increase RWD adoption and evidence in HTAs.

The pan-European Innovative Medicines Initiative GetReal (IMI-GetReal) consortium, for example, has engaged key stakeholder groups, including academia, industry and HTA bodies, to define robust methods for RWD and evidence collection and interpretation (GetReal Institute, 2021). The output can serve as a framework for HTA bodies to interpret and appraise these evidence types.

In recent years, HTA bodies have also increasingly issued guidance and regulations on the usage of RWE in their decision-making process. The French HTA body Haute Autorité de santé, for example, has recently published methodological guidelines for industry, containing key recommendations for producing quality RWE studies (IQVIA, internal expertise). Similarly, NICE has provided a framework that guides the use of RWE to inform HTAs (NICE, 2021e). Moreover, NICE indicated in its five-year strategy that it aims to "become scientific leaders by driving the research agenda, using RWD to resolve gaps in knowledge and drive forward access to innovations for patients" (NICE, 2021c).

To mitigate evidence uncertainties of new health technologies at the time of decision making, the German G-BA launched its Gesetz für mehr Sicherheit in der Arzneimittelversorgung (GSAV) law in 2020. If G-BA considers the evidence on aspects like safety immature, this law obliges manufacturers to gather additional evidence on aspects like effectiveness in post-marketing studies (IQVIA, internal research).

Managed entry agreements

Guidance and regulations by national HTAs on the usage of RWE is often part of 'managed entry agreements'. While there are several types of such novel payment agreements, they all allow sharing risks between the manufacturers and the HTA bodies in case of evidence uncertainty. Reimbursement of a health technology may then be based on the actual clinical outcomes for patients, as derived from

post-marketing (real-world) evidence generation.
Considering the high price and uncertainty of gene therapy outcomes, these agreements may be particularly relevant for this therapy type. Indeed, Jørgensen *et al*. (2020) found that for gene therapies, outcome-based reimbursement models are increasingly adopted in the EU4 & UK.

2.2. Literature review

The objective of this literature review is twofold. The first objective is to inform the practice of HTA of gene therapies and gain insights into the methodological challenges encountered (track 1). In addition, literature could potentially explain RWD/RWE in gene therapy HTAs and possibly its role in these appraisals (track 2).

2.2.1. Search Strategy & Process

To gather relevant literature, the PubMed® registry was used in combination with the electronic abstract and citation database Scopus. Additional literature was retrieved through both forward and reverse referencing. Appendix A1 provides an overview of the search queries used.
Literature types comprised peer-reviewed articles, grey literature (i.e. reports, non-academic research) and book chapters. Documents should also be published in a language comprehensible to the author, i.e. English, Dutch or German. Literature was considered relevant if it discusses: HTA frameworks, HTA methodological challenges, real-world evidence usage in HTAs. No exclusion criteria on publication year were applied to capture a broad knowledge field. Most relevant literature was published in 2018 or later; the relative few documents published before were included. However, to optimise the number of relevant hits, exclusion criteria included focusing on specific diseases or geographic areas other than EU4, UK or US.
Once the search was completed, relevant papers were identified. The selection process involved four stages. First, the title and abstract of retrieved documents were screened for relevant content. Next, introductions and conclusions of selected documents were scanned. The final stage included screening the full text.

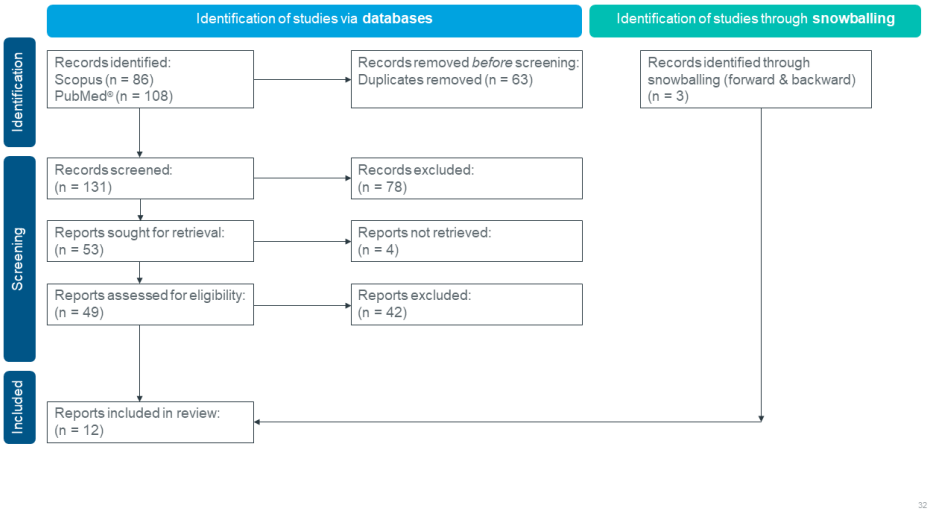


Figure 2.1. PRISMA-based overview of the search for relevant literature track one: current practices of gene therapy HTAs.

Track one: current practices of gene therapy HTAs

An initial search query using the keywords ((”health technology assessment*” OR HTA) AND (”Gene therap*” OR GTMP*)) yielded 38 results in Scopus. The search query was adjusted to capture a broader knowledge field in response to the limited output. To this end, “market access” was included as an alternative keyword for HTA. Moreover, since the previous section highlighted the similarities between cell and gene therapy types, findings for a similar therapy type were also considered relevant for this research. Multiple keywords referring to similar therapy types were therefore included, i.e. ”Cell therap*” OR ”Cell and gene therap*” OR ”CGT” OR ”Advanced Therapy Medicinal Product*” OR ATMP. The search yielded a total of 131 documents (Figure 2.1).

Based on the title and abstract screening, 78 records were excluded. Many of the excluded articles appeared to discuss challenges for cell & gene therapies in the trajectory towards market access but did not mention current practices. Of the 49 reports assessed for eligibility, nine were considered relevant. Similar to the excluded documents at the previous stage, many detailed aspects of market access were not considered relevant for this literature review. Other excluded documents discussed the feasibility of novel payment schemes or related aspects. While adopting payment schemes may be considered part of the HTA process, literature on what such schemes entail and their associated challenges is considered out of scope.
In addition, three documents were retrieved through a backward snowballing approach.

Track two: RWD/RWE usage in HTAs

To inform on real-world evidence usage in gene therapy HTAs, the aforementioned search query (track one) was initially supplemented with the keywords AND (”real-world evidence” OR RWE OR ”real-world data” OR RWD)). This search yielded 12 documents, hinting towards scarcity of literature on this aspect.
It should be noted, however, that there are multiple types of real-world evidence sources (Makady *et al*., 2017a). Therefore, documents that mention the use of a particular real-world evidence type in the considered context but do not refer to it as ’real-world evidence’ could be missed in this search strategy. Consequently, an additional search query was formulated to include different types of real-world evidence and real-world data sources (which are aligned with the search algorithm keywords as described in chapter 3.2). This query considerably increased the number of hits (from 6 to 25 in Scopus). However, the 11 documents that, based on title and abstract, would be considered for full-text analysis were already obtained in the other search queries described here—expanding the queries to include the terminology as mentioned above was therefore not considered necessary.

Following the scarcity mentioned above, subsequent search queries were reformulated to omit therapy-specific keywords and instead focus on broader literature, regardless of the therapy type of HTAs. The search query was limited to title and abstracts to optimise the relevance of documents. Combined, the two search engines retrieved 232 documents for this search (139 without duplicates) (Figure 2.2). From this increased output, it appears that the existing body of literature does not explicitly link real-world evidence usage in HTAs to gene therapies. This observation embodies a knowledge gap that the presented research aimed to solve: the lack of empirical evidence in RWD/RWE usage in gene therapy HTAs.

Based on the title and abstract screening, 83 out of the 139 documents were omitted. Irrelevant documents focused on specific diseases or had a geographic focus that did not include EU4, UK or US.
Of the 53 documents assessed for eligibility in full-text screenings, 16 were considered relevant. Excluded documents focused on the methodological aspects to realise the full potential of RWD/RWE, i.e. statistical analyses or multiple-criteria decision analysis. Such knowledge was considered out of scope for the presented literature review.
In addition, five documents were retrieved through a snowballing approach.

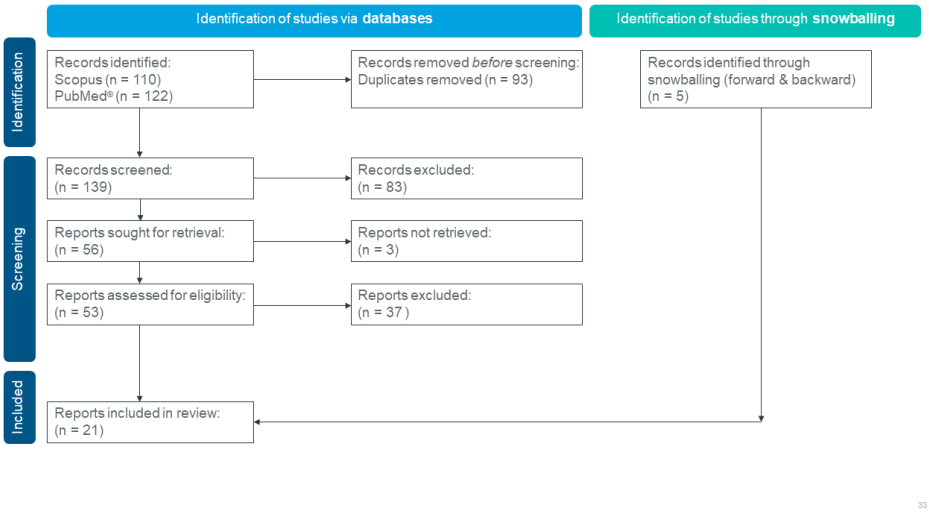


Figure 2.2. PRISMA-derived overview of the search for relevant literature track two: RWD/RWE usage in HTAs.

2.2.2. Current practices of curative therapy HTAs

The observed misalignment between novel curative therapies and established HTA methods is broadly covered in the literature. However, works that provide insights on the current practice of HTAs of such therapies appear to be scarce. Three documents were found to analyse HTA outcomes of curative therapies, albeit in varying levels of detail.

HTA outcomes and considerations differ per jurisdiction

The work by Ten Ham *et al*. (2021) presents the most detailed analysis. It is also the only one that utilised a methodological framework to categorise curative therapy HTA outcomes. Their work identifies and structures key considerations in the reimbursement recommendations in England, Scotland and the Netherlands.

In line with the earlier identified scarcity of literature, the authors observed a lack of empirical evidence on HTAs of ATMPs. They aimed to address it by reviewing the HTA outcomes of EMA-approved ATMPs. The authors found that reimbursement recommendations and underlying considerations differed across the studied jurisdictions for the same therapy. While this finding is in agreement with earlier described variation in HTA guidelines and practices (Kleijnen *et al*., 2012), most considerations appear to relate to the clinical effectiveness and cost-effectiveness. Not surprisingly, these are key uncertainties related to curative therapies, as will be discussed in the next section.

Similar to Ten Ham *et al*. (2021), Gozzo *et al*. (2021) found that the perceived added value of curative therapies by HTA bodies differed substantially. The authors compared HTA outcomes in Germany, France and Italy, specifically evaluating the view on the 'added value' that a novel therapy would add to existing healthcare practices. In doing so, the authors do not consider type of recommendation (positive, restricted or negative). This appears to be an important limitation, as HTA bodies may weigh the added value of therapy differently in their final recommendation (Kleijnen *et al*., 2012).

Indeed, following their finding that analysed HTAs agree on the added value of only two ATMPs, the authors acknowledge that the heterogeneity of HTA recommendations is likely related to varying HTA practices and acceptance of uncertainty (Gozzo *et al*., 2021). Consequently, the work states that it will be crucial to understand the causes of disagreement among the HTA bodies to increase patient access.

While te analysis presented by Gozzo *et al*. (2021) does provide initial insights into the alignment of HTA bodies, the underlying considerations are not considered at the level of detail presented by Ten Ham *et*

al. (2021). To illustrate: 'added value' as analysed by Gozzo *et al*. (2021), is merely one aspect of HTA outcomes and recommendations, and it is unclear how other factors may contribute to the outcome. To this point, the framework utilised by Ten Ham *et al*. (2021) may be particularly relevant. The authors categorised the identified considerations in pre-defined domains of the EUnetHTA core model®. As this framework allows for both the production and sharing of HTA information (EUnetHTA, 2015), it may very well contribute to the increased understanding of causes of disagreement among the HTA bodies.

Unlike the works mentioned above, Faulkner *et al*. (2019) quantify HTAs that address particular core dimensions per therapy type, rather than comparing HTAs on a jurisdiction level. The authors evaluated 100 HTAs of four technology types from five markets, including the UK and France. Their findings appear to suggest the presence of technology-specific challenges. Nevertheless, the authors conclude that HTA bodies may not be applying therapy-specific analysis frameworks. While this may be true, such frameworks could be missed by taking the aggregate of HTAs of five different countries, each having varying HTA guidelines and practices (Kleijnen *et al*., 2012). Similarly, if HTAs adapt over time to novel technologies, this would not be apparent from the presented aggregate analysis.
Moreover, the scope of this research is limited to therapeutics that gained regulatory approval in 2016 and 2017. Since 2017, at least six more cell & gene therapies have been commercially launched in Europe alone (IQVIA, internal research), the current situation may differ considerably.

Similarly, the tolerance for evidence uncertainty may vary between HTA bodies

In addition to the clinical effectiveness and cost-effectiveness considerations identified by Ten Ham *et al*. (2021), the authors emphasise that considerations relating to the ethical and legal aspects may bear substantial weight in the HTA outcome. The observation illustrates that when a therapy receives orphan designation, a higher degree of uncertainty is generally accepted in decision-making. Indeed, Pochopień *et al*. (2021) note that "[...] disease severity and unmet needs are important factors to be included in value assessments. This is reflected in the higher acceptability of uncertainty in clinical evidence and the cost-effectiveness analyses".
Similarly, Gozzo *et al*. (2021) note that the three ATMPs that are available in each of the analysed countries but not equally reimbursed are indicated for not life-threatening diseases or for diseases with other treatment options available. Lower quality of evidence may not be accepted in these cases. Then again, the tolerance for evidence uncertainty in different circumstances may vary across HTA bodies (Gozzo *et al*., 2021).

Appraisal of gene therapies appears to be inherently linked to uncertainties

Evidence uncertainty is a prominently mentioned factor in the broad literature base that analyses the misfit between novel curative therapies and established HTA methods. While the variety of identified challenges and proposed solutions appear to reflect the multifaceted nature of HTAs, most appear to relate to uncertainties regarding cost-effectiveness or clinical effectiveness due to evidence generation challenges.

Uncertainties in cost-effectiveness The challenges that Angelis *et al*. (2020) highlight, for example, mainly relate to cost-effectiveness. The authors categorise the main challenges encountered in assessing cell & gene therapies as cost estimation, benefit estimation or affordability. In accordance, the authors propose a specific set of adaptations per category to re-calibrate HTA frameworks. However, as these adaptations are not tailored to specific HTA bodies, they may not be equally applicable to different jurisdictions. For example, the authors propose capping the price at the maximum willingness per quality-adjusted life-year in determining efficiency. This is only relevant in countries where HTA bodies conduct cost-effectiveness analyses (IQVIA, internal research).

These cost-effectiveness challenges are similar to those described by Pochopień *et al*. (2021). However, rather than generalising problems relating to gene therapy HTAs, the authors argue that the challenges encountered in gene therapy HTAs should be put into perspective of both the severity of disease and the unmet need. To integrate these aspects, the work presents a framework from which there is no one challenge for gene therapies and their target diseases. Instead, the challenges lie on a spectrum that depends on the disease severity and the unmet need. However, whether the model holds with data from practice remains unclear.

In response to the cost-effectiveness challenges, the authors briefly discuss alternative funding mechanisms that share the financial risk related to such uncertainties. As these payment mechanisms evaluate the treatment outcome in the real-world clinical setting, they are, per definition, linked to real-world evidence generation. However, this evidence may not be available at the initial HTA recommendations.

Uncertainties in clinical effectiveness In addition to uncertainties relating to cost-effectiveness, Angelis *et al*. (2020) and Pochopień *et al*. (2021) also acknowledge the evidentiary uncertainties on the treatment effects in HTAs of ATMPs.
These challenges are more thoroughly discussed by Annemans and Makady (2020), who distinguish four main types of uncertainties that are inherent to treatments of rare diseases. These uncertainties relate to the population, the disease and its current management, the new treatment and the health ecosystem. Categorising them make up the first of three blocks of the methodological TRUST4RD tool to guide stakeholders in defining uncertainties and evidence gaps when assessing gene therapies. It should be noted that categories do not appear to be mutually exclusive. The authors do not specify how to handle HTA elements that fit multiple uncertainty types.
Interestingly, the authors emphasise the potential role of real-world evidence to reduce the identified evidence gaps while stressing the importance of collecting as much as possible data during the development phase of new treatments. However, methodological details on incorporating or appraising this kind of evidence in HTAs is lacking. Interestingly, out of the seven analysed documents that identify the challenges mentioned here, only two briefly mention the usage of real-world evidence as part of a potential solution. The focus appears to be more on emphasising the need for iterative dialogues. Moreover, involvement of multi-stakeholders to identify what uncertainties matter most and suggest using impact scoring. However, methods to derive such an impact score are not provided in Annemans and Makady (2020).

The essence of dialogues was also anticipated by Ronco *et al*. (2021). However, the authors acknowledge that they did not find evidence of such dialogues in their study on ATMPs in the EU5 countries. This is interesting since there are several support platforms and programs in both the US and Europe to facilitate such interactions (Overbeeke *et al*., 2021).
In a similar vein, the work by Coyle *et al*. (2020) notes that international initiatives such as the EUnetHTA and the EVIDENT database are essential in facilitating the adoption of real-world evidence. The authors take a more holistic approach to identifying adjustments in policy and assessment methodologies to improve gene therapies. The work advocates the inclusion of additional value elements into current frameworks and recognises the potential of real-world evidence in mitigating evidence uncertainty.

Managed access agreements may mitigate uncertainties

In part, the above-mentioned uncertainties are embraced in HTA reports by introducing managed access agreements. In a retrospective comparative analysis of recently launched gene therapies in the EU4, UK and US, Jørgensen *et al*. (2020) found that outcome-based reimbursement mechanisms have gained traction. However, in a follow-up study, the same authors found that such innovative payment mechanisms are more accepted in the studied EU countries than in the US (Jørgensen and Kefalas, 2021). Similarly, from literature and expert panels, Godman *et al*. (2021) conclude that despite appar-

ent disadvantages, further growth of managed access agreements is likely. The authors argue that the anticipated introduction of curative therapies with the inherent uncertainties will push the field towards adopting such payment schemes. Moreover, the foresee outcome-based reimbursements become increasingly feasible after adopting more sophisticated IT infrastructures across countries. Indeed, both Jørgensen *et al*. (2019) and Kefalas *et al*. (2018) found that having appropriate real-world data collection infrastructures is key in further facilitating these payment mechanisms.

2.2.3. Real-world evidence usage in HTAs

Interestingly, none of the articles that analyse and compare key considerations in HTAs of ATMPs, discuss the usage and value of real-world evidence. However, as noted in the previous section, papers that describe methodological challenges in HTAs of curative therapies occasionally discuss the potential of real-world evidence usage. However, none of them explicates the role that real-world evidence plays or could play in HTAs of curative therapies. Given that real-world evidence usage in HTAs not specific to ATMPs appears to be broadly covered in literature, the scope was expanded to capture real-world evidence usage in HTAs, regardless of the therapy type. General developments and considerations in this field were also relevant for gene therapies.

The apparent recognition of the potential value of real-world evidence for HTAs may not be reflected in the adoption

Five documents were analysed that explicate the role of real-world evidence in HTAs. Three of them provide empirical insights on the usage of real-world evidence.

The work by Makady *et al*. (2018), for example, examined the use of real-world data in melanoma HTAs using a retrospective, comparative analysis of HTA reports. From the analysis that the authors present, it appears that real-world data inclusion has not increased over time (Makady *et al*., 2018). Interestingly, Milliano (2019) reported opposite findings, noting that real-world evidence usage in oncology drugs HTA actually increased in four of the same jurisdictions from 2013 to 2018. However, it should be noted that the latter reports on general oncology drug HTAs, rather than melanoma HTAs specifically. Moreover, Makady *et al*. (2018) rightfully acknowledge that their conclusions should be taken cautiously, owing to differences in practices between agencies and varying numbers of reports published per year.

Other than the quantification of HTAs per year, Makady *et al*. (2018) also differentiated between relative effectiveness assessments and cost-effectiveness assessments in HTAs. They found that real-world data inclusion was 30% more common in cost-effectiveness assessments than relative effectiveness assessments. Real-world data mainly inform epidemiological information (i.e. prevalence/ incidence) in relative effectiveness assessments and long-term effectiveness and costs in cost-effectiveness assessments.

The work by Lee *et al*. (2021) reports similar findings for the cost-effectiveness assessments. The authors analysed real-world evidence usage in cost-effectiveness assessments by the American HTA body Institute for Clinical and Economic Review. Real-world evidence was found to mainly inform disease progression, health care resource utilisation or costs. Both Lee *et al*. (2021) and Makady *et al*. (2018) found that registry data was the most frequent source of data, followed by database data.

The work by Bullement *et al*. (2020) presents similar findings. The authors specifically considered the real-world evidence inclusion in cancer drug cost-effectiveness analyses by NICE. However, the authors state that inclusion is mainly related to patients' health-related quality of life. Real-world data inclusion informing epidemiological information is not mentioned. This is interesting as Makady *et al*. (2017b) found that HTA agencies generally do recommend the usage of real-world data in these cases.

Formal guidance on real-world evidence usage in HTAs appears insufficient to stimulate adoption

Indeed, guidance provided by regulatory bodies and HTA agencies is an often recurring aspect in the literature on real-world evidence usage in HTAs.

Makady *et al*. (2017b) reviewed HTA agencies' policies on the use of real-world data and found that the evidence requested by HTA agencies appears to vary with the context for which it is used. For example, while real-world data usage was accepted among all analysed HTA agencies for initial reimbursement decisions, it was not explicitly recommended. For parameters used in pharmacoeconomic analyses, however, agencies did specifically recommend using national real-world data sources. In such cases, real-world data may provide evidence on, i.e. epidemiological data (prevalence and incidence) or (relative) treatment effects.

As mentioned earlier in this literature review, policy considerations like orphan drug designations may increase the acceptance of less robust evidence due to the high unmet need (Ten Ham *et al*., 2021). Similar to the findings described here, agencies' acceptance of real-world data to provide evidence on treatment effects in such cases appears to vary, considering that while some agencies deem this acceptable, others explicitly advise against it (Makady *et al*., 2017b). The same authors note that policies may prominently feature the hierarchies of evidence that agencies use to classify evidence quality. However, they raise the question of whether practices derived from evidence-based medicine are still applicable to real-world data usage for HTAs. They argue that such hierarchies tend to downgrade real-world data without differentiating the type of data that randomised clinical trials (e.g. efficacy data with high internal validity) and different forms of real-world data (long-term effectiveness data with high external validity) can provide.

That is not to say that HTA bodies do not appreciate the concept of real-world data and evidence. On the contrary, it appears from the systematic literature review by Bowrin *et al*. (2019) that most HTA bodies do recognise both the benefits and limitations of real-world evidence usage. However, in line with the above, the authors found that formal guidance on leveraging it in cost-effectiveness modelling was scarce.

Similarly, Kent *et al*. (2021) agree that to ensure the generation of high-quality evidence suitable for decision making, HTA bodies should issue clear guidance on data quality standards and best practice methods. Several initiatives exist to establish frameworks for the use of real-world evidence in decision-making and stimulate the adoption of real-world evidence in HTAs (Annemans and Makady, 2020; Facey *et al*., 2020; Alliance for Regenerative Medicine, 2019). To this point, the authors note that to ensure adoption and streamline evidence generation, the output of these initiatives should be developed collaboratively. They argue that these outputs could also include clear guidance on when non-randomised studies can be considered; this would include real-world data. A similar point is brought up by Makady *et al*. (2017b), who highlight the need for policy alignment of HTA agencies within Europe on real-world data usage in HTAs and provide guidance on practical aspects of its collection and analysis. The authors point out that a "harmonised set of policies on real-world use for HTA would provide market authorisation holders with the ability to plan alternative evidence generation pathways which rely less on randomised controlled trials, and more on real-world studies; the latter theoretically yielding outcomes more relevant for HTA purposes".
Similar to the varying degree to which HTA bodies provide guidance on real-world evidence usage, their attitude towards the value of real-world evidence also appears to vary.

Attitudes towards real-world evidence usage in HTAs appear to vary across different agencies

Three documents were found to cover the attitudes towards real-world evidence usage in HTAs and the consequent role in the assessment.

Sievers *et al*. (2021) aim to explicate this in semi-structured interviews with industry experts and HTA bodies. The authors suggest that the evidence requirements of the different stakeholders' conflict. Whereas regulators may demand real-world evidence to support long-term safety evaluations, HTA bodies generally require evidence that allows for comparative assessments. According to the interviewees, real-world evidence currently appears to satisfy only the former, as most HTA bodies still prefer evidence from randomised controlled trials. While these findings should be placed into context, considering that the HTA bodies interviewed have historically been more on the fence regarding real-world evidence usage in HTAs (IQVIA, internal research), this finding appears to be broadly supported in the analysed literature.
Indeed, Makady *et al*. (2017b) found that the big EU4 HTA agencies adopt similar hierarchies of evidence in accordance with principles of evidence-based medicines. Consequently, agencies unanimously place real-world data sources on a lower level of quality and reliability than randomised controlled trials. In a similar vein, Katkade *et al*. (2018) note that RWD/ RWE have "the potential to support, improve, and potentially accelerate the delivery of safe and cost-effective therapeutic interventions".
As such, RWD/RWE usage in HTAs appears to mainly serve as a supplement to randomised controlled trials. However, literature explicating whether this also holds for gene therapies and empirical evidence on the extent to which RWE/RWD plays a role in the final HTA outcome is lacking.

The results presented by Vreman *et al*. (2019) show that negative HTA recommendations or (economic) restrictions do not apply more often for conditionally approved drugs without controlled evidence. This implies that in HTA, the use of uncontrolled studies is not a decisive factor to come to a negative or a restricted recommendation. Nevertheless, Makady *et al*. (2018) note that real-world data usage for effectiveness is more likely to be negatively appraised in relative effectiveness assessment HTAs. However, more explicit barriers to the adoption and appraisal of real-world evidence have been mentioned in the literature.

The barriers impeding real-world evidence usage in HTAs appear to mainly relate to the challenges of evidence generation

Seven documents that present barriers to adopting real-world evidence were analysed. Half of the documents identify methodological challenges.

Methodological challenges A flaw that each of these documents acknowledges is the potential bias of real-world evidence. Indeed, in a systematic literature review of 14 articles, Bowrin *et al*. (2019) identified confounding bias as the main limitation in real-world evidence usage. In a similar vein, in half of the 11 semi-structured interviews conducted by Sievers *et al*. (2021), selection bias was mentioned as a challenge.

Another prominently mentioned methodological challenge that is apparent from both above-mentioned works is the lacking randomisation (Bowrin *et al*., 2019; Sievers *et al*., 2021). This should be expected, considering the earlier identified preference of HTA bodies for evidence derived from randomised clinical trials. Along the same line, Roberts and Ferguson (2021) report that poor internal validity is a limiting factor.

Lack of quality The authors further report that transparency is a key barrier in adopting real-world data in HTAs. Similarly, the quality of real-world data was perceived as a major challenge by five interviewees in Sievers *et al*. (2021). The authors note that there are often no quality control infrastructures in place to ensure the completeness of data collection. This point is also raised by Simpson and Ramagopalan (2021), who describe the concerns of the German HTA body IQWiG on real-world data sources. With the exception of high-quality registries, IQWiG berates the insufficient data quality and completeness of real-world data sources.

Lack of standardisation However, it should be noted that the extent to which data is considered insufficiently complete or low quality may vary across data sources. To this point, Sievers *et al*. (2021) note that there is a lack of standardisation in real-world data collection, leading to differences between countries, regions, and hospitals. Moreover, as noted earlier in this review, the degree to which HTA bodies accept such sub-optimal evidence sources may vary. Indeed, Kanavos *et al*. (2019) note that a key challenge from an HTA perspective remains the variable acceptance of real-world evidence. The authors note that despite infrastructure improvements, issues remain with access to data in several jurisdictions, including privacy issues, the lack of incentives for data sharing, availability and use, and the ongoing debate about real-world evidence distrust.
Consequently, the work stresses the necessity to discuss what needs real-world evidence will fulfil. Such discussions should span across borders according to Facey *et al*. (2020). The same authors state that jurisdictions should agree on real-world data requirements and the associated infrastructure, development of data analytics methods for HTA, and transparency in real-world evidence studies.

Lack of clarity In line with the above, the lack of clarity among stakeholders appears to impede the adoption of real-world evidence usage in HTAs. Facey *et al*. (2020) did case studies with policymakers and HTA bodies and identified a lack of clarity about the Payer/HTA questions that could be answered by real-world data and on how its quality could be assessed.
Similarly, Roberts and Ferguson (2021) found that lack of training on how to evaluate observational studies was a fundamental challenge to adopting real-world evidence. Interestingly, while a lack of guidance from payers to industry players was earlier identified in this review, Roberts and Ferguson (2021) note that there is also no guidance available to payers themselves. The authors argue that the complexity of evidence available to the payer has increased exponentially with data available from different clinical trials and real-world settings. Payers may not know how to appreciate and interpret the entire body of evidence available fully. This appears to necessitate the need for advanced tools to analyse the high volumes of data efficiently, according to the authors (Roberts and Ferguson, 2021).

Through focus groups, interviews and surveys, Malone *et al*. (2018) aimed to map the perceptions and acceptance of real-world evidence among US payers. The authors found that many participants indicated that a lack of experience conducting their own analyses and interpreting those of other HTA bodies formed a considerable barrier to adopting real-world evidence. Nevertheless, the work suggests that HTA bodies are open to and interested in evaluating observational studies if given the proper guidance or tools. This reiterates the necessity of earlier mentioned initiatives to educate stakeholders, establish frameworks and guidance and facilitate the adoption of real-world evidence.

2.3. The literature review laid a foundation for further research

The presented literature review has covered several relevant aspects of this research (Table 2.1). From the documents gathered in track 1, it appears that there is a consensus that the unprecedented potential value of gene therapies challenge established HTA methods.
From the literature, it is clear that the considered therapies mainly concern orphan diseases, which generally allow for less stringent evidence requirements. The associated challenges that are commonly encountered mainly relate to cost-effectiveness and the uncertainty regarding long-term safety and -effects and the appropriate evidence generation. While the recommendations for adaptations to tackle these challenges appear to be abundant, empirical evidence on how curative therapies are assessed in practice was found to be scarce.
Track 1, therefore, laid a foundation for the first sub-question of this research (*How are gene therapies currently assessed in HTAs?*), but more empirical data is considered necessary.

Similarly, track 2 laid a strong foundation for the second research sub-questions (*What role does*

RWD/RWE play in the HTA appraisal process of gene therapies? and *How has RWD/RWE usage in gene therapy HTAs evolved?*). From the retrieved literature, it appears that the potential of real-world evidence is generally recognised but only sporadically linked to curative therapies. RWD/RWE mainly serves a supportive role in HTAs, specifically delivering evidence on epidemiological information (i.e. prevalence/ incidence) in relative effectiveness assessments and long-term effectiveness and costs in cost-effectiveness assessments. Moreover, it is clear from the studied literature that HTA bodies prefer evidence from randomised controlled trials over real-world evidence. Whether RWD/RWE usage in HTAs has increased over time is unclear from the literature. It is clear, however, that the potential is increasingly recognised.
While the literature informs sub-question 2, additional empirical data from the retrospective comparative analysis will be key to formulate an answer.

Track 2 also informed a substantial part of sub-question 3 (*What factors impact RWE usage in gene therapy HTA appraisals?* and *The steps to be taken to extend RWE usage in HTA appraisals of gene therapies?*).
The literature prominently mentions the methodological factors impeding widespread RWD/RWE adoption. These include confounding bias, lacking randomisation and a lack of transparency. Other considerable barriers appear to be formed by the lack of guidance & policies and a lack of standardisation among HTA bodies and health technology manufacturers.
Steps to be taken are mainly related to the factors impeding RWE/RWD usage. The retrieved literature consistently mentions stakeholder discussion and alignment to tackle methodological challenges and misalignment. This would include consensus on RWD/RWE terminology and mitigating methodological challenges. Similarly, in response to the lack of consistent guidance across agencies on RWE/RWD usage, the need for guidance for both payers and pharmaceutical companies is suggested. Finally, the introduction of data collection standardisation is hypothesised to ensure high-quality data to facilitate adoption.
Again, it should be noted that the retrieved literature mainly explicates these aspects for RWD/RWE usage in HTAs not specific to gene therapies. Validating these findings in interviews will therefore be necessary.

From the presented literature review, an ample knowledge gap was identified. While initial findings inform specific aspects of the formulated sub-questions, they are rarely explicitly stated to apply to gene therapies. Moreover, the existing literature fails to specify what role RWD/RWE currently plays and could potentially play in the future in gene therapy HTAs.
Understanding this aspect is deemed essential in the recently accepted joint clinical assessment legislation. Empirical data will be precious to provide insights into the potential feasibility challenges of this legislation.

Table 2.1. Key findings from the literature review.CEA: cost-effectiveness assessment. HTA: Health technology assessment. REA: Relative effectiveness assessment. RWD: Real-world data. RWE: Real-world evidence.

Sub-question	Preliminary finding	Description	Reference
1	1	Gene therapies mainly concern orphan diseases, which generally allow for less stringent evidence requirements	Section 2.2.2

continues on next page

	2	The challenges associated with gene therapies mainly relate to cost-effectiveness and the uncertainty regarding long-term clinical effectiveness due to challenges in appropriate evidence generation	Section 2.2.2
2	3	It appears that RWD/RWE mainly serve a supportive role in HTAs, specifically delivering evidence on epidemiological information in REAs and long-term effectiveness and costs in CEAs	Section 2.2.3
	4	HTA bodies prefer evidence from randomised controlled trials over real-world evidence.	Section 2.2.3
3	5	Methodological factors impeding widespread RWD/RWE adoption include confounding bias, lacking randomisation and a lack of transparency	section 2.2.3
	6	Other considerable barriers appear to be formed by the lack of guidance & policies and a lack of standardisation among HTA bodies and health technology manufacturers	section 2.2.3
	7	There is a need for guidance for both payers and pharmaceutical companies on RWD/RWE usage	section 2.2.3
	8	Stakeholder alignment, i.e. consensus on RWD/RWE terminology and mitigating methodological challenges, are key steps in facilitating its increased adoption	section 2.2.3
	9	The introduction of data collection standardisation will ensure high-quality data needed to facilitate increased adoption of RWD/RWE	Section 2.2.3

3

Methodology

This chapter will elaborate on the research methodology used to answer the earlier stated research questions. First, the research design is presented, followed by a description of the data collection and validation approaches.

3.1. Research design

The research design (Figure 3.1) comprised three stages: data gathering, analysis, and validation. It combined two approaches: desk research (e.g. literature review) and a comparative retrospective analysis of published gene therapy HTAs.

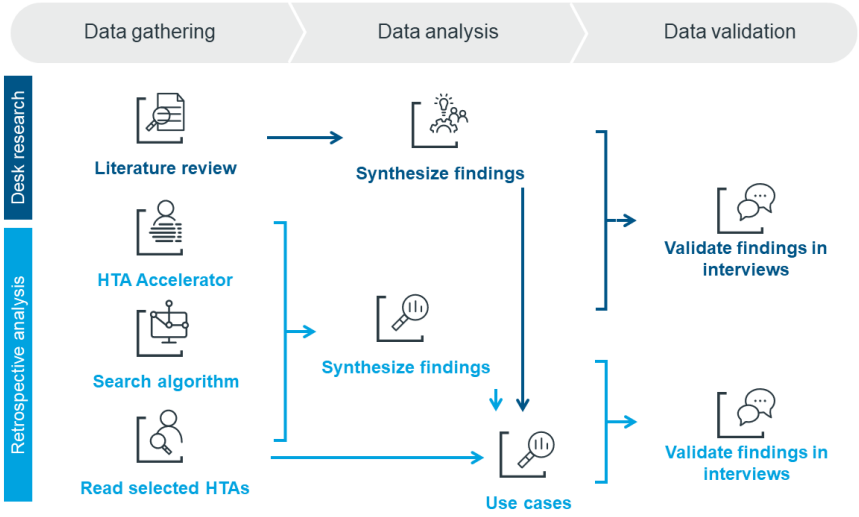


Figure 3.1. Research design for the presented research. HTA: Health technology assessment.

The literature review laid a foundation for the research questions in the data gathering and analysis stage. Findings from the retrospective analysis and three use-cases were synthesised that were validated in semi-structured interviews.

3.2. Data gathering

Literature review

A critical literature review was used to position this research relative to the existing literature body and inform the development of the research questions. This literature review was presented in the previous chapter. It provided the research with the most recent literature and developments on gene therapy HTAs and the usage of real-world evidence and data in HTAs.
Reviewing the literature involved selecting, analysing and synthesising relevant literature to identify related work and methods (Sekaran and Bougie, 2016).

While the presented literature review should not be considered a systematic literature review, it aimed to be extensive. To this end, the search strategies followed a similar approach to the 'Preferred Reporting

Items for Systematic Reviews and Meta-Analyses' (PRISMA) statement (Moher *et al.*, 2009). However, the quantification of excluded articles with substantiating reasons was left out of scope.

Retrospective analysis

All gene therapies that received centralised marketing authorisation by the European Medicines Agency until July 1, 2021, were included in the retrospective comparative analysis. (Figure 3.1). Eligible gene therapies were identified from the American Society of Gene & Cell Therapy Q2 quarterly data report (ASGCT and Pharma Intelligence, 2021). Some identified therapies are not strictly a gene therapy but a combined cell & gene combination therapy (Strimvelis®, Kymriah®, Yescarta®, Tecartus®and Libmeldy®). However, as the EMA did designate these therapies as gene therapies in their corresponding European public assessment reports, these therapies were taken along in the analysis.

Table 3.1. Gene therapies in Europe and their approval details. Cut-off date July 1, 2021. ADA-SCID - Adenosine deaminase-severe combined immunodeficiency. ALL: Acute Lymphoblastic Leukaemia. DLBCL: Diffuse Large B-cell Lymphoma. PMBCL: Primary mediastinal B-cell lymphoma. SMA: Spinal muscular Atrophy.

Product	Generic name	Manufacturer	Indication	Orphan designation	Market authorisation
Imlygic®	Talimogene laherparepvec	Amgen	Metastatic melanoma	✗	December 2015
Strimvelis®	Autologous CD34+ cells transduced with a lentiviral vector containing the human ADA gene	Orchard Therapeutics	ADA - SCID	✓	June 2016
Kymriah®	Tisagenlecleucel	Novartis	DLBCL	✓	August 2018
			ALL	✓	August 2018
Yescarta®	Axicabtagene ciloleucel	Kite Pharma (Gilead)	DLBCL & PMBCL	✓	August 2018
Luxturna®	Voretigene neparvovec	Spark Therapeutics (Roche)	Leber's congenital amaurosis	✓	November 2018
Zynteglo®	Betibeglogene autotemcel	bluebird bio	Transfusion-dependent beta thalassemia	✓	May 2019
Zolgensma®	Onasemnogene abeparvovec	Novartis	SMA	✓	May 2020
Tecartus®	Brexucabtagene autoleucel	Kite Pharma (Gilead)	Mantel cell lymphoma	✓	December 2020
OTL-200/ Libmeldy®	Autologous CD34+ cells encoding ARSA gene	Orchard Therapeutics	Metachromatic leukodystrophy	✓	December 2020

Two European HTA agencies were selected for the retrospective comparative analysis: NICE (England) & G-BA (Germany). The HTA framework adopted by these two agencies is generally considered prime examples of relative effectiveness assessments and cost-effectiveness assessments, respectively (IQVIA, internal research). Consequently, they represent two ends on an evidence requirement spectrum, wherein other HTA bodies and archetypes can be placed. As such, a combined analysis of HTA reports published by NICE & G-BA is assumed to produce findings largely applicable to other HTA bodies.

HTA reports of the identified gene therapies were retrieved from the HTA agency websites by searching

for the products brand and generic name. If products were authorised for multiple indications, reports for each indication were included.
Only completed HTAs were included; HTAs that were suspended or ongoing were excluded. For NICE, the committee papers & final appraisal documents were used. While the committee papers explicate the evidence submitted, the final appraisal document details how specific evidence was valued. For G-BA the document 'Nutzenbewertung G-BA' (benefit assessment) and 'Tragende Gründe zum Beschluss' (Justification) & 'Beschlusstext' (Resolution) were used. The former summarises relevant evidence, and whether the evidence was accepted, the latter two documents explicate the appraisal of the evidence. In addition to the above, the document 'Modul 4 - Dossier zur Nutzenbewertung' (dossier for benefits assessment) was used for the G-BA case studies.

Data extraction

Contrary to existing literature that delivers empirical evidence on current curative therapy HTA practices (i.e. Makady *et al*. (2018) and Ten Ham *et al*. (2021)), no second author was available to validate data extraction.
A multifaceted approach was therefore used to explicate the usage of real-world evidence & -data in the identified HTA reports (Figure 3.2). By triangulating the retrieved data, higher confidence in the results was obtained (Sekaran and Bougie, 2016).

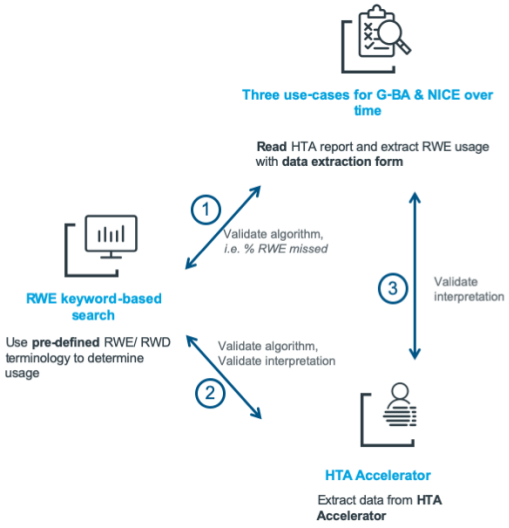


Figure 3.2. Data extraction methodology. The numbered arrows refer to validation approaches. HTA: Health technology assessment. RWD: Real-world data. RWE: Real-world evidence.

HTA Accelerator Initial data extraction was performed using IQVIA's proprietary HTA Accelerator software.
IQVIA's HTA Accelerator is an online platform that tracks publicly available HTA records (IQVIA, 2021). The data is collected from payer assessments and regulatory approvals, clinical trials and price information. The platform provides insights from HTA reports of more than 100 agencies in 41 countries across 250 primary diseases. To prevent misinterpretation or loss of knowledge due to language barriers, HTA summaries are translated into English by local experts and native speakers.
Pre-defined parameters allow for data extraction on various parameters, including real-world evidence usage. Within this parameter, the following information is captured: the name of the real-world evidence source, the type of real-world evidence, what area it supports, whether the HTA body accepted the evidence, and the rationale and additional details.

However, initial data extractions with the HTA Accelerator appeared to insufficiently cover real-world evidence/ real-world data usage, especially in older reports. Therefore, an algorithm was developed to search for real-world evidence terminology in the HTA reports and complement the output from the HTA Accelerator.

Search algorithm Individual HTA reports were searched for pre-defined real-world evidence/ real-world data terminology using a search algorithm.
RWD/RWE terminology for the search algorithm was obtained through publicly accessible databases and glossaries (RWE Navigator, 2021a; IMI GetReal, 2016; National Health Council, 2021). A list was compiled and complemented with words derived from IQVIA internal expertise. Keywords were adjusted to minimise off-target hits (Appendix A4).

HTA reports were retrieved in portable document format (pdf) and converted to text files. The individual text files were then fed into the algorithm, which extracted all occurrences of the pre-defined keywords, including the ±50 surrounding words (Appendix A3).
The retrieved 101-word text extracts were then assessed for relevance. If the relevance could not be derived, the paragraph containing that text extract was read in the original HTA report. If still in doubt, the original reference was retrieved for review. If the reference was not accessible, the study was not included. Sub-studies were merged: i.e. if three studies refer to a single study (i.e. Melody), then only 'Melody' was used. Similar to Lee *et al*. (2021), RWE studies were counted as many times as it was used to support different areas in the HTA.
Extracted data included the name of the evidence source used and its type, the area supported and whether it was accepted. Acceptance was categorised as 'yes' (following an explicit statement that the RWE was accepted/ included), 'no' (following an explicit statement that the RWE was not accepted/ excluded) or 'not identified' (if neither a positive nor negative statement regarding the RWE inclusion was identified).

In addition to the search algorithm, HTA reports of three therapies (Imlygic®, Yescarta®and Zolgensma®) were studied in more detail. For this purpose, a data extraction form was developed.

Data extraction form A data extraction form was adapted from the studies by Makady *et al*. (2018) and Lee *et al*. (2021) to allow for a systematic approach to data extraction. Similar to Lee *et al*. (2021), the extraction form included a (1) general information section (e.g. title of HTA report, report number, product name and date of publication), (2) a section on RWE/ RWD characteristics (e.g. name of evidence source used and its type), (3) a section on what area the evidence supported, and (4) an appraisal section (Appendix A5).
Based on IQVIA internal expertise, categorisation of areas supported was done according to the PICO (population, intervention, comparator and outcome) framework (Richardson *et al*., 1995).

3.3. Data analysis

The output from the different data extraction methods was compiled and analysed.
Extracted RWD/RWE usage was cross-compared between the two HTA reports for the same gene therapy. The motivation for this comparison was twofold. First, comparing RWD/RWE usage allows for synthesising preliminary findings on differences in usage or appraisal of certain RWD/RWE types. Moreover, comparing allows to complement RWD/RWE usage initially missed by the HTA Accelerator and the search algorithm (i.e. if a study was not mentioned close to a pre-defined keyword).

Illustrative use cases

Three illustrative use-cases provided context to the identified RWD/RWE usage. Three therapies were chosen based on the average HTA publication date. The (on average) earliest published HTA report (Imlygic®), middle (Yescarta®) and latest published report (Zolgensma®) were assumed to be representative for the time that they were published. As such, the aggregate findings for these three use cases may also provide insights in how RWD/RWE usage has evolved over time and what relevant considerations were in the HTA outcomes. A complete overview on identified RWD/RWE usage is provided in Appendix A8.

3.4. Data validation

In the presented research, interviews were used to verify the validity of the preliminary findings from the literature review and retrospective analysis.

Interviews

Interviewees (Table 3.2) were recruited via email. Industry experts had extensive experience in RWD/RWE usage in HTAs, albeit not specifically for gene therapy HTAs. However, they were well-informed of recent developments related to gene therapy HTAs (and noted that many factors and considerations overlap with other therapy types). Industry expert inputs provided a perspective that was to a degree representative of both manufacturers and HTA bodies.
Interviewed academics had published multiple relevant articles on cell & gene therapy HTAs. Their input provided a perspective on gene therapy HTAs specifically, the current state of literature and ongoing initiatives (driven by academia).

Prior to the interview, interviewees signed an informed consent form. The template for this form is included in a separate appendix that can be provided upon request.
Interviews were held in a semi-structured format, following an interview protocol (Appendix A2). Compared to structured interviews, semi-structured interviews allow for more leeway for following up on perspectives, ideas and topics raised by the interviewee. This enables the interviewer to make better use of the knowledge-producing potentials of dialogues (Leavy, 2014).
Interviews were held via Microsoft Teams. The interview was recorded using the Microsoft Teams functionality with the interviewee's consent. Afterwards, the interviews were transcribed manually in a light edited form. Expressions such as 'uh' or 'hmm', pauses and repetition of words were omitted. The anonymised transcripts are included in a separate appendix and can be provided upon request.

Table 3.2. Overview of interview participants. HTA: Health technology assessment. RWD: Real-world data. RWE: Real-world evidence.

ID	Description	Expertise
I1	Industry expert	Germany: HTA and RWD/RWE
I2	Industry expert	RWD/RWE usage in HTAs across Europe
I3	Industry expert	Health economics
A1	Academia	Health economics
A2	Academia	HTAs of cell & gene therapies

The obtained interview transcripts were then analysed, a process that generally involves data reduction, data display, and the drawing of conclusions (Miles and Huberman, 1994)
Data reduction comprises selecting, coding and categorising the interview data (Sekaran and Bougie, 2016). According to the same authors, coding refers to the analytical process of reducing, arranging and integrating qualitative data to conclude. The purpose is to help draw meaningful conclusions about the data.

Coding involves labelling units of text to group them into categories later (Sekaran and Bougie, 2016). This activity was performed using the computer-assisted qualitative data analysis software ATLAS.ti 22 (ATLAS.ti Scientific Software Development GmbH., Berlin, Germany).
Interview transcripts were imported to ATLAS.ti and read line-by-line. Text segments deemed relevant or interesting by the researcher were assigned codes. After this open coding phase, the number

of codes was reduced through axial coding, where codes were merged and overarching categories formulated (Flick *et al.*, 2004). In the final phase, selective coding was applied. This allowed the researcher to select and integrate the organised data from axial coding in a cohesive manner to derive findings.
It should be noted that while the coding process is described sequentially, the overall approach was non-linear. Between stages, the appropriateness of codes and categories was revised and re-applied to the collected data in order for the theory to evolve.

In the subsequent data display step, the aggregate findings were visualised in a network. This helped the researcher to understand and interpret relationships in the obtained, reduced data.
Insights generated in the data reduction and data display stage were interpreted in the final analytic stage to draw conclusions and compare findings.

Validating data extraction methodologies

Two aspects of the data extraction methodology were validated (Figure 3.2); completeness of extracted data using the algorithm and interpretation of identified RWD/RWE appraisals. The outcomes are presented in Appendix A6.
To derive the completeness of extracted data using the algorithm, the output was first compared to the HTA Accelerator data (Figure 3.2; arrow 2). However, as pointed out before, data in the HTA Accelerator may not always sufficiently cover RWD/RWE data in a report. For three indications (Imlygic®, Yescarta®and Zolgensma®), data from the search algorithm was therefore also compared to the data extracted from reading the HTA reports for the use-cases (Figure 3.2; arrow 1). Two assumptions are made here. The first assumption is that reading the report is the most thorough and reliable approach to extract RWD/RWE and can therefore be considered the golden. By comparing the search algorithm output to the golden standard, an impression is obtained that the percentage of RWD/RWE usage is missed by using the search algorithm alone. The second assumption is that the sample used in this validation is representative of the other reports.
A similar approach is used to validate the interpretation of identified RWD/RWE appraisals from use cases (Figure 3.2; arrow 3). However, in this case, the HTA Accelerator was considered the golden standard, as experts have interpreted the reports with more knowledge than the researcher.

4 Results

4.1. Orphan designation benefits gene therapy HTA outcomes

From the literature review, a lack of empirical evidence on gene therapy HTAs became apparent. Similarly, literature did not suffice to clearly explicate how gene therapies are currently assessed. As such, the first step in performing this research was to gain insights into the recent developments in gene therapy HTAs to answer sub-question 1 (*How are gene therapies currently assessed in HTAs?*).

Table 4.1. Identified gene therapy HTAs and initial reimbursement recommendations of in-scope HTA bodies. Cut-off November, 2021. Initial reimbursement recommendations: Negative recommendations (orange), Restricted recommendations (light green) and Positive recommendations (green). '-': No HTA-report identified as of July 2021. AMNOG: Arzneimittelmarkt-Neuordnungsgesetz (translates to pharmaceuticals market reorganisation act). CDF: Cancer drug frund. EMA: Europen Medicines Agency. G-BA: Gemeinsamer Bundesausschuss. HST: Highly specialised technology. NICE: National Institute for Health and Care Excellence. STA: Standard technology appraisal. *Table layout modified from Ten Ham et al. (2021).*

Product	Germany (G-BA)			England (NICE)		
	Program	Date	Outcome	Program	Date	Outcome
Imlygic®	AMNOG	December 2016	No added benefit	STA	September 2016	Recommended for restricted population with discount
Strimvelis®	-	-	-	HST	February 2018	Recommended
Kymriah®- DLBCL	AMNOG	March 2019	Non-quantifiable added benefit	STA	March 2019	Recommended with managed access through CDF
Kymriah®- ALL	AMNOG	March 2019	Non-quantifiable added benefit	STA	December 2018	Recommended with managed access through CDF
Yescarta®	AMNOG	May 2019	Non-quantifiable added benefit	STA	January 2019	Recommended with managed access through CDF
Luxturna®	AMNOG	October 2019	Considerable added benefit	HST	October 2019	Recommended
Zynteglo®	AMNOG	May 2020	Non-quantifiable added benefit	STA	*Submission date October 2019*	Suspended but initial documents available
Zolgensma®	AMNOG*	November 2021	No added benefit	HST	July 2021	Recommended for restricted population (beyond EMA label)
Tecartus®	AMNOG	August 2021	Non-quantifiable additional benefit	STA	February 2021	Recommended with managed access through CDF
Libmeldy®	AMNOG	November 2021	Considerable additional benefit	HST	*Submission date February 2020*	In progress but initial documents available

Nineteen HTA reports published by G-BA and NICE were identified for the ten in-scope gene therapies (Table 4.1). Two of these HTA reports were still ongoing or suspended (NICE: Libmeldy®(submission

date October 2019) and Zynteglo®(submission date February 2020), respectively). NICE did publish initial documents that allowed for deriving RWD/RWE usage in these HTA submissions but not the RWD/RWE appraisal. Therefore, these reports were included for the total RWD/RWE usage but excluded in the appraisal comparison between NICE and G-BA.

Nine of the identified HTA reports issued a positive recommendation (G-BA, n = 7; NICE n = 2), six a restricted recommendation (NICE, n = 6), two a negative recommendation (G-BA, n = 2) and two no recommendation (NICE, n = 2) (Table 4.1). It should be noted, however, that the G-BA and NICE frameworks and implications of orphan designation differs. As such, it may not be appropriate to strictly compare gene therapy HTA outcomes as being positive, restricted or negative.

G-BA did not use any special programs to assess the gene therapies. In line with German regulation, all therapies that have an orphan designation and annual sale of below €50 million, received a positive recommendation (e.g. added benefit) by G-BA. Indeed, Imlygic®does not have an orphan designation and annual sales of Zolgensma®exceeded €50 million, thereby taking away its orphan designation privileges in a full assessment (Schulz *et al*., 2020).

At NICE, four therapies were assessed via the highly specialised technology route, which allows for a higher cost-effectiveness threshold. The two indications that received a positive recommendation (Strimvelis®; Luxturna®) were assessed via this route.

Table 4.2. Considerations in gene therapy HTA reports. Data retrieved from HTA Accelerator, complemented by data from Jørgensen and Kefalas (2021). Not exhaustive. Considerations are not mutually exclusive.

Consideration	Germany (G-BA)	England (NICE)
Key considerations in decision making (not exhaustive)		
Lack of long-term data on efficacy	2	2
Lack of long-term data on safety	2	-
Lack of long-term data on all patient-relevant endpoints	1	-
Lack of long-term follow-up	-	3
Uncertainty in overall survival	1	-
Uncertainty in long-term benefits	-	1
Uncertainty in cost-effectiveness	-	-
Special considerations		
Orphan drug	8	1
Burden of illness	-	2
End-of-life	-	3
Market access considerations		
Discount applied	-	6
Outcome-based agreement	4	2
Temporary decision	6	-
Continued evidence development agreement	4	3

Orphan designation was specifically mentioned in nine HTA reports (G-BA, n = 8; NICE n = 1) (Table

4.2). This difference is expected, as this designation only influences the G-BA appraisal process. Orphan designation per se does not lead to a special program in the NICE HTA framework. However, it is one of the requirements for the highly specialised technology trajectory that four gene therapies underwent.
Other than orphan designation, NICE has also formulated burden of illness (NICE, n = 2) and end-of-life (NICE, n = 3) criteria that some gene therapies may meet and that could influence the HTA outcome. Such additional considerations are not taken explicitly into account by G-BA.

Lack of long-term data was found to be a key consideration in rationales for decision making for both NICE and G-BA (G-BA, n = 5; NICE, n = 6) (Table 4.2). Indeed, the majority of G-BA assessments (n = 5) have a non-quantifiable benefit following lack of long-term data.

However, the approach to mitigating such evidence uncertainties slightly differs between both agencies. Whereas G-BA appears to mainly rely on temporary decisions (n = 6), NICE appears to resort to discounts (n = 6). The agencies have similarly applied outcomes-based agreements and continued evidence development agreements. Such agreements comprise collection of RWD/RWE, the usage of which is discussed in the next section.

4.2. RWD/RWE usage in gene therapy HTA varies

To answer sub-question 2a (*What role does RWD/RWE play in the HTA appraisal process of gene therapies?*), a retrospective comparative analysis was performed. In this analysis, RWD/RWE usage by G-BA and NICE in gene therapy HTAs was explicated.

While NICE uses both RWE and RWD, G-BA mainly uses RWE

RWD/RWE usage was derived from the identified HTA reports (Table 4.3). While the number of sources used for RWE was comparable, (G-BA, n = 66; NICE, n = 86), the number of RWD sources differed (G-BA, n = 7; NICE, n = 56). The difference in average RWD/RWE usage per HTA report (NICE, = 14; G-BA, N = 8) appears to be attributable to the low RWD usage of G-BA. However, areas supported by RWD in NICE HTA reports did not differ substantially from the areas supported by RWE. As such, it appears that the lower RWD uptake by G-BA is not linked to the area that it supports.

Table 4.3. RWD/RWE sources G-BA and NICE. RWD: Real-world data. RWE: Real-world evidence.

	G-BA	NICE
RWE sources	66	86
RWD sources	7	56
Total	73	142

The most common RWD sources are interviews (G-BA, n = 3; NICE, n = 29) and disease registries (G-BA, n = 4; NICE, n = 11), together accounting for 72% and 100% of total RWD usage by NICE and G-BA, respectively (Appendix A7).
For both agencies, the most frequently used RWE study design is a retrospective cohort study (G-BA, n = 27; NICE, n = 35), followed by non-randomised controlled trials (G-BA, n = 19; NICE, n = 20) and prospective cohort studies (G-BA, n = 14; NICE, n = 14).

From the presented data, it appears that the RWD types (i.e. interviews & disease registries) used by both agencies does not differ considerably. The same is true for RWE study designs (i.e. retrospective

cohort study & non-randomised controlled studies).

RWD/RWE supports different areas for NICE and G-BA

While the types of RWD/RWE used by G-BA and NICE are comparable, the reasons for inclusion, e.g. area of HTA supported appears to differ (Figure 4.1).
The areas supported by RWE and RWD individually do not differ substantially (Appendix A7). For both agencies, the main area supported by RWE studies is evidence on an external comparator to assess clinical benefit and safety (G-BA, n = 34; NICE, n = 31). Other areas substantially supported by RWD/RWE include information on the patient population (G-BA, n = 14; NICE, n = 37) and effectiveness of the intervention (G-BA, n = 18; NICE, n = 23).

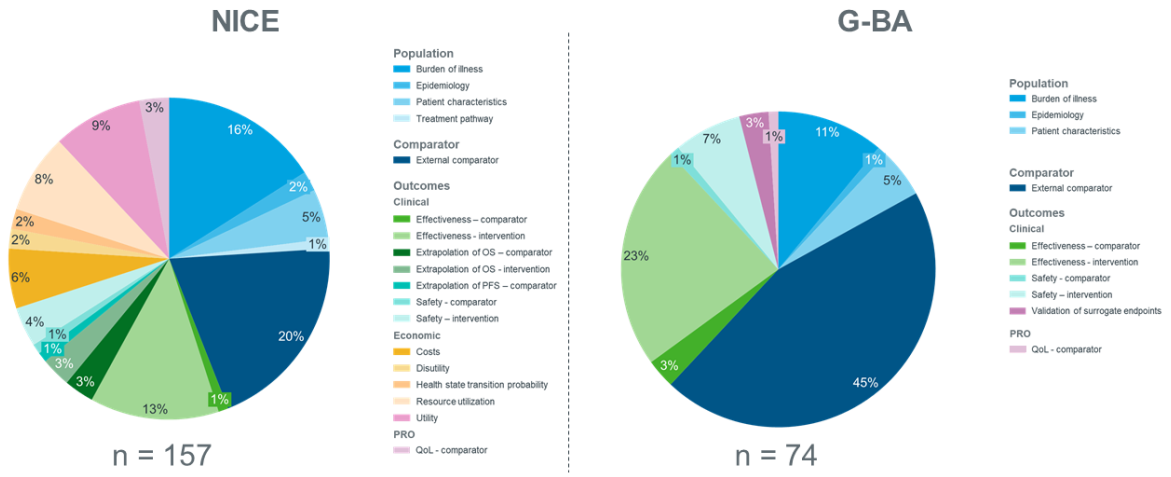


Figure 4.1. Areas supported by RWD/RWE in NICE and G-BA gene therapy HTAs. As one RWD/RWE source may be used to support different areas, the total volume of RWD/RWE usage (G-BA, n =74; NICE, n = 157) differs from the number of sources mentioned in table 4.3. OS: Overall survival. PFS: Progression free survival. QoL: Quality of life. RWD: Real-world data. RWE: Real-world evidence.

A considerable difference is evident from RWD/RWE used to support economic outcomes. While G-BA does not consider this area in their assessments, it accounts for 25% of RWE usage by NICE. This is expected, considering the different HTA archetypes of G-BA (REA) and NICE (CEA): while NICE takes economic considerations into account, G-BA does not.

4.3. RWD/RWE appraisal in gene therapy HTA varies

To find out whether one HTA agency is more amenable to RWD/RWE usage than the other, the appraisals of RWD/RWE were analysed.

Appraisal was categorised as ‘accepted’, ‘not accepted’ ‘not identified’, or ‘other’ (Figure 4.4). RWD/RWE sources categorised as ‘not identified’ or ‘other’ were not considered in this analysis. A complete overview on the RWD/RWE appraisal per area supported is provided in Appendix A7.

The acceptance rate (defined here as the ratio between ‘accepted’ and ‘not accepted’ RWD/RWE usage) appears to be higher for NICE than G-BA (G-BA, n = 18/39; NICE, n = 35/27) (Figure 4.4). The same is true for the ‘accepted’ RWD/RWE usage compared to the total volume of RWD/RWE usage (G-BA, n = 18/74; NICE, n = 35/118), albeit with a smaller difference.

Table 4.4. RWE/RWD sources G-BA and NICE. 'Other' refers to to RWD/RWE usage in suspended and ongoing reports (e.g. NICE, Zynteglo®; NICE, Libmeldy®).

	G-BA	NICE
Accepted	18	35
Not accepted	39	27
Not identified	17	56
Used in Analysis	74	118
Other	-	39
Total	74	157

RWD/RWE acceptance rates differ on areas supported for NICE and G-BA

While the overall acceptance rate by NICE may be higher, the acceptance could potentially differ between area supported.
To explicate whether the acceptance rates of RWD/RWE vary per area supported, the differently appraised RWD/RWE uses are split out per area supported per agency (Figure 4.2).

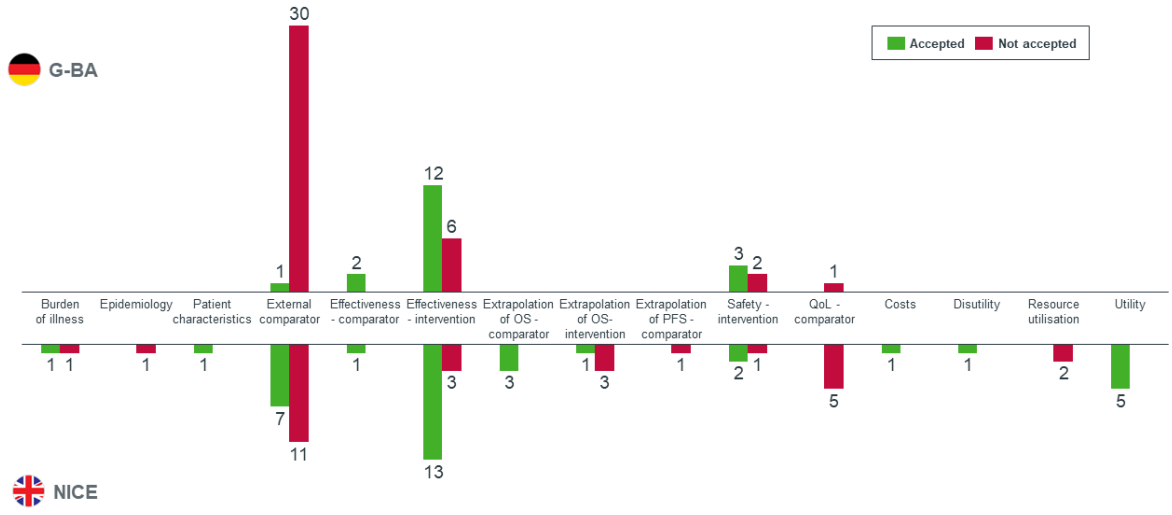


Figure 4.2. RWD/RWE acceptance per area supported for G-BA and NICE. OS: Overall survival. PFS: Progression free survival. QoL: Quality of life.

The acceptance rate for RWD/RWE on an external comparator is particularly low for G-BA, compared to NICE (G-BA, n = 1/30; NICE, 7/11). Another supported area where RWD/RWE acceptance differs considerably is the effectiveness of intervention (G-BA, n = 12/6; NICE, 12/2). It therefore appears that both agencies tend to accept RWD/RWE usage to support this area more often than not, albeit in different ratios.

NICE and G-BA appear to generally align on the appraisal of RWD/RWE

In the final part of this analysis, the extent to which NICE and G-BA evaluate the same RWD/RWE source differently is evaluated.

RWD/RWE sources and their appraisal were compared between the two agencies. Appraisals of the same RWD/RWE source by NICE and G-BA that were identified as 'accepted' or 'not accepted' were categorised as 'opposing' or 'corresponding'. Other RWD/RWE sources were categorised as 'unique' to either one of the agencies (Table 4.5), regardless of whether the appraisal was identified in the HTA reports.

Unique sources evaluated by NICE mainly comprised economic outcomes (i.e. costs, utility, resource utilisation) (n = 39), burden of illness (n = 19) and external comparator data (n = 14) (Appendix A7). For G-BA the majority of unique sources cited support data on external comparator (n = 15). Other areas supported comprise effectiveness of intervention (n = 4) information of population characteristics (n = 3) and validation of surrogate endpoints (n = 2).

Table 4.5. RWE/RWD usage comparison NICE vs G-BA. Only includes RWD/RWE sources identified as accepted or not accepted for both agencies. *: not included in appraisals but different RWD/RWE uses included. n/a: not applicable.

Product	Opposing appraisals	Corresponding appraisals	Unique sources G-BA	Unique sources NICE
Imlygic®	-	-	-	7
Kymriah®(DLBCL)	-	3	3	10
Kymriah®(ALL)	1	4	2	6
Yescarta®	-	1	12	5
Luxturna®	-	-	-	13
Zynteglo®*	n/a	n/a	4	21
Zolgensma®	-	7	0	19
Tecartus®	1	4	2	7
Libmeldy®*	n/a	n/a	1	7

Table 4.6. Similarly appraised RWD/RWE in gene therapy HTA reports published by NICE and G-BA. Only includes RWD/RWE sources identified as accepted or not accepted for both agencies. QoL: Quality of life.

Area supported	Kymriah®-DLBCL	Kymriah®-ALL	Yescarta®	Zolgensma®	Tecartus®	Total
External comparator	1	3	-	-	3	7
Effectiveness - intervention	2	1	-	4	1	8
Effectiveness - comparator	-	-	1	-	-	1
Safety - intervention	1	-	-	2	-	3
QoL - comparator	-	-	-	1	-	1
Total	3	4	1	7	4	20

Corresponding appraisals mainly concerned the effectiveness (n = 8) and external comparator (n = 6) (Table 4.6).

Two instances were found where the same evidence was appraised differently (Table 4.7). In both cases, RWD/RWE was used to provide evidence on an external comparator.
While G-BA did not accept these studies following concerns on the applicability of the submitted studies, NICE did acknowledge similar weaknesses but still accepted the studies.

Table 4.7. Differently appraised RWD/RWE in gene therapy HTA reports published by NICE and G-BA. RWD: Real-world data. RWE: Real-world evidence.

Product	RWD/RWE source	Area supported	G-BA	NICE
Kymriah®- ALL	Jeha et al., 2006	External comparator	No information is available on the specific ALL diagnoses in the studied population	The committee accepted that the study had a number of limitations, but concluded that that it was appropriate to consider in its decision-making.
Tecartus®	McCulloch et al., 2020	External comparator	Insufficient comparability with clinical study on clinical effectiveness of the intervention (differences in study design, inclusion criteria, data collection, characteristics of study population)	"Using data derived from this study was considered to be more appropriate than using uncertain estimates from an indirect treatment comparison."

As such, the two instances mentioned above are in line with earlier identified difference in RWD/RWE acceptance rates of G-BA and NICE; NICE appeares more amenable to RWD/RWE usage than G-BA. However, when placing these two instances in the context of the total volume of RWD/RWE submitted (e.g. n = 7 instances for external comparator), it appears that NICE and G-BA align on the appraisal of evidence (n = 5) more often than not (n = 2).

4.4. Synthesising preliminary findings

Based on the results presented in the previous sections, the following preliminary findings were synthesised.

Table 4.8. Preliminary findings from the retrospective comparative analysis. Findings in *italics* were not probed in use cases or validated in interviews. HTA: Health technology assessment. RWD: Real-world data. RWE: Real-world evidence.

Sub-question	Preliminary findings	Description	Reference
1	*1*	*Gene therapy HTA outcomes differ between G-BA and NICE*	*Section 4.1*
	2	(Lack of) long-term data is a key consideration in rationales for decision making by both G-BA and NICE	Section 4.1
	3	G-BA and NICE take different approaches to mitigate evidence uncertainty	Section 4.1
2	4	NICE is more amenable to RWD/RWE usage in gene therapies than G-BA.	Section 4.2
	5	*For both G-BA and NICE, retrospective cohort studies are the most commonly cited RWD/RWE sources, followed by non-randomised controlled trials and prospective cohort studies*	*Section 4.2*

continues on next page

6	RWD/RWE supports different areas for NICE and G-BA	Section 4.2
7	RWD/RWE acceptance rates by NICE and G-BA differ per area supported	Section 4.2
8	RWD/RWE is generally appraised similarly between NICE and G-BA	Section 4.2

The preliminary findings presented in Table 4.8 are based on a rather superficial dataset. As such, use cases were performed to provide additional context to the identified RWD/RWE usage, and probe the synthesised findings. Findings 1.1 and 2.5 were considered sufficiently covered in literature and therefore deprioritised in these use cases.

4.5. Use cases illustrate RWD/RWE usage in HTAs

Three illustrative use cases (Imlygic®, Yescarta®and Zolgensma®) will be presented in this section. A complete overview on identified RWD/RWE usage is provided in Appendix A8.

Imlygic®

No RWD/RWE usage was identified for G-BA

In December 2016, G-BA published the final resolution and justification on Imlygic®(Gemeinsamer Bundesausschuss, 2016). As this therapy did not have an orphan designation, it did not have an additional medical benefit rating by default.

The evidence submitted by the company was primarily based on the OPTiM trial, a multinational phase three randomized clinical trial that ran from April 2009 to September 2014. However, the committee considered the comparator used in the OPTiM study inappropriate. In addition, no results of direct comparative studies were available for any patient population to demonstrate an added benefit of Imlygic®. The company was also unable to identify studies suitable for an indirect comparison.
No real-world evidence usage was mentioned to mitigate this lack of evidence. Hence, the committee noted that from the presented evidence, it was not possible to demonstrate an additional benefit compared with the the current standard care in the therapeutic area. Consequently, G-BA issued a negative recommendation, where added benefit of Imlygic was not proven for all three considered subgroups.

NICE mainly used RWD/RWE in cost-effectiveness assessments

Similar to HTA report submitted to G-BA, the evidence on effectiveness of the intervention was primarily based on the OPTiM trial. The committee concluded that the most clinically relevant comparator within the scope for this appraisal was ipilimumab (NICE, 2016).

To model long-term survival beyond the data that was obtained from the OPTiM trial, the company relied on registry data from the American Joint Committee on Cancer (AJCC), as well as UK life tables from the Office of National Statistic. While the committee did accept the multi-staged modelling approach taken, it concluded that it had not been presented with a plausible incremental cost-effectiveness ratio (ICER) for the intervention compared to the standard of care and did therefore not accept the submitted evidence. Interestingly, the critique by NICE mainly relates to the methodological assumptions underlying the model that was used to inform decision making, rather than the quality or appropriateness of the RWD/RWE sources.

The other four areas supported by RWD/RWE, were part of the CEA. However, their acceptance could not be derived from the HTA report.

A retrospective cohort study (MELODY) informed healthcare resource utilisation that would be associated with adopting Imlygic® in the National Health Service. Similarly, RWD from electronic medical records informed resource utilisation, as well as costs. Finally, the manufacturer derived utility decrement values from a time-trade off study that was conducted among 300 respondents in the general UK population.

In the end, NICE recommended that Imlygic should be restricted to people with melanoma for whom immunotherapy is not suitable or otherwise contraindicated.

The observed difference in RWD/RWE usage appears to be related to the area supported

While no RWD/RWE usage was identified for G-BA, it appears to play a supportive role in the case of NICE. RWD/RWE is used to provide benchmark data in informing extrapolations and assessments, specifically related to the cost-effectiveness. As such, the difference in RWD/RWE usage appears to be attributed to the area that it supports; G-BA does not consider cost-effectiveness in its HTAs.
Alignment on the evidence appraisal can therefore not be assessed in this case. It should be noted, however, that the HTA bodies do not align on the appropriateness of the comparator. Whereas, NICE does accept ipilimumab as a clinically relevant comparator, G-BA does not. This appears to be in line with the earlier observed difference in acceptance rates between both agencies, especially for this area.

Yescarta®

G-BA did not accept the majority of RWE submitted, due to lack of transparency

G-BA's assessment for Yescarta® was published in February 2019 (Geimeinsamer Bundesausschuss, 2019). Primary study outcomes were based on the ZUMA-1 trial, an ongoing phase I/II multi-center, open-label, single arm study that started in January 2015 and is expected to be completed by September 2035. In absence of a direct comparator, the manufacturer proposed to use the retrospective SCHOLAR-1 study as a proxy. In addition, the NCI 09-C-0082 supportive study is mentioned. However, this open, single-arm phase I dose-finding study was not used for the benefit assessment, as dosing amounts did not conform to regulatory requirements.
While G-BA noted uncertainties and possible differences between the patient populations in both the ZUMA-1 and SCHOLAR-1 studies, it did consider the indirect historical comparison sufficiently valid to assess the additional benefit of Yescarta®.

In addition to the above, 15 retrospective studies were submitted as alternative indirect historical comparators. Yet, the G-BA found that relevant differences to compare of the patient populations were not given and that most studies lacked information on the patient characteristics. These studies were therefore not used for the benefit assessment.

G-BA noted that an indirect historical comparison is highly sensitive to bias. Taking into account the other uncertainties regarding long-term effects, sample size and patient populations, G-BA concluded from the real-world evidence sources that an effect is present but cannot be quantified for both considered patient populations. However, as Yescarta® received an orphan designation, the therapy received a positive recommendation by default.
As such, G-BA provided a temporary positive recommendation, valid through May 2022. Yescarta® will then be re-assessed based on the results of the ZUMA-1 study after five years, as well as additional comparative evidence for relevant further knowledge gain for the benefit assessment.

Lack of sufficient data quality and completeness in real-world evidence is apparent in this case from the 15 retrospective studies that lack relevant patient information. This appears to be in line with earlier synthesised findings, where this impedes the uptake of RWD/RWE.

Immature survival data and limitations in the comparator data hampered assessment of the added benefit by NICE

NICE published its final appraisal document for Yescarta® in November 2018 (NICE, 2019). Similar to the G-BA appraisal, the results of the ZUMA-1 and SCHOLAR-1 studies were used.
While the evidence review group noted that "comparative effectiveness results from single-arm trials are prone to bias", the committee concluded that this approach was suitable. Yet, the SCHOLAR-1 study was not considered representative for patient populations in the NHS. In response to this critique, the manufacturer provided RWE from a patient cohort audit from an Oxford University Hospitals database to validate the appropriateness of SCHOLAR-1. However, due to the limited sample size of 41 patients, this audit was not further considered. Instead, the committee concluded that Yescarta® was "clinically effective compared with salvage chemotherapy, but immature survival data and limitations in the comparator data mean that the exact size of the benefit is unknown".

In addition to the above, three RWD soures and one RWE study were submitted to support evidence the burden of illness and on current treatment pathways. However, the acceptance of these sources was not identified.
Due to the uncertainty in available evidence, NICE recommended the use of Yescarta® for use through the Cancer Drug Fund, conditional on a managed access agreement where follow-up data is required. By February 2022, gathering five-year follow-up data from the ZUMA-1 clinical trial is anticipated to conclude. NICE will then evaluate its guidance for Yescarta®.

While both HTA bodies acknowledged weaknesses in the submitted RWE, it did allow for comparing evidence on clinical effectiveness

In both cases, RWD was used to provide and compare evidence on clinical effectiveness. While the RWD/RWE submitted was mostly unique to the HTA bodies, both did accept the SCHOLAR-1 study, albeit with concerns on bias and uncertainty.

Interestingly, more RWD/RWE sources were identified in the G-BA HTA report. This is in contrast with earlier synthesised finding that NICE reports use more RWD/RWE sources. Also, contrary to the previous use case, the areas supported by RWD/RWE were not found to differ substantially for the HTA bodies. Indeed, no RWD/RWE sources were identified to inform CEA parameters in the NICE HTA.

Zolgensma®

As annual sales exceeded €50 million, Zolgensma® did not benefit from its orphan designation in the G-BA assessment

Zolgensma is the first therapy to be assessed under GSAV regulations by the G-BA (Gemeinsamer Bundesausschuss, 2021). This means that since its budget impact was estimated to exceed €50 million by IQWiG, Zolgensma had to go through a full HTA process, without its orphan designation privileges. In such cases, the legislation requires a direct comparison with an appropriate comparative therapy. Moreover, pharmaceutical companies may be obligated to generate and collect post-launch evidence.

The company did not submit randomised clinical trial evidence data that would allow for a direct or adjusted indirect comparison with the appropriate comparator therapy (Biogen's Spinraza (nusinersen)). Instead, it included for patients with SMA type 1 the single-arm studies START, STR1VE-EU and STR1VE-US and for nusinersen the randomised clinical trial ENDEAR and the non-randomised single-arm study CS3A, as well as its extension study SHINE.

RWD/RWE on the safety and effectiveness of the intervention came from prospective cohort study LT-001 and the non-randomised controlled trial CL-101. Both evidence sources were accepted by G-BA. In addition, the company submitted the ongoing LT-002 observational follow-up study of these single-arm trials. However, since no data was available, this study was not considered.

Data from the non-randomised phase I dose comparison study STRONG was not considered as the intrathecal use in this study was not in conformity with the technical application. Other rejection rationales for RWD/RWE on the effectiveness of the comparator comprised it being a divergent intervention. Finally, the acceptance of seven RWD/RWE sources was not identified. These were mainly used to support evidence on the burden of illness and patient characteristics.

The committee noted that there are clear differences in the mean duration of disease, which is a very significant confounder. It concluded that due to the large uncertainties, the presented comparison are not relevant for the benefit assessment and cannot be used to derive an added benefit. An added benefit was therefore not proven.
For the other three indications, Zolgensma®was again found to offer no additional benefit over the comparative therapies for treating spinal muscular atrophy.

NICE used RWD/RWE to inform several aspects of its assessment

Similar to G-BA, the main clinical effectiveness evidence in NICE's final evaluation document was derived from two completed open-label single-arm studies, START and STR1VE-US. In addition, the company provided interim data of two ongoing single-arm studies: STR1VE-EU and SPR1NT, as well as a long-term follow-up study of START, LT-001. The latter is an prospective observational study anticipated to be completed by December 2033 and is considered RWE.

As none of the above-mentioned studies had a control arm, the company identified four potential natural history studies to estimate outcomes for best supportive care. Three of these were considered real-world data: a prospective study, a retrospective study and a database-derived study. The committee acknowledged flaws in each of these natural history studies but considered the prospective NeuroNext study the most suitable to estimate best supportive care outcomes.

To model long-term outcomes for different health states, the LT-001 study was complemented with three additional real-world evidence sources. This included a retrospective chart review, a prospective & retrospective study and UK life table data from the office for national statistics. Yet, the committee noted that "there were limited data to inform long-term outcomes in the model and that this was a key area of uncertainty" and concluded that although Zolgensma®is likely to have long-term benefits, the exact amount of benefit was uncertain.

To inform costs and utilities of different health state scenarios in the cost-effectiveness assessment, the company submitted multiple real-world data sources in the form of a cross-sectional and clinician proxy vignette study, as well as systematic patient surveys. Again, the committee noted considerable uncertainties in the model and underlying assumptions but concluded that "they appeared to be the most appropriate to use in decision making". Other RWD/RWE sources mainly served to support evidence on quality of life (used in CEA) but also on the burden of illness or patient characteristics. For the latter, acceptance was often not identified.

Taking all evidence and uncertainties into account, the committee deemed a managed access agreement most suitable. In three years, NICE will re-evaluate its guidance based on (real-world) evidence that is being collected to resolve some of the identified uncertainties.

RWD/RWE usage NICE and G-BA partially overlaps

In the case of NICE, RWD/RWE was used to inform multiple HTA aspects, albeit with some uncertainties on the appropriateness. Lack of sufficient data quality did not appear to hamper real-world evidence usage in this HTA. Similarly, the lack of randomisation was not explicitly mentioned as a downside. Lack of long-term data was mentioned multiple times as supporting rationale for not accepting RWD/RWE sources.
Again, the appraisal of evidence by NICE and G-BA was similar, but a larger volume of RWE/RWD

was submitted to NICE. Similar to earlier observations, this can in part be attributed to the differences in HTA archetype between NICE and G-BA.

Analysis of aggregate data from the use cases

No clear trend is apparent in RWD/RWE usage over time

The total volume of RWD/RWE in the use cases was visualised (Figure 4.3).

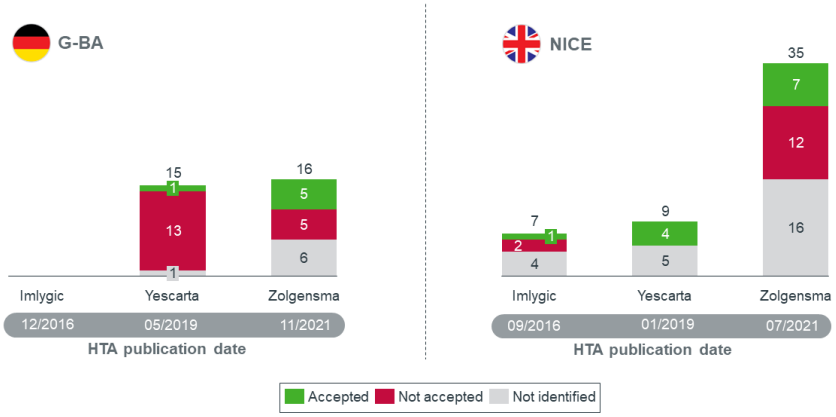


Figure 4.3. RWD/RWE appraisal of use cases over time. HTA: Health technology assessment.

Based on the limited data points, no clear trend is visible for the volume of RWD/RWE usage over time for both agencies. In a similar vein, no trends were identified in areas supported or acceptance rates.

Formulating insights from comparing RWD/RWE exclusion rationales

The RWD/RWE exclusion criteria identified in the use cases were categorised in Table 4.9.

Table 4.9. Quantifying exclusion rationales from the use cases. HTA: Health technology assessment. RWD: Real-world data. RWE: Real-world evidence.

Rationale	Description	G-BA	NICE
1	Lack of suitable effectiveness inputs	0	2
2	Lacking information / comparability not provided	11	2
3	Relevant differences of the patient characteristics	1	0
4	Measurements in performed study not conform regulatory requirements/ current practices	2	1
5	Limited data set	0	3
6	No data available	2	2
7	Methodology used in RWD/RWE source not appropriate	0	1
8	Not used as other RWD/RWE source was better suitable	0	1
9	Not used as RWD/RWE source is a duplicate	0	1

Based on the presented rationales, it appears that gene therapy manufacturers should take the following aspects into account when submitting RWD/RWE in gene therapy HTAs.
First, it appears that manufacturers should be critical towards the data that they submit, i.e. does it

provide information to substantiate its relevance and is the data set extensive (rationales 2, 5 & 6). Especially for G-BA this appears to be a confounding factor.
Another relevant consideration appears to be linked to the appropriateness of the RWD/RWE sources, i.e. whether it reflects practices in the current standard of care (rationales 4 & 8).

Findings to be validated in interviews

The use cases provided context to the earlier identified RWD/RWE usage. Moreover, the exclusion rationales were categorised to and the preliminary findings presented in Table 4.8 were probed.
Use cases supported finding 2.4 (in general more, RWD/RWE is submitted to NICE) and that NICE also has a higher acceptance rate than G-BA (finding 2.7). Moreover, in line with finding 2.6, RWD/RWE was often used to support CEA aspects in the NICE HTAs.

It should be noted that the synthesised findings and generated insights are based on a rather limited data set. Moreover, RWD/RWE usage may have been misinterpreted (as discussed in chapter 6). As literature delivering empirical evidence on RWD/RWE usage was found to be scarce, most of these findings could not be validated through literature alone. As such, interviews were conducted to validate findings 1.2, 1.3, 2.4, 2.6 - 2.8. Again, findings 1.1 and 2.5 were considered sufficiently covered in literature and therefore deprioritised.

5

Interviews

Interviews with industry experts and academics served to validate the preliminary findings described in the previous chapters. In addition, the interviews aimed to gather insights on potential challenges and opportunities in HTA body alignment, specifically in the context of future joint clinical assessments of gene therapies.

5.1. Validating preliminary findings

Preliminary findings included output from the literature review, as well as findings from the retrospective, comparative analysis.

Understanding the current practices in gene therapy HTAs

In the first part of the interview, preliminary findings on sub-question 1 (*How are gene therapies currently assessed in HTAs?*) were validated. The goal was to obtain a better understanding of the current practices in gene therapy HTAs, including specific considerations or frameworks applied in their appraisal.

While most special considerations in gene therapy HTAs are shared with other therapy types, curative potential may be considered unique

When asked for special considerations in gene therapy HTAs, three interviewees mentioned that orphan designation may influence the HTA outcome for some HTA bodies. One interviewee also mentioned burden of illness and end of life criteria considerations, but noted that these are specific to NICE; G-BA does not explicitly consider this.

However, each interviewee emphasised that the special considerations mentioned above are not unique to gene therapies. Interviewee A2 explained: *"I always find it really difficult to just look at HTAs of gene therapies. I think you should put them in a wider perspective so compare them to other types of HTAs and I think if you compare them to other types HTAs, [...] I think the key considerations aren't that different, right? Because it's the same framework they are being assessed in."*

The curative potential of gene therapies, however, was mentioned in two interviews to be an unique consideration for this therapy type. Other interviewees did not mention this consideration explicitly. However, part of this consideration may be captured in the uncertainty regarding long-term benefits. This is a challenge that every interviewee mentioned during the interview, albeit with the nuance it is not unique to gene therapies.

Similarly, HTA frameworks are not specific for gene therapies, although that may change over time

Interviewees agreed that there are currently no HTA frameworks that are specifically tailored to gene therapies. Indeed, all interviewees noted that the methodological and practical challenges encountered in gene therapy HTAs are also not necessarily unique to this therapy type. To this point, interviewee I3 noted that *"You would like to think that these innovations are very special and very unique, but there are just a few key challenges and they're kind of shared by these innovations [...]."*

The absence of specific frameworks therefore appear to be linked to the fact that the challenges are not unique. However, this could change in the coming years, according to interviewee I1: *"They [Zorg*

Instituut Nederland] basically say that they need to adjust their frameworks. I can see something similar happening for some of the HTA bodies, basically saying that our systems may not be fully equipped to assess cell and gene therapies and we may need to adjust our frameworks to account for special considerations." A special considerations in this context could then include curative potential. Other than this statement, no additional mentions on potential changes to HTA frameworks were identified in the interviews.

Discounts are not specific to gene therapy types and do not serve to mitigate uncertainty

From the preliminary findings described in the previous chapter, it appeared that discounts were commonly adopted in gene therapy HTAs. However, interviewee I2 noted that *"Discounts are a way to lower the price and decrease the budget impact. [...] I don't think it necessarily that something to do with the increased uncertainty [...] I don't think discounting per se is anything specific to gene therapies"*. When this findings was probed in interviews, all responses shared the view that discounts are not a way to mitigate uncertainties and that they are not specific to a particular therapy type.

Managed access agreements appear well-suited to mitigate evidentiary uncertainties

Instead, interviewees agreed that the concept of managed access agreements can be applied to mitigate evidentiary uncertainties. While three interviewees noted that this approach was increasingly common, interviewee I1 was a bit more reserved, noting that *"[...] it is not that we see an overwhelming amount of agreements that are more advanced than simple discounts. Currently, the impact is relatively limited but everyone is looking to change that"*.
From the interviewees' responses it became clear that both G-BA and NICE embrace such novel payment forms. To this point, interviewee I2 mentioned that "*G-BA of course has been the key development in the last years where they have started to ask for real-world registries, which mandate manufacturers to collect RWE on the product post-launch and then doing a reevaluation. NICE has had similar processes where they would reevaluate products in a certain number of years*". Indeed, mitigating uncertainty through continued evidence development as part of regulatory approvals or novel payment agreements was specifically mentioned by three interviewees.

The current and future role of real-world evidence in gene therapy HTAs

Continued evidence collection from a real-world setting appears to imply the collection of RWD and synthesis of RWE. As such, the focus of the interview shifted towards the role of RWD/RWE to validate preliminary findings on sub-question 2 (*What role does RWD/RWE play in the HTA appraisal process of gene therapies?* and *How has RWD/RWE usage in gene therapy HTAs evolved?*).

While the role of RWD/RWE may increase over time, it is not a 'silver bullet'

When asked to what extent RWD/RWE could mitigate challenges encountered in appraising gene therapies, all interviewees agreed that it at least has the potential to support decision making. A recurring caveat in the interviewees' responses was that RWD/RWE is not the single answer to these uncertainties, or as interviewee I3 put it: *"if anything, it probably moves the needle more from a no to a maybe"*. There was a general consensus among interviewees that RWD/RWE currently mainly serves to support particular areas in HTAs. Two interviewees noted that the area supported depends on the quality of the evidence and what insights can be derived from it. Common areas where RWD/RWE can play a supportive role according to the interviewees, included describing the patient population (n = 3), the national history of the disease (n = 2) and comparative effectiveness (n = 3). For NICE specifically, healthcare resource use and shape of the long term extrapolations were also mentioned by interviewee I3.

In general, interviewees' view on how the adoption of RWD/RWE may evolve over time is in line with the statement made by interviewee I1: *"[...] overall, the role of RWE is here to stay and also will*

have an increasing impact". However, where interviewees saw this increasing impact most likely to happen, remained mostly unknown from the responses. One interviewee (I2), mentioned supporting effectiveness or supporting the validity of certain endpoints as areas where RWD/RWE may play a pivotal role in future HTAs.

The role of real-world evidence differs between HTA agencies

The majority of interviews noted that RWD/RWE usage differs between HTA agencies. Interviewee A1 noted that *"different agencies have different methods, different processes and different principles of operation [...] they will have different preferences for additional source of evidence, including real-world evidence, right?"*

Industry experts I1 and I2 provided more detail on the differences in areas supported by RWD/RWE between G-BA and NICE. Both saw a substantial difference in RWD/RWE usage to support effectiveness, interviewee I1 explained: *"[...] if you look at effectiveness as an area supported, you will not find any case in Germany where that has been accepted. Whereas in the UK you won't find loads of cases, but there are cases where RWE really is accepted as a source to inform effectiveness."*

Accelerated regulatory pathways and international initiatives may drive RWD/RWE uptake

A few enabling factors for RWD/RWE uptake emerged from the interviews. International initiatives and collaborations were mentioned most often (n = 4). However, two interviewees noted that progress in such initiatives has been slow.

Other factors included technological advancements of clinical systems and techniques to analyse those data. Interviewee I1 explained: *"common data sets, developing standardization of what is collected in certain registries, what's collected in EMR [electronic medical records] across Europe will certainly enable the use of RWE"*.

The accelerated regulatory pathways was also recurrently mentioned as a factor that drives RWD/RWE usage. Interviewee I2 explained that *"products get proved based on more limited data sets with mandatory post authorization data collection and sometimes also RWE as part of the regulatory submissions [...]"*.

HTA bodies themselves may be a bottleneck in RWD/RWE uptake

In addition to the above, interviewee I2 also noted that the accelerated regulatory approval has *"put the burden on HTA bodies what they do with this evidence. We see that they are struggling and lagging behind [...]"* As such, it appears that while this development has lead to an increase in RWD/RWE submissions, it has also exposed that HTA bodies themselves may be a barrier in RWD/RWE uptake. It seems that this contrast can, at least in part, be attributed in the difference of evidence requirements. Interviewee I3 explained: *"there is this tension between what is being asked by regulatory bodies and what is needed by payers to feel confident that they are dealing with a value for money product"*.

Interviewees were generally aligned on barriers that impede RWD/RWE uptake

Another recurrently mentioned barrier was the lack of prescriptive guidance. Three interviewees saw this as a substantial barrier to increase the adoption of RWD/RWE. Interviewee I1 explained that *"at the moment, the vast majority of RWD/RWE is submitted because the manufacturer believes that it can help them, not because the HTA bodies requested it"*.

Other factors that impede increased RWD/RWE uptake that were mentioned, included the data privacy (n = 1), lack of standardisation (n = 2), risk of bias in the data (n = 1) and heterogeneity in patient populations (n = 1).

Interviewees did not mention RWD/RWE terminology as a factor that impedes RWD/RWE uptake. When probing for this factor, interviewee I2 responded that *"I don't think the alignment on the terminology of RWE plays a role per se, I mean they're very broad, right? There are differences, certainly in*

the types of studies which are accepted in one country from another, but I don't think there's confusion around the definitions per say that hinder uptake of RWE."

The HTA body archetype and inherent characteristics influence RWD/RWE uptake

The barriers and enablers covered in the paragraphs above appear to be generally applicable to RWD/RWE uptake by most HTA bodies. When interviewees were asked to differentiate RWD/RWE usage (e.g. openness to its usage) between G-BA and NICE, a clear difference became apparent in both areas supported and acceptance.

From the responses, it became clear that the German system has a very mechanistic way of looking at the data, where they apply the same methodology and threshold across therapy types. NICE, on the other hand, appears to be much more flexible in its approach.

Similar to the HTA body characteristics, their approaches to HTAs (REA & CEA for G-BA & NICE, respectively) also appear to influence RWD/RWE usage. Interviewee I1 explained why NICE's CEA archetype may be particularly compatible with RWD usage: *"there's a lot of assumptions that need to go in there anyway and if they sound plausible to the various committees and etc., then that's OK"*. As such, it appears that per definition, NICE considers aspects (e.g. cost-effectiveness) that are inherently more suitable to be substantiated by RWD/RWE than aspects considered in REAs.

Interviewees were generally aligned on NICE being more open to RWD/RWE usage than G-BA. While this difference may be partially attributed to the archetypes, the methodological strictness was also found to be relevant. Interviewee I1 explained: *" It comes down to the mechanistic view in Germany, where they apply the same methodology and threshold across therapy types. NICE is more flexible in its approach and the way it is dealing with RWD/RWE [...]."* Three interviewees mentioned this difference in methodological strictness as a considerable influence on the RWD/RWE usage.

Multiple factors influencing RWD/RWE uptake appear to connected

From the conducted interviews, several factors relevant to RWD/RWE uptake emerged. The aggregate findings and their relationships were visualised in a network (Figure 5.1).

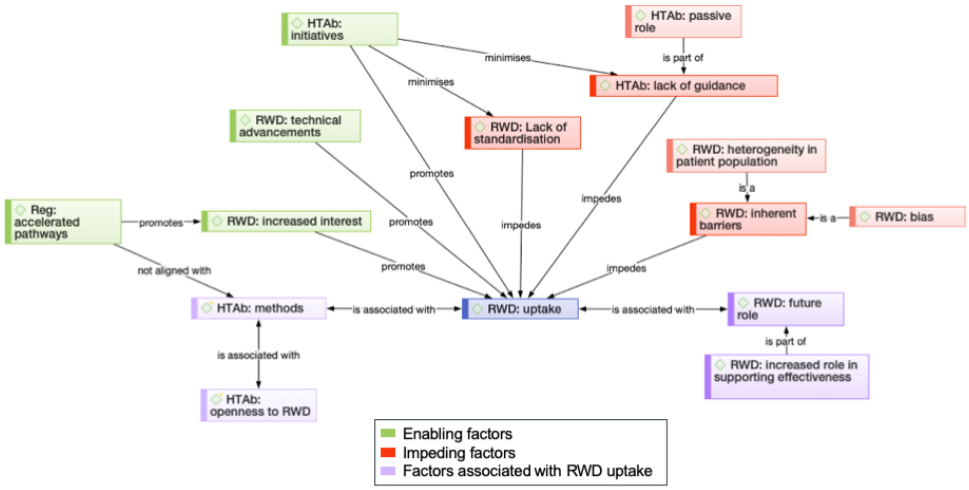


Figure 5.1. Conceptual network of factors influencing RWD/RWE uptake. Green: enabling factors. Red: impeding factors. Purple: factors associated with RWD uptake. HTA: Health technology assessment. HTAb: HTA bodies. Reg: regulatory agencies. RWD: Real-world data. RWE: Real-world evidence.

From this network, it appears that many factors that affect RWD/RWE uptake are interrelated. The current and future role of RWD/RWE are associated with the uptake of RWD/RWE. Several promoting and impeding factors affect RWD/RWE uptake. While international initiatives may mitigate some of

the barriers, including lack of guidance and standardisation, other barriers like confounding bias and heterogeneity among patient populations appear to be unsolvable by nature.

5.2. Exploring potential challenges and opportunities in HTA body alignment

All interviewees were aware of the recently introduced legislation on EU-wide joint clinical assessments. However, the majority noted that this legislation is still in its infancy and that many aspects are still unknown.

Joint clinical assessment pilots may solve some uncertainties

One of these unknown aspects is its implication for individual HTA bodies. Interviewee I2 rightfully pointed out that: *"The joint clinical assessment is going to be on the clinical data. Legislation says that individual country HTA bodies can ask for additional data but then need to substantiate why, i.e. we want to see local study data etc. That is still in place"*. Therefore, the uncertainties appear mainly relate to what national HTA bodies will do with the clinical assessment outcome and to what extent it satisfies their evidence requirements. Given these uncertainties, interviewees appeared slightly dubious towards the feasibility given the complexity and variability between countries. However, no strongly positive or negative views were identified.
Uncertainties may be partially solved through a number of pilots that will start soon. Interviewee I1 explained: *"That is going to be a big learning experience for the manufacturers, but also for local HTA bodies in terms of what data do we see and what else do we expect or get from manufacturers because they decide what extra data they provide."* Indeed, three interviewees mentioned that many aspects will become clear over the next years or remain to be seen.

Alignment on evidence requirements would be needed for HTA bodies to accept the outcome of a joint clinical assessment

Despite the legislation still being in its infancy, interviewees could foresee potential challenges and implications. A recurrently mentioned challenge relates to the alignment of evidence requirements. Interviewee I1 explained that if there are *"fundamental differences in how NICE looks purely at the clinical data vs. G-BA looking at same set of data, that is where alignment needs to take place"*. In a similar vein, interviewee I2 said that *"those fundamental things of what evidence countries accept and in what context, that's going to be crucial for the subsequent implementation"*.

Similarly, two interviewees foresaw potential challenges related to variable acceptance of RWD/RWE in various geographic contexts. Interviewee I1 explained: *"from my experience a lot of local RWD/RWE gets submitted to the HTA body of a particular country [...] You would need to harmonize the operation of let's say a registry across Europe. That is difficult because the treatment realities of the patients are different in different countries."* While standardisation of EU guidelines could potentially mitigate such differences, industry expert I1 could foresee related challenges: *"[standardisation] is more a physician-led discussion and not so much a political one. This is much more complicated because even within a single country there may be disagreements on how patients should be treated."*

European guidelines and standards could potentially mitigate fundamental differences between HTA bodies in a joint clinical assessment

According to interviewee I3, such guidance would be valuable: *"they [manufacturers] will need to prepare their RWE platform or package for their submission on an EU level, and I do think that the European*

Commission or the working group that's assigned to this needs to think about this". In a similar vein, interviewee I2 noted that guidelines *"would also then benefit companies because the key concern that many of the of our clients have is that this [joint clinical assessment] is just going to be an extra hurdle [...] I think what you want to try and do is to minimize the additional information that needs to go to other countries"*. To mitigate such a potential increase in administrative burden, the transferability of joint assessments to national HTAs should be optimised.

Alignment on PICO would be needed for HTA bodies to accept the outcome of a joint clinical assessment

Interviewees were generally aligned on what would be needed for national HTA bodies to accept the outcome of a joint clinical assessment. According to interviewee I2, *"the key alignment to make a meaningful sort of joint assessment and also to allow it to be used by other countries is that there needs to be alignment on the PICO [patient population, intervention, relevant comparators, outcomes (e.g. relevant endpoints)]."*
While the alignment on comparator was mentioned by all interviewees, alignment on the outcomes (e.g. clinical endpoints) was mentioned in two interviews. However, alignment on the intervention was not mentioned by another interviewee. Moreover, interviewee I1 did not foresee any challenges in the near future related to the patient population: *"at the moment, in the context of cell and gene therapies, these are relatively small overall populations. Alignment is therefore not going to be an issue at the moment"*.

How joint clinical assessments may affect RWD/RWE uptake remains unknown

Interviewees gave mixed responses when asked to what extent the joint clinical assessment may impact RWD/RWE usage. While interviewee I1 said *"I don't think that a joint clinical assessment is going to trigger more requests from HTA bodies. So still in the case of joint clinical assessments, it is still up to the manufacturer to decide what RWE study to run in what countries and then make that available"*, interviewee I2 said that *"[...] I think the EU regulation will give a sort of an injection into relooking at all the guidance documents that exist, which I think also make a realization that many of them were outdated or not specific enough"*.
However, interviewee I2 also noted that the extent to which HTA bodies would evaluate their RWD/RWE guideline and act upon it, may substantially differ. The interviewee explained: "*they [NICE and G-BA] have the expertise in house [to interpret RWD/RWE]. But if you look at smaller countries, they will not have the expertise, and that's again, I think where the joint EU HTA may come into play by addressing and filling that gap [...]*".

While many aspects of the joint assessment remain unknown, some relationships could be identified

Several elements relevant to RWD/RWE uptake emerged during the interviews. The aggregate findings and their relationships were visualised in a network (Figure 5.2).

From the conceptual network, it appears that there are two boundary conditions for optimal transferability of joint clinical assessment outcomes to national HTA bodies. The first condition is alignment on the different aspects of the PICO framework. The second condition is then alignment on the evidentiary requirements of evidence that is submitted for these PICO aspects. EU-wide consortia can potentially help facilitate achieving such alignments in EU guidelines, which would also help manufacturers focus on synthesising evidence that meets the required standards.

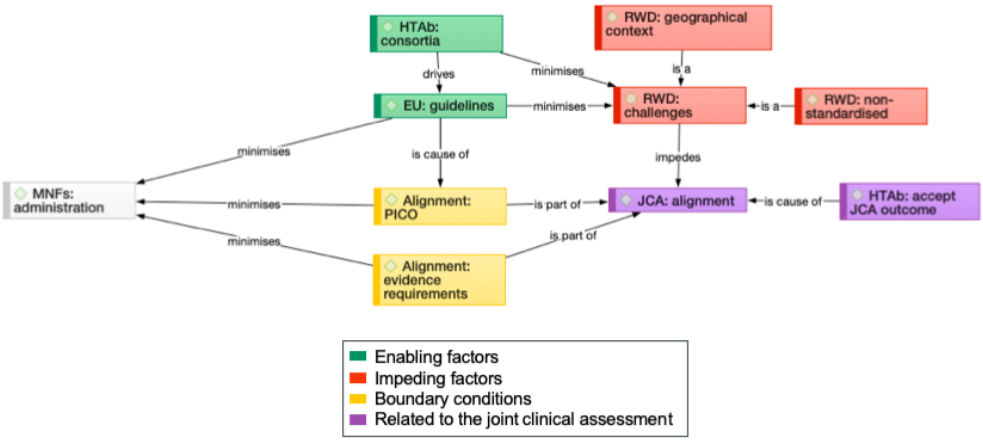


Figure 5.2. Alignment in joint clinical assessment: conceptual network. Red: impeding factors. Green: enabling factors. Yellow: boundary conditions. Purple: Factors related to the joint clinical assessment. Orange: manufacturers. EU: European Union. HTA: health technology assessment. HTAb: HTA body. MNFs: manufacturers. RWD: real-world data.

5.3. The interviews provided key insights to validate and interpret preliminary findings

Based on the results presented in the previous sections, the following key insights were synthesised (Table 5.1)

Table 5.1. Key insights from the semi-structured interviews. HTA: Health technology assessment. RWD: Real-world data. RWE: Real-world evidence.

Sub-question	Insight	Description	Reference
1	1	While most challenges encountered in gene therapy HTAs are shared with other therapy types, curative potential is a key differentiating factor	Section 5.1
	2	HTA bodies have no frameworks or payments schemes specifically tailored to gene therapies, but these may emerge over time	Section 5.1
2	3	While RWD/RWE has the potential to support decision making, it is not a silver bullet.	Section 5.1
	4	RWD/RWE currently mostly supports areas like the patient population, the national history of the disease and informing the comparative effectiveness.	Section 5.1
	5	RWD/RWE is here to stay and also will have an increasing impact, particularly in supporting effectiveness or supporting the validity of certain endpoints.	Section 5.1

continues on next page

3	6	Enablers for the uptake of RWD/RWE by HTA bodies include technological advancements of clinical systems and techniques, as well as international initiatives and collaborations.	Section 5.1
	7	Accelerated regulatory pathways also appear to drive RWD/RWE usage but the impact is limited as HTA bodies have different evidence requirements	Section 5.1
	8	Data privacy, lack of standardisation, risk of bias in the data and lack of prescriptive guidance act as substantial barrier to increase the adoption of RWD/RWE	Section 5.1
4	9	The uncertainties on joint clinical assessments appear to mainly relate to what national HTA bodies will do with the clinical assessment outcome and to what extent it satisfies their evidence requirements	Section 5.2
	10	The key alignment to make a meaningful sort of joint assessment and also to allow it to be used by other countries is that there needs to be alignment on the PICO aspects	Section 5.2
	11	Alignment on the fundamental evidence requirements needs to take place (e.g. what kind of evidence and in what context)	Section 5.2
	12	Potential challenges for a joint clinical assessment include the geographical context of RWD, as well as the lack of standardisation in the data	Section 5.2
	13	EU-wide guidelines may trigger issuance of updated national (prescriptive) guidelines on RWD/RWE usage, which would mitigate a potential increase of administrative burden for manufacturers	Section 5.2

6 Discussion

This study aimed to derive implications for alignment in future joint clinical assessments based on RWD/RWE usage in gene therapy HTAs. This chapter serves to interpret and discuss the findings and their implications. First, the research findings will be critically discussed. Second, the limitations of this study are described. Finally, the scientific and practical contributions are highlighted.

6.1. Discussion of findings

Throughout this thesis, preliminary findings and key insights were summarised at the end of each data gathering and data analysis chapter. Explicitly recapping the findings per research question served to keep an overview of what research questions were addressed in a chapter but it also yields an impression of how these findings changed throughout the research.

While special considerations apply to gene therapy HTAs, they are often not unique to this therapy type

The first step in performing this research was to gain insights into the recent developments in gene therapy HTAs, which were relevant for answering sub-question 1 (*How are gene therapies currently assessed in HTAs?*).

In accordance with findings from the literature, key uncertainties in gene therapies mainly concerned the lack of long-term data on clinical effectiveness and cost-effectiveness. However, considerations relating to the ethical and legal aspects appeared to bear counterweight to such uncertainties in gene therapy HTA outcomes. This has also been described in literature (Pochopień *et al*., 2021; Ten Ham *et al*., 2021). Yet, to what extent these considerations allow for increased acceptance of uncertainty, was unclear. Moreover, the outcomes of this leniency appears to differ between HTA bodies; whereas orphan designation leads to a positive recommendation by default in the G-BA framework, it does not lead to a special program in the NICE HTA framework. higher acceptability of uncertainty.
Interviewees emphasised that most considerations and challenges are not unique to gene therapies other than curative potential. Interviewee A2 explained: *"[...] I think the key considerations are not that different, right? Because it is the same framework that they are being assessed in"*. This finding corresponds with the work by Faulkner *et al*. (2019), who note that it appears that HTA bodies are not applying technology-specific frameworks. Curative potential alone has therefore not led to HTA bodies developing frameworks tailored to gene therapies. That is not to say that HTA bodies are not willing to change, as interviewee I1 noted *"They [Zorg Instituut Nederland] basically say that they need to adjust their frameworks. I can see something similar happening for some of the HTA bodies [...]"*.

Explicating the role of real-world evidence in gene therapy HTAs

Through a retrospective comparative analysis, findings relevant for answering the question *What role does RWD/RWE play in the HTA appraisal process of gene therapies?* were obtained.

RWD/RWE support different areas in NICE and G-BA gene therapy HTAs

A substantial difference in RWD usage between G-BA and NICE became apparent. This difference may be attributed to the methodological framework used by NICE. It inherently allows for more RWD

usage due to the increased number of assumptions compared to the G-BA framework. One could argue that RWD may be beneficial in making assumptions on particular aspects (i.e. insurance claims to derive costs) that may be less relevant in the G-BA gene therapy HTAs.

In line with the HTA archetypes of G-BA (REA) and NICE (CEA), a considerable difference was evident from RWD/RWE used to support economic outcomes. The work by Makady *et al*. (2018) reported similar observations: RWD/RWE was mainly used for delivering evidence on epidemiological information in REAs and long-term effectiveness and costs in CEAs. However, unlike the work by Makady *et al*. (2018), the areas supported in this research mainly comprised information on an external comparator, information on the patient population and effectiveness of the intervention for both G-BA and NICE.
It should be noted that the retrospective analysis by Makady *et al*. (2018) did focus on melanoma drugs and included other HTA bodies. As such, a comparison between these works may be inappropriate. However, as G-BA and NICE are generally considered prime examples of REA and CEA, respectively, a similar result would be expected. Moreover, the authors used IQWiG to represent Germany rather than G-BA. As IQWiG generally uses more epidemiological data (IQVIA internal expertise, 2021), this could, in part, explain the observed difference.

While G-BA and NICE are generally aligned on the appraisal of RWD/RWE, the acceptance rates differ

NICE was found to have a considerably higher RWD/RWE acceptance rate than G-BA (56% and 32%, respectively). Interviewees were generally aligned on NICE being more open to RWD/RWE usage than G-BA. While this difference may be partially attributed to the archetypes, the methodological strictness was also relevant.
Despite these inherent differences, it appeared from the retrospective analysis that G-BA and NICE are generally aligned on the appraisal of RWD/RWE. Two instances were found where the identified appraisal by both HTA bodies were opposites. In line with the identified difference in leniency for RWD/RWE usage, NICE did accept both RWD/RWE sources, and G-BA did not.
From the acceptance rates of RWD/RWE per area supported, both agencies tend to accept RWD/RWE usage to support the effectiveness of the intervention area more often than not, albeit in different ratios. This finding is in contrast with the experience of industry expert I2, who noted that: *"[...] if you look at effectiveness as an area supported, you will not find any case in Germany where that has been accepted. Whereas in the UK, you won't find loads of cases, but there are cases where RWE really is accepted as a source to inform effectiveness."* However, given the lack of empirical evidence, literature does not suffice to validate either of these findings.

Identified factors influencing RWD/RWE uptake in gene therapy HTAs align with the literature

Interviews served to validate initial findings and gain insights needed in answering sub-question 3 (*What factors impact RWD/RWE usage in gene therapy HTA appraisals?* & *What are steps to be taken to extend RWD/RWE usage in HTA appraisals of gene therapies?*)

The existing body of literature prominently mentioned methodological factors impeding widespread RWD/RWE adoption. These include confounding bias, lacking randomisation, standardisation and transparency, as well as a lack of guidance & policies. These factors were not explicitly linked to particular therapy types.
This is in line with the identified knowledge gap where RWD/RWE usage in HTAs is not specifically linked to gene therapies. However, from interviews it appeared that the factors identified from literature also applied to gene therapies.
Contrary to the finding described by Makady *et al*. (2017a), interviewees did not mention varying definitions of RWD/RWE as a barrier to its uptake. Interviewee I2 noted that *"I do not think the alignment*

on the terminology of RWE plays a role per se [...]". As such, the ambiguities around RWD/RWE definitions may have been solved over the years. International collaborations and initiatives between multiple stakeholders are likely to have contributed if this is true. Interviewees described such international collaborations as a key driving force for the uptake of RWD/RWE.
A conceptual network of factors that influence RWD/RWE uptake was established based on gathered interview data. From the network, initiatives indeed appear to be able to mitigate some of the identified barriers. However, other barriers appear to be unsolvable by nature and maybe mitigated over time through technological advances. Interestingly, however, Hampson *et al*. (2018) noted that such advances, i.e. the increasing ability to generate and interpret large amounts of data, have also lead to the increasing complexity of RWE study designs. This, in turn, may form a barrier to uptake, as HTA bodies may not know how to interpret the data. As such, HTA bodies themselves may end uop being a barrier to RWD/RWE uptake. Then again, international collaborations such as the GetReal Institute may help knowledge sharing (i.e. best practices) between HTA bodies to mitigate such barriers.

Steps to be taken to increase RWD/RWE uptake mainly relate to mitigating identified barriers

Based on the above, the following steps are likely to contribute to increased uptake of RWD/RWE usage among HTA bodies:

- Foster and expand ongoing initiatives and inter-organisational collaborations to facilitate knowledge transfer and knowledge sharing (i.e. on best practices on RWD/RWE usage) between HTA bodies and other relevant stakeholders
- The output of these collaborations should be ideally be formalised in two ways. Guidelines for HTA bodies themselves on interpreting the increasingly complex RWD/RWE would help HTA bodies embrace and appreciate RWD/RWE in their assessment frameworks. This is particularly valuable for HTA bodies with insufficient in-house resources.
  In addition, HTA bodies should be issuing prescriptive guidance for gene therapy manufacturers on requirements of RWD/RWE sources
- Finally, within the EU-wide efforts to introduce standardisation of data collection, data control infrastructure should be expanded to ensure high-quality data, which would facilitate increased adoption of RWD/RWE

Inter-organisational alignment on assessment elements and evidentiary requirements is needed

This section will discuss the implications of these findings for achieving alignment between these HTA bodies. However, merely suggesting a general alignment of gene therapy HTA outcomes would undermine the multifaceted nature of HTAs and their complexity arising from interpretation in the national standard of care. Indeed, the joint clinical assessment legislation is limited to the clinical assessment of a health technology (European Commission, 2021). To optimise the transferability to national HTA bodies, alignment on assessment elements and evidentiary requirements is necessary.

Alignment on assessment elements

The majority of similarly appraised sources supported evidence on the outcome (effectiveness of the intervention) and evidence on an external comparator. Relatively few unique sources were included for the former, which is expected since these often comprise effectiveness studies by the manufacturers. However, many unique sources were submitted to both G-BA and NICE for evidence of an external comparator. This makes sense, considering the varying appropriateness of external comparators in the context of the local standard of care. This illustrates that a substantial barrier in alignment may arise from the comparative aspect in the joint clinical assessment.

Similarly, both agencies mainly considered unique sources for the patient characteristics aspect of the PICO framework. Again, this makes sense given the local standard of care.

These findings were validated in interviews, from which it indeed appeared that alignment on the comparator aspect was deemed most challenging. While alignment on the outcomes (e.g. clinical endpoints) and alignment on the intervention were considered less challenging, alignment on the patient population was not considered a considerable barrier in the foreseeable future. This is due to the small patient populations that gene therapies generally target. However, it should be noted that this may change if gene therapies would target more significant patient populations in the future.

Alignment on evidence requirements

To improve synergy among HTA bodies, Wang *et al*. (2018) recommend using real-world evidence to support relative effectiveness assessments in HTAs. Given the variable RWD/RWE acceptance rates for different areas supported in HTA reports, this recommendation appears too broad.
In line with the work by Kanavos *et al*. (2019), it appears that HTA bodies' risk tolerance and attitudes towards RWD/RWE usage may form a potential barrier to the feasibility of joint clinical assessments. HTA bodies should have a discussion and align on evidentiary requirements (e.g. what type of RWD/RWE is accepted and in what context).
Similarly, alignment should be reached on requirements in terms of the geographical context of the submitted RWD/RWE. While EU-wide patient registries could potentially mitigate regional differences, patient data privacy may pose challenges in its feasibility.
Consortia and EU-wide initiatives such as EUnetHTA may play a pivotal role in establishing such prescriptive guidelines and bringing relevant stakeholders together.

Practical implications for future joint clinical assessments of gene therapies

Despite many unknowns, it appears that the envisioned joint clinical assessment may evolve in a hierarchy structure, where a single actor or committee determines the assessment outcome (European Commission, 2021). However, establishing a framework to guide this assessment would likely involve aligning the national HTA bodies. Considering the diversity of national HTA bodies and other stakeholders, including decision-makers, policymakers and manufacturers, the HTA landscape may be interpreted as a network (Bruijn and Heuvelhof, 2018).

Achieving consensus in the HTA network Problems in the context of networks and varying interests of actors may often be considered 'unstructured' or 'wicked' (Rittel and Webber, 1973). The authors specify multiple characteristics of such problems, one of which is the lack of a definitive formulation for the problem. However, one could argue that there appears to be a clear definition for the problem (e.g. establishing a joint assessment framework that is transferable to national HTA bodies). In a similar vein, there may be a true solution to the problem (e.g. achieving alignment of actors on the assessment criteria and evidentiary requirements) that can be tested (i.e. through the extent to which the outcome is integrated into the HTA process national HTA bodies). Wicked problems do not have true solutions or an ultimate test of a solution (Rittel and Webber, 1973). As such, it appears that despite the different interests of these actors, the problem at hand may not be considered 'unstructured' or 'wicked'.
To formulate a way forward, the decision tree on policy analysis support as presented by Enserink *et al*. (2010) is used. The problem must be solved in a consensual process of multiple HTA bodies with aligned interests (e.g. a joint clinical assessment outcome that can be used directly in national HTA body assessments). However, their consensus on the technical information may not be achieved, following the identified differences in HTA characteristics and RWD/RWE amenability. Considering these factors, an interactive analysis approach appears to be most suitable for further research, where stakeholders are involved in defining the scope and analysis tools.

Interviews made it apparent that ongoing collaboration efforts between national HTA bodies have been slow, resulting in incremental change. The same may be expected for how the output of the above-mentioned interactive analysis approach would be used.
Knowledge diffusion models may be leveraged to promote future research output and enhance its implementation into practice. One of these models is the Ottawa Model of Healthcare Research Use, as presented by Logan and Graham (1998). This non-linear framework consists of six elements that explain the uptake of knowledge by actors: the practice environment (1), potential adopters (2), the innovation (3), strategies for the transfer of the innovation into practice (4), the evidence adoption (5), and health-related outcomes (6). It should be noted that these elements are interrelated, reflecting the complexity of the knowledge transfer process (Logan and Graham, 1998).
While applying this framework to formulate implementation strategies is considered out-of-scope, it does display the need for knowledge transfer and translation strategies (element 4) in the studied context.

Knowledge transfer strategies Transfer of knowledge (i.e. technical information or best practice) is generally aided through easily understandable tools (Formoso *et al*., 2022).
While an HTA in itself can be considered a tool for knowledge transfer (Battista, 1996), tools like the EUnetHTA Core Model® (EUnetHTA, 2015) may be particularly valuable in transferring codified knowledge (Newell *et al*., 2020). This emphasises that international collaborations and initiatives may play a key role in facilitating such alignments and the corresponding knowledge transfer (Pichler *et al*., 2019; Wang *et al*., 2018). They may also play a pivotal role in developing EU guidelines that capture the outcomes of such alignment efforts.

Inter-organisational knowledge transfer and knowledge sharing may also lead organisational learning (Lane and Lubatkin, 1998). This could help national HTA bodies agree on the knowledge involved in achieving alignment on assessment elements and evidence requirements.
However, organisational learning capabilities may differ substantially between national HTA bodies. From the seminal work by Grant (1996), the learning capabilities of an organisation is understood to relate to the characteristics of the recipient and donor organisations, as well as the nature of knowledge and the inter-organisational dynamics (Easterby-Smith *et al*., 2008). Given that considerable differences in resources and existing knowledge are available to HTA bodies (Kálo *et al*., 2016), their absorptive capacity may differ considerably (Newell *et al*., 2020). This point was also raised by interviewee I2, who noted that "*if you look at smaller countries, they will not have the expertise*".
These differences may cause friction, as the perceptions of the outcomes relative to the required input may differ per HTA jurisdiction. Moreover, such differences may affect the priority that national HTA bodies give to international knowledge-sharing initiatives. This, in turn, could impede the proposed collaboration (Gray, 1985).

Transferability to national HTA bodies Similarly, the attitudes of national HTA bodies towards alignment efforts may be driven by the extent to which the joint clinical assessment outcome is compatible with their existing appraisal frameworks. Low compatibility may alter the beliefs of an HTA body on the benefits of joint assessments, thereby impeding inter-organisational collaboration (e.g. alignment) efforts with other actors (Gray, 1985).
Some HTA bodies may be able to innovate their HTA process to make it compatible with future joint assessment outcomes. Others, however, may be less suitable for process innovation due to the variability mentioned above in resources and existing knowledge among HTA bodies.

Converting implications into actionable insights

Following the aspects described above, achieving inter-organisational alignment may prove complex. The following aspects may be taken into account in further research and defining strategies to achieve

inter-organisational alignment:

- An interactive analysis approach appears most suitable for further research. However, adoption of this research output may be slow, owing to varying cultural, political and economic contexts of the national HTA bodies. Therefore, knowledge diffusion models may be leveraged to formulate strategies for increased adoption.
- Knowledge transfer and sharing are an essential aspect of organisational learning and should be stimulated through appropriate tools
- Organisational learning and incentive for HTA process innovation may vary per accessibility to resources, existing knowledge and compatibility with existing frameworks. Such differences and their effect on collaboration should be considered in formulating alignment strategies.

Gene therapy manufacturers should account for differences in RWD/RWE usage by HTA bodies in their market access strategy

While the previous sections discussed the presented findings from an inter-organisational (multi- HTA body) perspective, this section aims to interpret the findings from the perspective of a gene therapy manufacturer. These insights will be relevant in answering the final sub-question (*What are the implications of RWD/RWE usage by HTA bodies for gene therapy manufacturers?*)

Manufacturers should acknowledge varying methodological strictness of HTA bodies when submitting RWD/RWE in gene therapy HTAs

The presented work RWD/RWE reports varying acceptance rates between HTA bodies. Gene therapy manufacturers should consider these differences when submitting evidence for gene therapy HTAs to minimise inefficiencies. Inefficiencies in this context comprise submitting evidence that, based on retrospective analyses, has a high likelihood of being negatively appraised.
To this point, the presented research provides relevant insights:

- From the retrospective analysis, RWD/RWE was found to have the lowest acceptance rate if it supports evidence on an external comparator. Based on exclusion rationales in the use cases, manufacturers should be critical towards the data they submit, i.e. does it provide information to substantiate its relevance and is the data set extensive. Especially for G-BA, this appears to be a confounding factor
- RWD/RWE supporting the effectiveness of the intervention is relatively often accepted by both NICE and G-BA. This makes sense, given that there may not be alternative data available to provide this evidence. Manufacturers should expand conversations with HTA bodies to understand when RWD/RWE sources are considered ‘rich’ (sufficient data available) to increase the acceptance rates further. Similarly, manufacturers should be critical to the appropriateness of the RWE study design (e.g., methods used in line with regulatory standards and current practices).

Manufacturers should be actively involved in shaping future joint clinical assessments of gene therapies

While the points mentioned above highlight the need for critical reflection of manufacturers on the RWD/RWE that they submit, this reflection should ideally be guided by conversations with HTA bodies. As discussed in previous sections, prescriptive guidelines would be a key enabler for RWD/RWE uptake. This would drive a more mature RWD/RWE field with an increased understanding of its advantages and disadvantages in specific contexts.
The call for guidelines becomes increasingly essential with the prospective joint clinical assessments of gene therapies. While many aspects are still unknown, a higher administrative burden for manufacturers may be lying in wait. Based on the established conceptual network, a burden may be mitigated

by EU-wide prescriptive guidelines and inter-organisational alignment on both the assessment framework and evidentiary requirements. As such, gene therapy manufacturers need to take an increasingly active role in the innovation system by participating in such conversations and facilitating enabling initiatives (Edquist, 2011). Future research could explain such an innovation system to understand better how innovation and knowledge sharing can be optimised.

6.2. Discussion of research limitations

The limitations of this study can be categorised as relating to the research design or the reliability and validity of the presented data and analyses.

Research design

HTA reports published by two HTA bodies were included in the retrospective analysis, one of which (NICE) is not part of the EU. Therefore, the scope of this comparison may give a limited view, and the results may not be representative of the other EU member states. Moreover, in the retrospective analysis, only gene therapy HTAs were considered, while the context in which the findings were interpreted comprised EU-wide legislation, not specific for gene therapies. Therefore, the generalisability of the derived implications for the EU joint clinical assessment legislation may also be limited.

Another limitation relates to the definition of RWD/RWE and the categorisation of the areas supported in the retrospective analysis. In literature, the varying definition of RWD (and consequently RWE) is well-recognised (Makady *et al.*, 2017a). The results should be interpreted with this in mind, as it may change with a different RWD/RWE definition. To minimise such potential misalignments, broadly accepted definitions of RWD/RWE were adopted in this research (IMI GetReal, 2016).
The categorisation of areas supported was done following the PICO framework with subcategories defined following IQVIA internal research (2021). While the categories are backed by literature, they may be somewhat subjective. Similarly, the setting in which the research was performed may not be generalisable. While an academic institution led the research, the researcher benefited from IQVIA internal expertise. Procedures such as the categorisation of data may therefore not be representative of the academic context but rather the commercial context.

Finally, the theories and frameworks utilised for interpreting the results of sub-questions 4 and 5 were not derived through a systematic literature review. Due to time constraints, such a literature review was considered out of scope. Future research may explicate whether better applicable frameworks are available in the literature. Future works should also build on the presented work and apply the framework that is deemed most suitable to formulate implementation strategies in the context of innovation systems.

Reliability and validity of data

Data for the retrospective, a comparative analysis was extracted by a single researcher. As such, data may have been missed, or interpretation of data may have resulted in the wrong categorisation of RWD/RWE usage. Similar works mitigated similar uncertainties by calculating an inter-rater reliability (Vreman *et al.*, 2019; Makady *et al.*, 2018). However, due to the absence of a second researcher, this metric could not be established in the presented research. An alternative approach to data validation was therefore used, where three data sources were combined to (1) derive the completeness of extracted data using the algorithm derived and (2) interpretation of identified RWD/RWE appraisals (Appendix A6). By triangulating the data, higher confidence in the results is obtained (Sekaran and Bougie, 2016).

Another limitation relates to the data collected through semi-structured interviews. The limited sample

size and background of interviewees may not represent the studied context, which may reduce the generalisability of the performed research.
Moreover, qualitative data analysis of the interviews is prone to the researcher's subjectivity. Since one researcher performed the research, the interjudge reliability could not be established (Kassarjian, 1977). Therefore, the reliability of the assigned quotations, codes, and categories remains unknown.

6.3. Discussion of contributions

Scientific contributions

The potential of RWD/RWE in gene therapy HTAs have increasingly been recognised and embraced by HTA bodies. However, the existing body of literature fails to deliver empirical evidence on the extent to which RWD/RWE is used or its role in gene therapy HTAs.
By addressing the identified knowledge gap, the presented research makes multiple scientific contributions:

- From the literature review, two relevant scarcities in literature became apparent. The number of works that deliver empirical evidence on HTAs of ATMPs is scarce. Similarly, works that deliver empirical evidence on RWD/RWE usage in HTAs were equally scarce. The presented empirical evidence adds to the current knowledge base in two ways; it provides empirical evidence on HTAs of gene therapies (1) and RWD/RWE usage in gene therapy HTAs (2). Consequently, this research contributes to the existing knowledge base in a third way: to the best of the author's knowledge, no other work reports on a combination of these two aspects
- In addition to the above, the presented research utilised a novel approach for the retrospective analysis. Whereas previous works (i.e. Makady *et al*. (2018) and Lee *et al*. (2021)) used a data extraction form, this work combined data from different sources, including a search algorithm, which, to the best of the author's knowledge, has not been applied in this context in literature. Moreover, the presented research used interviews to validate findings from the retrospective analysis, while the works as mentioned earlier did not
- The presented work also contributes to the existing literature on joint clinical assessments. Allen *et al*. (2017) note that "in order to move forward to a more harmonised HTA environment within Europe, it is first necessary to understand the variation in HTA practices within Europe". The research, therefore, contributes to an increased understanding of the variation in HTA practices. In doing so, the work delivers insights for future joint clinical assessments from a novel perspective; whereas published works (i.e. Kisser *et al*. (2021) and Vreman *et al*. (2020)) do deliver empirical evidence and propose the need for alignment, none of them substantiates this need from the presented perspective (e.g. differences in RWD/RWE usage in gene therapy HTAs by different HTA bodies)
- Finally, the findings in this work were interpreted using theories of knowledge management and knowledge diffusion. Similarly, the work proposes different theories and frameworks, which may serve to interpret the generated insights for joint clinical assessments in future research.

Managerial contributions

According to Pochopień *et al*. (2021) improving HTA methods "will be driven by academics and not HTA agencies, as the latter tend to be conservative, asking for increasingly large volumes of evidence, without an appetite for innovative methodology". By building on the above-mentioned scientific contributions, this work contributes twofold to the managerial aspects of driving HTA process innovation.

- While the observed difference in methodological strictness and implied need for alignment in joint

clinical assessments has already been covered in literature (Kisser *et al*., 2021; Vreman *et al*., 2020), the generated insights provide nuance that other works do not. To this point, the work proposes that, based on differences in RWD/RWE usage, aspects of the PICO framework may be more difficult to align on. More specifically, while the comparator aspect presents the most considerable barrier to alignment, alignment on patient characteristics is not considered a barrier due to the small patient populations that gene therapies target.
These nuances may aid a more targeted strategy development to optimise transferability of the European joint clinical assessments outcome to national HTA bodies. To this point, enabling knowledge management practices were highlighted (i.e. the need for implementation strategies and the need for knowledge transfer and organisational learning, as well as the notion that future alignment strategies should take differences in existing knowledge and available resources of HTA bodies into account for better collaboration)

- The second managerial contribution of this work is that it offers insights for gene therapy manufacturers on differences in RWD/RWE usage by HTA bodies. Although it appeared from the interviews that industry experts have a sense of the recent developments, empirical evidence, as presented here, is lacking. Based on exclusion rationales of RWD/RWE, this work proposes that when manufacturers consider submitting RWD/RWE, they should be critical towards the data that they submit (i.e. does it provide information to substantiate its relevance and is the data set extensive) and the appropriateness of the RWE study design (e.g. are methods used in line with regulatory standards and current practices). To guide such reflections and prevent the potential additional administrative burden from joint clinical assessments, the work proposes that manufacturers should take an increasingly active role in the innovation system.

Societal contributions

Gene therapies have the unprecedented potential to target the underlying cause of genetic conditions and potentially prevent, treat or cure genetic conditions and hereditary diseases in the future. Increased patient access to such innovations is therefore considered of societal relevance.

While HTAs are one of the three essential domains for the adoption of innovative technologies (Gardner and Webster, 2016), established HTA frameworks may no longer suffice. Given the prospect of an increasing number of cell & gene therapies seeking market access in the coming years, tailoring HTAs to gene therapies is becoming more urgent.
Joint clinical assessments of these technologies may present opportunities to allow for faster and more uniform assessments, as well as improved patient access to innovative health technologies (European Commission, 2021; Kanavos *et al*., 2019). These insights from this research may contribute to a better understanding of what would be needed to optimise the transferability of the European joint clinical assessments to national HTA bodies. By optimising this process, EU-wide patient access to these curative therapies may be increased in the future.

7 Conclusions

This chapter provides conclusions to the research questions stated in chapter 1, as well as recommendations for future research. Subsequently, the link to the researcher's study programme will conclude this chapter.

7.1. Conclusions

This study aimed to identify feasibility challenges for EU-wide gene therapy joint clinical assessments, based on the current and future role that real-world data and real-world evidence play in HTA outcomes. A critical literature review laid the foundation for further research, including a retrospective, comparative analysis of gene therapy HTAs published by G-BA and NICE between December 2015 and November 2021. Interviews with industry experts and academics validated initial findings and explored potential implications for the recently approved joint clinical assessment legislation.
To answer the guiding research question: "What are the implications of RWD/RWE usage by HTA bodies for achieving alignment in future joint clinical assessments of gene therapies?", the output of five sub-questions was combined. These questions will be answered separately below.

Sub-question 1: *How are gene therapies currently assessed in HTAs?*

The appraisal process of gene therapies appears to be similar to that of other therapy types. Most challenges and considerations in gene therapy HTAs are not unique, and HTA bodies have no frameworks or payments schemes tailored explicitly to gene therapies. However, this may change in the following years due to increased attention and experience.

Based on the existing literature, the key challenges associated with gene therapies mainly relate to the uncertainty regarding long-term safety and -effects and cost-effectiveness. While these challenges may not be unique to gene therapies, appraising the curative potential of these therapies appears to trouble existing HTA frameworks.
Such difficulties could not be derived from the nineteen gene therapy HTA reports retrieved for G-BA and NICE. However, both agencies acknowledged the lack of long-term data in their appraisals. While both adopted novel payment schemes to mitigate evidentiary uncertainties, such methods are not unique to gene therapies. Similarly, while special considerations like orphan disease were often found to apply in gene therapy HTAs, they are also not unique to this therapy type. However, the influence that such special considerations' influence on the HTA process differs between HTA bodies.

Sub-question 2a: *What role do RWD/RWE play in the HTA appraisal process of gene therapies?*

While RWD/RWE mainly play a supportive role in gene therapy HTAs, acceptance rates and areas supported differ between G-BA and NICE.

In line with the HTA archetypes of NICE and G-BA, the RWD/RWE was found to support different areas in the respective gene therapy HTAs. However, from the higher total volume of RWD/RWE usage and the higher acceptance rate of RWD/RWE by NICE, NICE is more amenable to RWD/RWE. As such, RWD/RWE mainly serves to support areas in the HTAs (e.g. patient population, the national history of the disease and informing the comparative effectiveness), but the extent varies per HTA body. This

difference may, in part, be attributed to the NICE methodology, which inherently allows for more RWD usage due to the increased number of assumptions when compared to the G-BA framework.

Sub-question 2b: *How has RWD/RWE usage in gene therapy HTAs evolved?*

While RWD/RWE has gained traction among the HTA community, the uptake in practice appears to be lagging. From the retrospective analysis and use-cases, no clear trend was visible for the volume of RWD/RWE usage over time for both G-BA and NICE. The same was true for areas supported and acceptance rates. Therefore, it remains unclear how RWD/RWE usage has evolved, but no increased uptake was observed contrary to the increased interest.

Sub-question 3a: *What factors impact RWD/RWE usage in gene therapy HTA appraisals*

In line with the earlier observation that most challenges encountered in gene therapy HTAs are not unique, the RWD/RWE usage was not found to differ for gene therapy HTAs.
The main methodological factors impeding widespread RWD/RWE adoption include confounding bias, lacking randomisation and a lack of transparency. Moreover, the lack of (prescriptive) guidance was a considerable barrier to uptake. Enablers included technological advancements of clinical systems and techniques and international initiatives and collaborations.

Sub-question 3b: *What are steps to be taken to extend RWD/RWE usage in HTA appraisals of gene therapies?*

Steps to increase RWD/RWE uptake mainly relate to the impeding factors. Based on literature and interview input, the following steps were proposed:

- Foster and expand ongoing initiatives and inter-organisational collaborations (i.e. EUnetHTA, GetReal Institute) to facilitate knowledge transfer and knowledge sharing (i.e. on best practices on RWD/RWE usage) between HTA bodies and other relevant stakeholders
- The output of these collaborations should be ideally be formalised in two ways: Establish guidelines for HTA bodies themselves on how to interpret the increasingly complex RWD/RWE (1), and HTA bodies should be issuing prescriptive guidance for gene therapy manufacturers on requirements of RWD/RWE sources (2). Regulatory agencies and HTA bodies have started mandating continued (real-world) evidence development through registries. Expanding and integrating such regulations will extend the RWD/RWE usage and uptake in HTAs
- Finally, within the EU efforts to introduce standardisation of data collection, data control infrastructure should be expanded to ensure high-quality data, which would facilitate increased adoption of RWD/RWE

Sub-question 4: *What are the implications of RWD/RWE usage by HTA bodies for their alignment in the assessment of gene therapies?*

Based on the RWD/RWE usage by G-BA and NICE, a substantial barrier in alignment may arise from the comparative aspect in HTAs. Given the varying appropriateness of external comparators in the context of the local standard of care, it is expected that an RWD/RWE source on a particular comparator may be considered relevant in one country but not in another.

The number of opposing RWD/RWE appraisals was limited compared to the number of similarly appraised RWD/RWE sources. Following the PICO (e.g. patient population, intervention, comparator and outcome) framework, most similarly appraised sources supported the comparator aspect and the

outcome of the PICO framework. Unlike the comparator aspect, HTA agencies did not consider a substantial amount of additional unique RWD/RWE sources for the outcome. This is somewhat encouraging for potential alignment on the outcome aspect.

Sub-question 5: *What are the implications of RWD/RWE usage by HTA bodies for gene therapy manufacturers?*

The retrospective analysis shows that RWD/RWE has the lowest likelihood of acceptance if it was submitted to support evidence on an external comparator. On the other hand, RWD/RWE supporting the effectiveness of the intervention was found to be relatively often accepted by both NICE and G-BA. Gene therapy manufacturers may consider these differences when submitting RWD/RWE in HTAs.
To increase the likelihood of acceptance, this research proposes that gene therapy manufacturers should be critical towards the RWD/RWE that they submit (i.e. does it provide information to substantiate its relevance and is the data set extensive) and towards the appropriateness of the RWE study design (e.g. are methods used in line with regulatory standards and current practices).
To further increase the acceptance rates, manufacturers should expand conversations with HTA bodies to understand better when RWD/RWE sources are considered 'rich' (sufficient data available). Based on the established conceptual network, a burden may be mitigated by EU-wide prescriptive guidelines and inter-organisational alignment on both the assessment framework and evidentiary requirements. Knowledge sharing and knowledge transferring appear key in establishing these. As such, gene therapy manufacturers should continue and expand collaboration in inter-organisational initiatives.

Research question: *What are the implications of RWD/RWE usage by HTA bodies for achieving alignment in future joint clinical assessments of gene therapies?*

The presented retrospective comparative analysis unveiled differences in RWD/RWE usage in gene therapy HTAs. Substantial variations in how HTA bodies assess the added benefit of a new health technology may pose feasibility challenges to joint clinical assessments. In this context, inter-organisational alignment on the comparator aspect of the future joint clinical assessments will be a crucial challenge. Moreover, alignment on the evidentiary requirements is deemed necessary to optimise transferability of the joint clinical assessment outcome to national HTAs.

Achieving alignment in a multi-stakeholder environment may be challenging, especially considering the observed methodological differences between HTA bodies. Moreover, differences in available resources and existing knowledge result in differences in absorptive capacity among HTA bodies.
Such differences should be considered in future alignment strategies, as they may impede collaboration efforts. This is important because international collaborations and initiatives play a key role in facilitating alignments and the corresponding knowledge transfer. The existing collaborations should be leveraged, and new initiatives stimulated to facilitate joint clinical assessments of gene therapies.

7.2. Future research

Based on the presented research's scope, results, and limitations, several opportunities for future research were identified.

- Due to time constraints, the scope was limited to gene therapy HTA reports published by G-BA and NICE. To increase the generalisability, the scope should include other HTA bodies, such as the French HTA body Haute Autorité de Santé or the Dutch Zorginstituut Nederland
- Data for the retrospective, a single researcher extracted a comparative analysis. As such, data

may have been missed, or interpretation of data may have resulted in the wrong categorisation of RWD/RWE usage. Therefore, it is highly recommended that another researcher repeat the research so that inter-rater reliability can be established

- Similar to the previous point, the qualitative data analysis of the interviews is prone to the researcher's subjectivity. It is recommended that a second researcher reviews the transcripts and assigns codes and categories to establish the interjudge reliability
- While semi-structured interviews served to validate the findings presented in this research, the limited sample size of interviewees and their background may not represent the studied context. Therefore, it is recommended to validate the findings in interviews with experts from various backgrounds, including HTA bodies
- Finally, future research could build on the insights presented in this work. To this end, several 'stepping stones' were provided. An interactive analysis approach is suggested for further research, where stakeholders are involved in defining the scope and analysis tools. Moreover, future research could apply theoretical frameworks to formulate future implementation strategies. Finally, future research could explicate the innovation system for joint clinical assessments to understand better how innovation and knowledge sharing can be optimised in this context.

7.3. Reflection

In the early stages of this research, I spent a substantial amount of time getting familiar with the thesis topic and scoping the research. The methodological differences between HTA bodies and how HTA outcomes fit differently in their national healthcare systems, was somewhat overwhelming at first. Nonetheless, I think it is essential to understand and appreciate these differences, especially in the context of the recently approved joint clinical assessment legislation. Many aspects of how this concept will be turned into reality remain to be seen. Yet, the observed differences in RWD/RWE usage appear indicative of fundamental differences between HTA bodies.
From this research, it appears that alignment on evidentiary requirements will, at least to a certain extent, be needed to achieve transferability of joint clinical assessment outcomes to national HTA bodies. This would imply an increased alignment on the appraisal of RWD/RWE, which may prove to be difficult. I think that in formulating ways forward, varying methodological strictness and amenability to RWD/RWE by HTA bodies should be taken into account. It appears that this kind of evidence is inevitably going to play an increasing role in gene therapy HTAs and we should prevent a situation where national HTA bodies do not accept a joint assessment outcome because they do not agree with the evidence submitted. This could maintain a disparity of access for patients to these types of products, while harmonisation of HTA bodies should serve to achieve the opposite.

Recently, many developments in HTAs have been going on, with HTA bodies adjusting their frameworks and increasingly recognising the potential and acknowledging the limitations of RWD/RWE. This is a good start, but there is a long road ahead to reach the potential of joint assessments and achieve a sustainable pricing model for countries to provide sustainable patient access to curative therapies. Knowledge transfer and knowledge sharing initiatives will play a key role through academia, industry and policy-makers in the innovation system. I hope that presented research offers relevant insights to potentially contribute to this purpose.

7.4. Link to Management of Technology

According to TU Delft (2019), a thesis submitted for the study programme of 'Management of Technology', should reflect that "graduates learn to explore and understand how firms can use technology to

design and develop products and services that contribute to improving outcomes, such as customer satisfaction, corporate productivity, profitability and competitiveness."
In this thesis, HTAs are considered an essential aspect of enabling access to and facilitating the adoption of innovative health technologies. Indeed, HTAs allow for the transfer of knowledge derived from scientific research to the decision making process (Battista, 1996). Moreover, methodological differences between HTA bodies should be mitigated to optimise the transferability of a joint clinical assessment of gene therapies. Alignment involves managing knowledge processes in an EU-wide multi HTA body environment, linking the research to the Management of Technology study programme. Finally, the findings were interpreted from the perspective of gene therapy manufacturers to derive insights for improving the acceptance RWD/RWE in future gene therapy HTA submissions.

Relevant courses from the curriculum included 'Inter- and intra-organisational decision making' (MOT1452) and 'Leadership and Technology Management' (MOT1524) for formulating the implication of the research. The courses 'Research Methods' (MOT2312) 'Master Thesis Preparation' (MOT2004) provided the opportunity to learn relevant methodologies and best practices in conducting research.

Bibliography

Allen, Nicola, Lawrence Liberti, Stuart R. Walker, and Sam Salek (2017). A comparison of reimbursement recommendations by European HTA agencies: Is there opportunity for further alignment? *Frontiers in Pharmacology* **8**:348.

Alliance for Regenerative Medicine (2019). *Recommendations for Timely Access to Advanced Therapy Medicinal Products (ATMPs) in Europe*. URL: www.alliancerm.org.

Angelis, A., H. Naci, and A. Hackshaw (2020). Recalibrating Health Technology Assessment Methods for Cell and Gene Therapies. *PharmacoEconomics* **38**:12, 1297–1308.

Annemans, L. and A. Makady (2020). TRUST4RD: Tool for reducing uncertainties in the evidence generation for specialised treatments for rare diseases. *Orphanet Journal of Rare Diseases* **15**:1.

APM Health Europe (2021). *ATMPs 'not guinea pigs' as first through the EU's joint HTA - ISCT*.

ASGCT and Pharma Intelligence (2021). *Gene, Cell, & RNA Therapy Landscape - Q2 2021 Quarterly Data Report*.

Banta, David (2009). What is technology assessment? *International Journal of Technology Assessment in Health Care* **25**:SUPPL.S1.

Battista, R. N. (1996). *Towards a paradigm for technology assessment. In: Peckham M, Smith R, eds. The scientific basis of health services.* London: BMJ Publishing Group.

BeNeLuxA (2021). *HTA*. URL: https://beneluxa.org/hta.

Bowrin, Kevin, Jean Baptiste Briere, Pierre Levy, Aurélie Millier, Emilie Clay, and Mondher Toumi (2019). Cost-effectiveness analyses using real-world data: an overview of the literature. *Journal of Medical Economics* **22**:6.

Bruijn, Hans de and Ernst ten Heuvelhof (2018). *Management in Networks*.

Bullement, Ash, Tanja Podkonjak, Mark J. Robinson, Eugene Benson, Ross Selby, Anthony J. Hatswell, and Gemma E. Shields (2020). Real-world evidence use in assessments of cancer drugs by NICE. *International Journal of Technology Assessment in Health Care* **36**:4.

Coyle, D., I. Durand-Zaleski, J. Farrington, L. Garrison, J.-M. Graf von der Schulenburg, W. Greiner, L. Longworth, A. Meunier, A.-S. Moutié, S. Palmer, K. Zhao, and K. Shah (2020). HTA methodology and value frameworks for evaluation and policy making for cell and gene therapies. *European Journal of Health Economics* **21**:9, 1421–1437.

Easterby-Smith, Mark, Marjorie A. Lyles, and Eric W.K. Tsang (2008). Inter-organizational knowledge transfer: Current themes and future prospects. *Journal of Management Studies* **45**:4.

Eder, Claudia and Claudia Wild (2019). Technology forecast: advanced therapies in late clinical research, EMA approval or clinical application via hospital exemption. *Journal of Market Access & Health Policy* **7**:1.

Edquist, Charles (2011). Design of innovation policy through diagnostic analysis: Identification of systemic problems (or failures). *Industrial and Corporate Change* **20**:6.

Enserink, Bert, Leon Hermans, Jan Kwakkel, Wil Thissen, Joop Koppenjan, and Pieter Bots (2010). *Policy Analysis of Multi-Actor Systems*. The Hague: Eleven International Publishing, 32–43.

Erdös, Judit, Sabine Ettinger, Julia Mayer-Ferbas, Cecilia de Villiers, and Claudia Wild (2019). European Collaboration in Health Technology Assessment (HTA): goals, methods and outcomes with specific focus on medical devices. *Wiener Medizinische Wochenschrift* **169**:11-12.

Ernst & Young LLP (2021). *Q1 2021 Biopharma earnings analysis and industry outlook*.

EUnetHTA (2015). *The HTA Core Model® Guiding principles on use*.

EunetHTA (2021). *Creating, facilitating and promoting sustainable Health Technology Assessment (HTA) cooperation in Europe*. URL: https://www.eunethta.eu/.

European Commission (2021). *Health Technology Assessment: Commission welcomes the adoption of new rules to improve access to innovative technologies*. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip_21_6771.

European Medicines Agency (Sept. 2021). *Advanced therapy medicinal products: Overview*. URL: https://www.ema.europa.eu/en/human-regulatory/overview/advanced-therapy-medicinal-products-overview.

Facey, Karen M., Piia Rannanheimo, Laura Batchelor, Marine Borchardt, and Jo De Cock (2020). Real-world evidence to support Payer/HTA decisions about highly innovative technologies in the EU - Actions for stakeholders. *International Journal of Technology Assessment in Health Care* **36**:4.

Farid, Samar F. (2019). *Conceptual framework of the impact of health technology on healthcare system*.

Faulkner, E., D.S. Spinner, M. Ringo, and M. Carroll (2019). Are Global Health Systems Ready for Transformative Therapies? *Value in Health* **22**:6, 627–641.

FDA (2019a). *Framework for FDA's Real-World Evidence Program*.

– (2019b). *Statement from FDA Commissioner Scott Gottlieb, M.D. and Peter Marks, M.D., Ph.D., Director of the Center for Biologics Evaluation and Research on new policies to advance development of safe and effective cell and gene therapies*.

Flick, Uwe, Ernst von Kardorff, and Ines Steinke (2004). *A Companion to qualitative research*, 270–275.

Formoso, Giuolio, Ana Jeroncic, Laura Bonvicini, Olivera Djuric, Judit Erdos, Annamaria Pezzarossi, and Luciana Ballini (2022). Synthesizing quantitative and qualitative information on multiple comparisons of health interventions to facilitate knowledge transfer: an example from an EUnetHTA multi-HTA. *International Journal of Technology Assessment in Health Care* **38**:1.

Garattini, Livio and Anna Padula (2020). HTA for pharmaceuticals in Europe: will the mountain deliver a mouse? *European Journal of Health Economics* **21**:1.

Gardner, John and Andrew Webster (2016). The social management of biomedical novelty: Facilitating translation in regenerative medicine. *Social Science and Medicine* **156**:1.

Geimeinsamer Bundesausschuss (2019). *Nutzenbewertungsverfahren zum Wirkstoff Axicabtagen-Ciloleucel (Primäres mediastinales großzelliges B-Zell-Lymphom)*.

Gemeinsamer Bundesausschuss (2016). *Nutzenbewertungsverfahren zum Wirkstoff Talimogen laherparepvec (Melanom, Stadium IIIB, IIIC, IVMI1a)*.

– (2021). *Nutzenbewertungsverfahren zum Wirkstoff Onasemnogen-Abeparvovec (Überschreitung 50 Mio € Grenze: Spinale Muskelatrophie)*.

GetReal Institute (2021). *GetReal Institute*. URL: https://www.getreal-institute.org/.

Godman, Brian, Andrew Hill, Steven Simoens, Gisbert Selke, Iva Selke Krulichová, Carolina Zampirolli Dias, Antony P. Martin, Wija Oortwijn, Angela Timoney, Lars L. Gustafsson, Luka Voncina, Hye Young Kwon, Jolanta Gulbinovic, Dzintars Gotham, Janet Wale, Wânia Cristina Da Silva, Tomasz Bochenek, Eleonora Allocati, Amanj Kurdi, Olayinka O. Ogunleye, Johanna C. Meyer, Iris Hoxha, Admir Malaj, Christian Hierländer, Robert Sauermann, Wouter Hamelinck, Guenka Petrova, Ott Laius, Irene Langner, John Yfantopoulos, Roberta Joppi, Arianit Jakupi, Ieva Greiciute-Kuprijanov, Patricia Vella Bonanno, Jf Piepenbrink, Vincent de Valk, Magdalene Wladysiuk, Vanda Marković-Peković, Ileana Mardare, Jurij Fürst, Dominik Tomek, Mercè Obach Cortadellas, Corinne Zara, Caridad Pontes, Stuart McTaggart, Tracey Lea Laba, Øyvind Melien, Durhane Wong-Rieger, Seung Jin Bae, and Ruaraidh Hill (2021). Potential approaches for the pricing of cancer medicines across Europe to enhance the sustainability of healthcare systems and the implications. *Expert Review of Pharmacoeconomics and Outcomes Research* **21**:4.

Gozzo, Lucia, Giovanni Luca Romano, Francesca Romano, Serena Brancati, Laura Longo, Daniela Cristina Vitale, and Filippo Drago (Oct. 2021). Health Technology Assessment of Advanced Therapy Medicinal Products: Comparison Among 3 European Countries. *Frontiers in Pharmacology* **12**:1.

Grant, Robert M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal* **17**:SUPPL. WINTER, 109–122.

Gray, Barbara (1985). Conditions Facilitating Interorganizational Collaboration. *Human Relations* **38**:10.

Green, Eric D., James D. Watson, and Francis S. Collins (2015). *Human Genome Project: Twenty-five years of big biology*.

Hampson, G., A. Towse, S.D. Pearson, W.B. Dreitlein, and C. Henshall (2018). Gene therapy: Evidence, value and affordability in the US health care system. *Journal of Comparative Effectiveness Research* **7**:1, 15–28.

Hettle, Robert, Mark Corbett, Sebastian Hinde, Robert Hodgson, Julie Jones-Diette, Nerys Woolacott, and Stephen Palmer (2017). The assessment and appraisal of regenerative medicines and cell therapy products: An exploration of methods for review, economic evaluation and appraisal. *Health Technology Assessment* **21**:7, 131–138.

HTA Glossary (Sept. 2021). *health technology*. URL: http://htaglossary.net/health+technology.

IMI GetReal (2016). *WP1: Deliverable D1.3 Glossary of definitions of common terms*.

IQVIA (2021). *HTA Accelerator to support strategic payer decisions and market access*. URL: https://www.iqvia.com/solutions/real-world-evidence/health-economics-and-value/hta-accelerator.

Jaksa, Ashley, James Wu, Páll Jónsson, Hans Georg Eichler, Sarah Vititoe, and Nicolle M. Gatto (2021). *Organized structure of real-world evidence best practices: Moving from fragmented recommendations to comprehensive guidance*.

Jean, Nicole St, Carla Pinto, Ines Tenente, and Grace Murray (2018). *Collaboration is key to accelerating diagnostics access to optimize benefits of precision medicines*.

Jørgensen, J. and P. Kefalas (2021). The use of innovative payment mechanisms for gene therapies in Europe and the USA. *Regenerative Medicine* **16**:4, 405–421.

Jørgensen, Jesper, Eve Hanna, and Panos Kefalas (2020). Outcomes-based reimbursement for gene therapies in practice: the experience of recently launched CAR-T cell therapies in major European countries. *Journal of Market Access & Health Policy* **8**:1.

Jørgensen, Jesper, Laura Mungapen, and Panos Kefalas (2019). Data collection infrastructure for patient outcomes in the UK – opportunities and challenges for cell and gene therapies launching. *Journal of Market Access & Health Policy* **7**:1.

Kálo, Zoltán, Adrian Gheorghe, Mirjana Huic, Marcell Csanádi, and Finn Boerlum Kristensen (2016). HTA implementation roadmap in Central and Eastern European countries. *Health Economics* **25**:1, 179–192.

Kanavos, Panos, Aris Angelis, and Michael Drummond (2019). *An EU-wide approach to HTA: An irrelevant development or an opportunity not to be missed?*

Kassarjian, Harold H. (1977). Content Analysis in Consumer Research. *Journal of Consumer Research* **4**:1.

Katkade, Vaibhav B., Kafi N. Sanders, and Kelly H. Zou (2018). Real world data: An opportunity to supplement existing evidence for the use of long-established medicines in health care decision making. *Journal of Multidisciplinary Healthcare* **11**:1.

Kefalas, Panos, Omar Ali, Jesper Jørgensen, Nick Merryfield, Tim Richardson, Adam Meads, Laura Mungapen, and Matthew Durdy (2018). Establishing the cost of implementing a performance-based, managed entry agreement for a hypothetical CAR T-cell therapy. *Journal of Market Access & Health Policy* **6**:1.

Kent, Seamus, Maximilian Salcher-Konrad, Stefania Boccia, Jacoline C. Bouvy, Chiara de Waure, Jaime Espin, Karen Facey, Mary Nguyen, Juan Carlos Rejon-Parrilla, and Pall Jonsson (2021). The use of nonrandomized evidence to estimate treatment effects in health technology assessment. *Journal of Comparative Effectiveness Research* **10**:14.

Kisser, Agnes, Joschua Knieriemen, Annette Fasan, Karolin Eberle, Sara Hogger, Sebastian Werner, Tina Taube, and Andrej Rasch (2021). Towards compatibility of EUnetHTA JCA methodology and German

HTA: a systematic comparison and recommendations from an industry perspective. *European Journal of Health Economics*.

Kleijnen, Sarah, Elisabeth George, Scott Goulden, Anne D'Andon, Pauline Vitré, Boguslawa Osiska, Rafal Rdzany, Steffen Thirstrup, Belen Corbacho, Bence Z. Nagy, Hubert G. Leufkens, Anthonius De Boer, and Wim G. Goettsch (2012). Relative effectiveness assessment of pharmaceuticals: Similarities and differences in 29 jurisdictions. *Value in Health* **15**:6.

Lane, Peter J. and Michael Lubatkin (1998). Relative absorptive capacity and interorganizational learning. *Strategic Management Journal* **19**:5.

Leavy, Patricia (2014). *The oxford handbook of qualitative research: Second edition*.

Lee, Woojung, Victoria Dayer, Boshen Jiao, Josh J. Carlson, Beth Devine, and David L. Veenstra (2021). Use of real-world evidence in economic assessments of pharmaceuticals in the United States. *Journal of Managed Care and Specialty Pharmacy* **27**:1.

Leyens, Lada and Angela Brand (2016). Early Patient Access to Medicines: Health Technology Assessment Bodies Need to Catch Up with New Marketing Authorization Methods. *Public Health Genomics* **19**:3.

Logan, Jo and Ian D. Graham (1998). Toward a comprehensive interdisciplinary model of health care research use. *Science Communication* **20**:2.

Makady, A, A van Veelen, P Jonsson, O Moseley, A D'Andon, A de Boer, JL Hillege, O Klungel, and W Goettsch (2018). Using Real-World Data (RWD) in Health Technology Assessment (HTA) Practice: A Comparative Study of 5 HTA Agencies. *Value in Health* **20**:9.

Makady, Amr, Anthonius de Boer, Hans Hillege, Olaf Klungel, and Wim Goettsch (2017a). What Is Real-World Data? A Review of Definitions Based on Literature and Stakeholder Interviews. *Value in Health* **20**:7.

Makady, Amr, Renske ten Ham, Anthonius de Boer, Hans Hillege, Olaf Klungel, and Wim Goettsch (2017b). Policies for Use of Real-World Data in Health Technology Assessment (HTA): A Comparative Study of Six HTA Agencies. *Value in Health* **20**:4.

Malone, Daniel C., Mary Brown, Jason T. Hurwitz, Loretta Peters, and Jennifer S. Graff (2018). Real-World Evidence: Useful in the Real World of US Payer Decision Making? How? When? And What Studies? *Value in Health* **21**:3.

Miles, M B and A M Huberman (1994). *Qualitative Data Analysis: An expanded Sourcebook (2nd Ed.)*

Milliano, Thom de (2019). *The impact of Real-World Evidence on demonstrating clinical value in HTAs*.

Moher, David, Alessandro Liberati, Jennifer Tetzlaff, Douglas G. Altman, Doug Altman, Gerd Antes, David Atkins, Virginia Barbour, Nick Barrowman, Jesse A. Berlin, Jocalyn Clark, Mike Clarke, Deborah Cook, Roberto D'Amico, Jonathan J. Deeks, P. J. Devereaux, Kay Dickersin, Matthias Egger, Edzard Ernst, Peter C. Gøtzsche, Jeremy Grimshaw, Gordon Guyatt, Julian Higgins, John P.A. Ioannidis, Jos Kleijnen, Tom Lang, Nicola Magrini, David McNamee, Lorenzo Moja, Cynthia Mulrow, Maryann Napoli, Andy Oxman, Bá Pham, Drummond Rennie, Margaret Sampson, Kenneth F. Schulz, Paul G. Shekelle, David Tovey, and Peter Tugwell (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine* **6**:7.

Murad, M. Hassan, Noor Asi, Mouaz Alsawas, and Fares Alahdab (2016). *New evidence pyramid*.

National Health Council (2021). *Real-World Evidence Glossary*. URL: https://nationalhealthcouncil.org/additional-resources/real-world-evidence-glossary/.

Newell, Sue, Josh Morton, Marco Marabelli, and Robert Galliers (2020). *Managing Digital Innovation*. Red Globe Press.

NICE (2013). Guide to the methods of technology appraisal 2013. *National Institute for Health and Care Excellence* April.

– (2016). *Talimogene laherparepvec for treating unresectable metastatic melanoma. Technology appraisal guidance [TA410]*.

– (2019). *Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after 2 or more systemic therapies. Technology appraisal guidance [TA559]*.
– (2021a). *Changes to NICE drug appraisals: what you need to know*. URL: https://www.nice.org.uk/news/feature/changes-to-nice-drug-appraisals-what-you-need-to-know.
– (2021b). *Glossary*. URL: https://www.nice.org.uk/glossary?letter=q.
– (2021c). *NICE strategy 2021 to 2026*.
– (2021d). *Onasemnogene abeparvovec for treating spinal muscular atrophy. Highly specialised technologies guidance*.
– (2021e). *Real World Evidence Framework*.

O'Rourke, Brian, Wija Oortwijn, and Tara Schuller (2020). The new definition of health technology assessment: A milestone in international collaboration. *International Journal of Technology Assessment in Health Care* **36**:3.

Overbeeke, E. van, S. Michelsen, M. Toumi, H. Stevens, M. Trusheim, I. Huys, and S. Simoens (2021). Market access of gene therapies across Europe, USA, and Canada: challenges, trends, and solutions. *Drug Discovery Today* **26**:2, 399–415.

Pichler, Franz, Wija Oortwijn, Alric Ruether, and Rebecca Trowman (2019). Defining capacity building in the context of HTA: A proposal by the HTAi Scientific Development and Capacity Building Committee. *International Journal of Technology Assessment in Health Care* **35**:5.

Pochopień, M., T. Qiu, S. Aballea, E. Clay, and M. Toumi (2021). Considering potential solutions for limitations and challenges in the health economic evaluation of gene therapies. *Expert Review of Pharmacoeconomics and Outcomes Research*.

Richardson, W. S., M. C. Wilson, J. Nishikawa, and R. S. Hayward (1995). The well-built clinical question: a key to evidence-based decisions. *ACP journal club* **123**:3.

Rittel, Horst W.J. and Melvin M. Webber (1973). Dilemmas in a general theory of planning. *Policy Sciences* **4**:2.

Roberts, Melissa H. and Gary T. Ferguson (2021). *Real-World Evidence: Bridging Gaps in Evidence to Guide Payer Decisions*.

Ronco, V., M. Dilecce, E. Lanati, P.L. Canonico, and C. Jommi (2021). Price and reimbursement of advanced therapeutic medicinal products in Europe: are assessment and appraisal diverging from expert recommendations? *Journal of Pharmaceutical Policy and Practice* **14**:1.

RWE Navigator (2021a). *Generating real-world evidence*. URL: https://rwe-navigator.eu/use-real-world-evidence/generate-real-world-evidence/.
– (2021b). *Sources of real-world data*. URL: https://rwe-navigator.eu/use-real-world-evidence/sources-of-real-world-data/.

Schubert, Tino and Tobias Vogelmann (2019). *Market Access in der Medizintechnik*.

Schulz, Sandra, Anna Marie Passon, Matthias Perleth, Michael Kulig, Nina Paschke, and Katja Matthias (2020). The Evaluation of Orphan Drugs by the German Joint Federal Committee-An Eight-Year Review. *Deutsches Arzteblatt international* **117**:50.

Sekaran, Uma and Roger Bougie (2016). *Research Methods for Business: A Skill Building Approach*.

Sherman, Rachel E., Steven A. Anderson, Gerald J. Dal Pan, Gerry W. Gray, Thomas Gross, Nina L. Hunter, Lisa LaVange, Danica Marinac-Dabic, Peter W. Marks, Melissa A. Robb, Jeffrey Shuren, Robert Temple, Janet Woodcock, Lilly Q. Yue, and Robert M. Califf (2016). Real-World Evidence — What Is It and What Can It Tell Us? *New England Journal of Medicine* **375**:23.

Sievers, Hannah, Angelika Joos, and Mickaël Hiligsmann (2021). Real-world evidence: Perspectives on challenges, value, and alignment of regulatory and national health technology assessment data collection requirements. *International Journal of Technology Assessment in Health Care*.

Simpson, Alex and Sreeram V. Ramagopalan (2021). R WE ready for reimbursement? A round up of developments in real-world evidence relating to health technology assessment. *Journal of Comparative Effectiveness Research* **10**:10.

Stack Overflow (2017). *I want to extract a certain number of words surrounding a given word in a long string(paragraph) in Python 2.7*. URL: https://stackoverflow.com/questions/43449773/i-want-to-extract-a-certain-number-of-words-surrounding-a-given-word-in-a-long-s.

Ten Ham, Renske M T, Geert W J Frederix, Olivia Wu, Wim Goettsch, Hubert G M Leufkens, Olaf H Klungel, and Jarno Hoekman (2021). Comparative-Effectiveness Research Key Considerations in the Health Technology Assessment of Advanced Therapy Medicinal Products in Scotland. *Value in Health*.

Trosman, Julia R., Stephanie L. Van Bebber, and Kathryn A. Phillips (2011). Health technology assessment and private payers' coverage of personalized medicine. *Journal of Oncology Practice* **7**:3 SUPPL.

TU Delft (2019). *Final Assessment Form MSc Thesis Project CoSEM | EPA | MOT*.

Van Nooten, Floortje, Stefan Holmstrom, Julia Green, Ingela Wiklund, Isaac A.O. Odeyemi, and Teresa K. Wilcox (2012). Health economics and outcomes research within drug development: Challenges and opportunities for reimbursement and market access within biopharma research. *Drug Discovery Today* **17**:11-12.

Velasco-Garrido, Marcial and Reinhard Buss (2005). *Policy brief Health technology assessment An introduction to objectives, role of evidence, and structure in Europe*.

Vreman, Rick A., Jacoline C. Bouvy, Lourens T. Bloem, Anke M. Hövels, Aukje K. Mantel-Teeuwisse, Hubert G.M. Leufkens, and Wim G. Goettsch (2019). Weighing of Evidence by Health Technology Assessment Bodies: Retrospective Study of Reimbursement Recommendations for Conditionally Approved Drugs. *Clinical Pharmacology and Therapeutics* **105**:3.

Vreman, Rick A., Aukje K. Mantel-Teeuwisse, Anke M. Hövels, Hubert G.M. Leufkens, and Wim G. Goettsch (2020). Differences in Health Technology Assessment Recommendations Among European Jurisdictions: The Role of Practice Variations. *Value in Health* **23**:1.

Wang, Ting, Neil McAuslane, Lawrence Liberti, Hubert Leufkens, and Anke Hövels (2018). Building Synergy between Regulatory and HTA Agencies beyond Processes and Procedures—Can We Effectively Align the Evidentiary Requirements? A Survey of Stakeholder Perceptions. *Value in Health* **21**:6.

Ylä-Herttuala, Seppo (2015). Glybera's second act: The curtain rises on the high cost of therapy. *Molecular Therapy* **23**:2.

Zaprutko, Tomasz, Dorota Kopciuch, Krzysztof Kus, Piotr Merks, Monika Nowicka, Izabela Augustyniak, and El Bieta Nowakowska (2017). Affordability of medicines in the European Union. *PLoS ONE* **12**:2.

A1

Literature review process

Status quo of curative therapy HTAs

Table A1.1. Overview of search queries for literature review.

Search Engine	Domain	Query	Hits	Used
Track 1: status quo of curative therapy HTAs				
Scopus	TITLE-ABS-KEY	((”health technology assessment*” OR HTA) AND (”Gene therap*” OR GTMP))	38	No
Scopus	TITLE-ABS-KEY	((”health technology assessment*” OR HTA OR ”Market Access”) AND (”Cell and gene therap*” OR CGT OR ”Cell therap*” OR ”Gene therap*” OR ATMP OR ”Advanced Therapy Medicinal Product*”))	86	Yes
PubMed®	All	((”health technology assessment*” OR HTA OR ”Market Access”) AND (”Cell and gene therap*” OR CGT OR ”Cell therap*” OR ”Gene therap*” OR ATMP OR ”Advanced Therapy Medicinal Product*”))	108	Yes
Track 2: explicating real-world evidence usage in HTAs				
Scopus	TITLE-ABS-KEY	((”health technology assessment*” OR HTA OR ”Market Access”) AND (”Cell and gene therap*” OR CGT OR ”Cell therap*” OR ”Gene therap*” OR ATMP OR ”Advanced Therapy Medicinal Product*”) AND (”real world evidence” OR ”RWE” OR ”real world data” OR ”rwd”))	6	No
Scopus	TITLE-ABS-KEY	((”health technology assessment*” OR HTA OR ”Market access”) AND (”real world evidence” OR ”RWE” OR ”real world data” OR ”rwd”))	144	No
Scopus	TITLE-ABS	((”health technology assessment*” OR HTA OR ”Market access”) AND (”real world evidence” OR ”RWE” OR ”real world data” OR ”rwd”))	110	Yes

continues on next page

PubMed®	All	((”health technology assessment*” OR HTA OR ”Market access”) AND (”real world evidence” OR ”RWE” OR ”real world data” OR ”rwd”))	295	No
PubMed®	Title-Abs	((”health technology assessment*” [Title/Abstract] OR HTA[Title/Abstract]) AND (”real world evidence” [Title/Abstract] OR ”RWE” [Title/Abstract] OR ”real world data” [Title/Abstract] OR ”rwd”[Title/Abstract]))	122	Yes
Scopus	TITLE-ABS-KEY	((”health technology assessment*” OR hta OR ”Market Access”) AND (”Cell and gene therap*” OR cgt OR ”Cell therap*” OR ”Gene therap*” OR atmp OR ”Advanced Therapy Medicinal Product*”) AND(”real world evidence” OR ”RWE” OR ”real world data” OR ”rwd” OR ”cluster RCT” OR cohort* OR ”common comparator” OR cross-sectional OR database* OR ”extenstion stud*” OR ”hospital data” OR indirect OR ”insurance claim” OR kaplan OR ”meta-analys*” OR ”non-randomised” OR ”observational stud*” OR ”patient-power*” OR ”patient-report*” OR pharmacoviligance OR ”pragmatic RCT” OR ”prescription data” OR ”propensity score” OR ”prospective stud*” OR proxy OR ”electronic health record*” OR registr* OR ”retrospective stud*” OR ”social media” OR ”supplement* to RCT” OR ”Health survey” OR ”uncontrolled stud*” OR ”vignette stud*”))	25	No

A2

Interview protocol

Introduction

All interviewees signed an informed consent form prior to the interview.

Interviewer

I'd like to thank you for willing to participate in this interview as part of my master thesis. First I will introduce myself, the research and the objective of today's interview.

1. **Introduction researcher & study**: [...] Currently, I am writing my thesis on RWD/RWE usage in gene therapy HTA reports, specifically looking at Germany (G-BA) and England (NICE).
   The purpose of this research study is to deliver empirical evidence on how health technology assessment bodies embrace real-world data in the appraisal of gene therapy medicinal products. Moreover, this study aims to identify potential challenges in aligning these agencies in a joint clinical assessment.
2. **Goal of the interview**: Interviews serve to validate initial findings and explore additional considerations that may not have been captured in the performed analyses.
3. **Confidentiality**: Your name will be kept confidential. Findings from this discussion will be collated with other respondents and presented in the final research in aggregated or anonymous form.
   During the course of this interview, you will not be requested to share information that you are not allowed to share. Please let me know if a question requires you to reveal confidential information.
4. **Time duration**: The interview will take you approximately 45 - 60 minutes to complete.
5. **Other**: Your participation in this study is entirely voluntary and you can withdraw at any time. You are free to omit any question. If any questions (or other questions) arise at any point during the interview, please feel free to ask them.

If you consent, this interview will be recorded and transcribed for note taking purposes. Is that okay with you?
If yes: start recording
If no: proceed without recording

RWD is henceforth defined as an umbrella term for data collected outside the setting of randomised controlled trials (IMI GetReal, 2016). RWE is hereafter defined as the evidence derived from the analysis and/or synthesis of RWD (IMI GetReal, 2016).

Interviewee

1. Could you briefly describe your experience with gene therapy HTAs, RWD/RWE usage and/or alignment of HTA bodies in joint clinical assessments?

Gene therapy HTAs

Objective: The goal of this section is to understand the current practices in gene therapy HTAs, including challenges encountered and specific considerations applied in their appraisal.

1. What are key considerations in the rationales of gene therapy HTA appraisals?
   Probe: orphan designation, burden of illness, end of life
2. Based on literature, it appears that long-term uncertainty, lack of long-term data on efficacy & safety and uncertainty in cost-effectiveness are key methodological challenges associated with gene therapy HTAs

- How do different HTA bodies (specifically looking at NICE and G-BA) mitigate these challenges?
  Probe: discounts, outcome-based agreements, temporary decisions, continued evidence development agreements

The role of real-world evidence in gene therapy HTAs

Objective: The goal of this section is to understand the role that RWD/RWE plays and could play in gene therapy HTAs

1. To what extent would you say that RWD/RWE could mitigate the earlier identified methodological challenges encountered in appraising gene therapies?
2. How would you say that the role of RWD/RWE in this context has evolved and will evolve over time?
3. Specifically looking at G-BA and NICE, what would you say are key differences in their RWD/RWE usage?
   Probe: areas supported, acceptance rates
   (a) How would you say that these differences impact their appraisal of RWD/RWE in gene therapy HTAs?
   Probe: areas supported, acceptance rates

Factors impacting real-world evidence usage

1. What are key barriers impeding RWD/RWE uptake, implementation and utilisation in gene therapy HTAs?
   Probe: lack of alignment on definition, lack of guidance, methodological challenges (i.e. confounding bias, lacking randomisation, lacking transparency)
2. What would you say are key enablers for increased RWD/RWE usage in gene therapy HTAs?
   Probe: Data collection standardisation, alignment of stakeholders on definition and added value
3. How would you say that these barriers and enablers differ for NICE and G-BA?

Exploratory: alignment of HTAs in the context of joint clinical assessments

1. What factors and considerations would be relevant in aligning gene therapy HTA outcomes by G-BA and NICE?
2. The European committee recently introduced legislation that enables EU-wide HTA assessments.
   - What opportunities and challenges do you foresee in these joint clinical assessments?
   - What are key implications of such legislation in the context of gene therapy HTAs?
   - What are key implications of such legislation in the context of RWD/RWE usage?
3. What other opportunities do you see for national and international collaborations in (gene therapy) HTAs in the next three years?
   (a) What would some strategies to facilitate these collaborations?

Other

1. Are there other things that we have not covered, but are relevant to consider in the discussed context?

Thank you for your participation.
Would you like to receive a link to my thesis, once it has been submitted?

If applicable: end recording

A3

Search algorithm

Algorithm modified from open-source code (Stack Overflow, 2017).

```
#import required packages
import os
import pandas as pd
import re

#load HTA report text file
with (open(os.path.expanduser(”~/Desktop/HTAs/GBA_M4_Imlygic.txt”),
encoding=”utf8”, errors='ignore')) as f:
    text = f.read()

#define and load excel sheet with pre-defined HTA keywords to search for
df_keywords = pd.read_excel('HTA_keywords.xlsx', sheet_name=3)
join_keywords = df_keywords.values.tolist()
key_words = [''.join(ele) for ele in join_keywords]

#define search
def search(target, text, context=100):
    words = re.findall(r'\w+', text)

    matches = (i for (i,w) in enumerate(words) if target in w.lower())
    output = []
    for index in matches:
        if index < context //2:
            output.append(words[0:context+1])
        elif index > len(words) - context//2 - 1:
            output.append(words[-(context+1):])
        else:
            output.append(words[index - context//2:index + context//2 + 1])
    return output

df = pd.DataFrame(columns=[”keywords”, ”words”])
for keywords in key_words:
    words = search(keywords, text, context=100)
    for w in words:
        w = ” ”.join(w)
        df = df.append({”keywords”: keywords, ”words”: w}, ignore_index=True)

#export text extracts that contain key words to Excel
print(df)
df.to_excel(”HTA_output.xlsx”)
```

A4

RWD/RWE terminology

Table A4.1. RWD/RWE terminology- English. List used for NICE reports, RWD: Real-world data. RWE: Real-world evidence.

RWD/ RWE	Type	Source	Keyword for algorithm
RWD	Administrative databases	Makady *et al*. (2017a)	Database
RWD	Claims database	Makady *et al*. (2017a)	Claims | Database
RWD	Clinicial database	Makady *et al*. (2017a)	Database
RWD	Health surveys	Makady *et al*. (2017a)	Survey
RWD	Healthcare databases including health records	RWE Navigator (2021b)	Database | Record
RWD	Hospital data	Makady *et al*. (2017a)	Hospital
RWD	Insurance claims	IQVIA internal expertise (2021)	Insurance | Claim
RWD	Patient / Physician interviews	IQVIA internal expertise (2021)	Interview
RWD	Patient reported outcome	Makady *et al*. (2017a)	Patient-reported
RWD	Patient registries	RWE Navigator (2021b)	Registr
RWD	Pharmacoviligance data	IQVIA internal expertise (2021)	Pharmacoviligance
RWD	Pharmacy and health insurance databases	RWE Navigator (2021b)	Database | Insurance
RWD	Prescriptipn data	IQVIA internal expertise (2021)	Prescription
RWD	Post-marketing studies	Makady *et al*. (2017a)	Post-marketing
RWD	Social media	RWE Navigator (2021b)	Social
RWD	Real-world data	IMI GetReal (2016)	Real-world
RWE	Case-control	RWE Navigator (2021a)	Case
RWE	Case report	RWE Navigator (2021a)	Case
RWE	Case series	RWE Navigator (2021a)	Case
RWE	Cohort study	RWE Navigator (2021a)	Cohort
RWE	Cohort multiple RCT (cmRCT)	RWE Navigator (2021a)	Cohort
RWE	Comprehensive cohort study	RWE Navigator (2021a)	Cohort
RWE	Cluster RCT	RWE Navigator (2021a)	Cluster

continues on next page

RWE	Cross-sectional	RWE Navigator (2021a)	Cross-sectional
RWE	Extension study	IQVIA internal expertise (2021)	Extension
RWE	Experimental vignette studies	IQVIA internal expertise (2021)	Vignette
RWE	Indirect treatment comparison	IQVIA internal expertise (2021)	Indirect
RWE	National history study	IQVIA internal expertise (2021)	Natural
RWE	Non-interventional study	IQVIA internal expertise (2021)	Non-interventional
RWE	Non-randomised controlled trial	RWE Navigator (2021a)	Non- randomised
RWE	Meta-analysis	IQVIA internal expertise (2021)	Meta-analys
RWE	Observational study	IQVIA internal expertise (2021)	Observational
RWE	Observational; prospective cohort study	FDA (2019a)	Prospective
RWE	Observational; retrospective cohort study	FDA (2019a)	Retrospective
RWE	Pragmatic RCT	RWE Navigator (2021a)	Pragmatic
RWE	Prospective outcomes study	IQVIA internal expertise (2021)	Prospective
RWE	Real-world evidence	IMI GetReal (2016)	Real-world
RWE	Retrospective chart review	IQVIA internal expertise (2021)	Retrospective | Chart review
RWE	Uncontrolled studies	IQVIA internal expertise (2021)	Uncontrolled

Table A4.2. RWD/RWE terminology- German. List used for G-BA reports, RWD: Real-world data. RWE: Real-world evidence.

RWD/ RWE	Type	Source	Keyword for algorithm
RWD	Alltagsbedingungen	Schubert and Vogelmann (2019)	Alltags
RWD	Assoziationsbeobachtungen	Schubert and Vogelmann (2019)	Beobachtung
RWD	Datenbank	IQVIA internal expertise (2021)	Datenbank
RWD	Einzelfallberichte	Schubert and Vogelmann (2019)	Einzelfall

continues on next page

RWD	Elektronische Patientenakten (EHR)	Schubert and Vogelmann (2019)	Patientenakten
RWD	Extensionsphase	IQVIA internal expertise (2021)	Extension
RWD	Extensionsprotkolls	IQVIA internal expertise (2021)	Extension
RWD	Fallzahlen	IQVIA internal expertise (2021)	Fallzahl
RWD	Krankenkassendaten	Schubert and Vogelmann (2019)	Krankenkassendaten
RWD	Patientenregister	Schubert and Vogelmann (2019)	Patientenregister
RWD	Primäre prospektive Datenerhebung	IQVIA internal expertise (2021)	Prospektiv
RWD	Real-world Daten	Schubert and Vogelmann (2019)	Real
RWD	Realen Versorgungssituation	IQVIA internal expertise (2021)	Real
RWD	Registerdaten	Schubert and Vogelmann (2019)	Register
RWD	Routinedaten	Schubert and Vogelmann (2019)	Routinedaten
RWD	Verlaufsbeobachtungen		Verlaufsbeobachtungen
RWE	Beobachtungsstudien	Schubert and Vogelmann (2019)	Beobachtung
RWE	Clusterrandomisierte Studien	IQVIA internal expertise (2021)	Clusterrandomisiert
RWE	Fall-Kontrollstudien	Schubert and Vogelmann (2019)	Fall-Kontrollstudien
RWE	Fallserie	Schubert and Vogelmann (2019)	Fallserie
RWE	Historisch kontrollierte Studien	IQVIA internal expertise (2021)	Historisch kontrollierte
RWE	Kohortenstudien	Schubert and Vogelmann (2019)	Kohorten
RWE	Nicht-vergleichenden studien	IQVIA internal expertise (2021)	Nicht-vergleichenden
RWE	Nicht-randomisierte studie	IQVIA internal expertise (2021)	Nicht-randomisierte
RWE	Pragmatische randomisierte Studie	Schubert and Vogelmann (2019)	Pragmatisch
RWE	Prospektive Studien	Schubert and Vogelmann (2019)	Prospektiv

continues on next page

RWE	Qualitativen studien	IQVIA internal expertise (2021)	Qualitative
RWE	Quasirandomisierten studien	Schubert and Vogelmann (2019)	Quasirandomisiert
RWE	Querschnittsstudie	Schubert and Vogelmann (2019)	Querschnitt
RWE	Retrospektive studien	Schubert and Vogelmann (2019)	Retrospektiv
RWE	Studien unter alltagsbedingungen	IQVIA internal expertise (2021)	alltag
RWE	Unkontrollierten Verlängerungsstudie	IQVIA internal expertise (2021)	Unkontrolliert
RWE	Vorher-Nachher Design	Schubert and Vogelmann (2019)	Vorher-Nachher

A5

Data extraction form

Modified from Makady *et al*. (2018) and Lee *et al*. (2021), categorisation of areas supported by RWE according to IQVIA internal expertise.

Part 1 - General information of the HTA report

- HTA body & title of HTA report
- Date of publication
- Indication

Part 2 - Characteristics of RWD/RWE

- Is real-world data / real-world evidence included in the HTA report
  - No
  - Yes

  If yes, continue to the next items
- Title of RWD/RWE
- Types of RWD/RWE, choose from table below:

Table A5.1. RWD/RWE categories in data extraction form. Not exhaustive. Derived from IQVIA internal research (2021. RWD: Real-world data. RWE: Real-world evidence.)

RWD / RWE	Category	Type
RWD	Patient registry data	Disease / condition registries
		Product registry
	Healthcare data	electronic patient/health/ medical record
	Adminstrative data	Prescriptions
		Hospital data
		Health insurance claims/ records
	Social media data	patient-powered research networks (PPRNs)
	Electronic source data	mobile device-generated data
		mobile health (mHealth)
		passive sensor devices
		mobile apps
		patient-generated data
		wearables
	Post-authorisation data	pharmacovigilance
		pharmacoepidemiology
	Surveys	patient surveys / interviews
	Surveys	physician surveys / interviews
RWE	Experimental	Pragmatic trial

continues on next page

	Population enrichment trial
	Cohort multiple trial
	Comprehensive cohort study
	Cluster trial
	Non-randomised controlled trial
	Large simple trials
	Experimental vignette study
Observational	Retrospective cohort study
	Prospective cohort study
	Case-control study
	Cross-sectional study
	Case series / interrupted time-series / before-and-after study
	Chart review studies
	Observational vignette study
Post-authorisation	Post-authorisation safety study (PASS)
	Post-authorisation efficacy study (PAES)
	Periodic safety update reports (PSUR)

Part 3 - Areas supported by RWD/RWE

- What area was supported by the use of RWD/RWE, choose from table below:

Table A5.2. Areas supported categorisation in data extraction form. Follwing the PICO framework, derived from IQVIA internal research (2021).

PICO category	Area supported	Definition
Population	Burden of illness	RWD was used to describe the burden of illness (e.g., disease mortality, risk factors, impact on HRQoL, unmet need) of the population (the indication and/or subgroups) being reviewed in the HTA.
	Epidemiology	RWD was used to estimate the size of the population (the indication and/or subgroups), i.e. prevalence and incidence, being reviewed in the HTA. Typically this data is used to establish the budget impact of the new treatment, and/or to claim special considerations such as rare disease.
	Patient characteristics	RWD was used to describe the population in terms of distribution by age, gender, ethnicity, socio-economic factors, co-morbidities and other factors treated locally in usual practice to demonstrate that the clinical trial represents the real-world patient populations.
	Treatment pathway	RWD was used to describe the current treatment pathway, e.g., to provide information what % of patients receive what treatment in the current treatment pathway either to support what products are part of current standard of care or what products these patients may receive either prior to and after the indication under review (i.e. under prior and subsequent therapies in real-world practice)
Intervention	Compliance, adherence, persistence - intervention	RWD was used to provide evidence on the compliance, adherence or persistence of the product being assessed in the HTA

continues on next page

	Treatment satisfaction	RWD was used to provide evidence on the treatment satisfaction or patient preference of the product being assessed in the HTA
Comparator	Appropriate comparator	RWD was used to demonstrate that the trial comparator is part of current standard of care in the country of the HTAB and hence should be accepted as an appropriate comparator.
	Compliance, adherence, persistence - comparator	RWD was used to provide evidence on the compliance, adherence or persistence of (one of) the comparator(s) included in HTA submission
	External comparator	RWD on external comparator (also referred to as external control, historical control or synthetic control) was used to assess clinical benefit and safety (e.g., in cases where the pivotal study is a single-arm trial, or where in case where no link could be established with RCT to do an ITC)
Outcomes (clinical)	Effectiveness - comparator	RWD was used to provide evidence on the effectiveness of the comparator(s) included in the HTA
	Effectiveness - intervention	RWD was used to provide evidence on the effectiveness of the product being assessed in the HTA
	Extrapolation of OS - comparator	RWD was used to extrapolate (or validate the extrapolation) of the effectiveness in terms of overall survival (OS) of the comparator product beyond the trial duration to estimate its long-term effectiveness (e.g., data used to model the natural history of the disease)
	Extrapolation of OS - intervention	RWD was used to extrapolate (or validate the extrapolation) of the effectiveness in terms of overall survival (OS) of the new product beyond the trial duration to estimate its long-term effectiveness.
	Extrapolation of PFS - comparator	RWD was used to extrapolate (or validate the extrapolation) of the effectiveness in terms of PFS or other xFS endpoints (e.g., MFS, DFS, EFS, RFS) of the comparator product beyond the trial duration to estimate its long-term effectiveness (e.g., data used to model the natural history of the disease)
	Extrapolation of PFS - intervention	RWD was used to extrapolate (or validate the extrapolation) of the effectiveness in terms of PFS or other xFS endpoints (e.g., MFS, DFS, EFS, RFS) of the new product beyond the trial duration to estimate its long-term effectiveness
	Safety - comparator	RWD was used to provide evidence on the safety of the comparator(s) included in the HTA
	Safety - intervention	RWD was used to provide evidence on the safety of the product being assessed in the HTA
	Validation of surrogate endpoints	Trial outcomes may represent physiological parameters, such as tumour response, blood haemoglobin level or lung function, which are not considered to be patient-relevant. However, these may serve as surrogate endpoints (proxies) for effectiveness outcomes of relevance to HTA, but the relationship between the surrogate and 'final' endpoint needs to be demonstrated quantitatively. RWD was used to validate the surrogate endpoint use in the trial (e.g., PFS) to hard endpoints (e.g., OS)
Outcomes (PRO)	QoL- comparator	RWD was used to provide evidence on the impact of the comparator(s) being assessed in the HTA on patient's QOL (e.g., EQ-5D or other PRO data collected through RWD). Note RWD to describe the general impact of the disease on QoL is captured under burden of illness.

continues on next page

	QoL- intervention	RWD was used to provide evidence on the impact of the product being assessed in the HTA on patient's QOL (e.g., EQ-5D or other PRO data collected through RWD). Note RWD to describe the general impact of the disease on QoL is captured under burden of illness.
Outcomes (economic)	Costs	RWD was used to collect information on health care costs or cost savings (e.g., cost of treating complications, cost of transplantation, cost of dialysis, cost of stay in ICU etc) used in the economic model
	Dis-utilities	RWD was used to collect information on disutilities used in the economic model associated with a specific event (e.g., complications or adverse events)
	Health-state transition probabilities	RWD was used to collect information on health-state transition probabilities used in the economic model
	Resource utilisation	RWD was used to collect information on health care resource utilisation (e.g., average length of stay in hospital for the specific indication) used in the economic model
	Utility	RWD was used to collect information on health state utilities used in the economic model

Part 4 - Final appraisal

- What was the impact of the RWD/RWE for decision-making?
  - **Accepted**, statement identifying a positive opinion on the role of data derived from RWD/RWE or statement on inclusion of the RWD/RWE source
  - **Not accepted**, statement identifying a negative opinion on the role of data derived from RWD/RWE or statement on exclusion of the RWD/RWE source
  - **Not identified**, no statement identified regarding the role or inclusion/exclusion of RWD/RWE
- What was the final recommendation of the dossier for effectiveness?
  - Positive or added benefit
  - Equal benefit or added benefit not proven
  - Negative or lesser benefit

A6

Data validation

Validation of data extraction methodology

To derive the completeness of extracted data using the algorithm, the output was compared to the HTA Accelerator data (Table A6.1).

Table A6.1. Search algorithm validation with HTA Accelerator data. ✓: found using algorithm. ✗: not found using algorithm. n/a: not applicable. HTA: Health technology assessment. RWD: Real-world data. RWE: Real-world evidence.

Therapy	HTA Body	RWD/RWE output HTA Accelerator	Found using algorithm
Imlygic®	G-BA	No information provided	n/a
	NICE	"RWE was used as supporting evidence" but no further information was provided	n/a
Strimvelis®	NICE	"RWE was used as supporting evidence" but no further information was provided	✓
Kymriah®- DLBCL	G-BA	SCHOLAR-1: a retrospective cohort study to support external comparator; Eyre, 2016; a retrospective cohort study to support external comparator	✓
	NICE	"RWE was used as supporting evidence" but no further information was provided	n/a
Kymriah®- ALL	G-BA	CIBMTR registry: patient disease registry to support effectiveness; MT103-205: a prospective cohort stody to inform effectiveness; Hijiya et al., 2001: an observational study to inform effectiveness; PEDICAR: a prospective cohort to support effectiveness; CTL019B2001X: a prospective cohort to support effectiveness	✓
	NICE	"RWE was used as supporting evidence" but no further information was provided	n/a
Yescarta®	G-BA	"RWE was used as supporting evidence" but no further information was provided	n/a
	NICE	"RWE was used as supporting evidence" but no further information was provided	n/a
Luxturna®	G-BA	No information provided	n/a
	NICE	"RWE was used as supporting evidence" but no further information was provided	n/a

continues on next page

Zynteglo®	G-BA	No information provided	n/a
	NICE	UK chart review: a retrospective chart review to support utility	✓
Zolgensma®	G-BA	No information provided	n/a
	NICE	Thompson et al., 2017: a cross-sectional study to support utility	✓
Tecartus®	G-BA	SCHOLAR-2: a retrospective cohort study to support external comparator; Eyre 2019: a retrospective cohort study to support external comparator; Jain 2018: an observational study to support external comparator; Martin 2016: a retrospective cohort study to support external comparator; McCulloch 2020: a retrospective cohort study to support external comparator; Epperla 2017: a retrospective cohort study to support external comparator; Wang 2017: an observational study to support external comparator;	✓
	NICE	McCulloch et al., 2020: a retrospective cohort study to support external comparator	✓
OTL-200/ Libmeldy®	G-BA	No information provided	n/a
	NICE	Mahmood et al., 2010: a retrospective cohort study to support burden of illness	✓

For three indications (Imlygic®, Yescarta®and Zolgensma®), data from the search algorithm was compared to the data extracted from reading the HTA reports for the use cases (Table A6.2).

Table A6.2. Search algorithm validation with data used with data extraction form. n/a: not applicable. DEF: Data extraction form. HTA: Health technology assessment. RWD: Real-world data. RWE: Real-world evidence.

Therapy	HTA Body	Algorithm			DEF		
		Accepted	Not accepted	Not identified	Accepted	Not accepted	Not identified
Imlygic®	G-BA	0	0	0	0	0	0
	NICE	1	2	4	1	2	4
Yescarta®	G-BA	1	12	1	1	13	1
	NICE	3	0	1	4	0	5
Zolgensma®	G-BA	2	2	0	5	5	6
	NICE	7	12	8	7	12	16

A similar approach was used to validate the interpretation of identified RWD/RWE appraisals from use cases with the HTA Accelerator (Table A6.3). Here, the HTA Accelerator data was considered the ‘golden standard’, as

experts have interpreted the reports with more knowledge than the researcher.

Table A6.3. Use case appraisal validation with HTA Accelerator data. cmark: found using algorithm. ✗: not found using algorithm. n/a: not applicable. Green: Accepted. Red: not accepted. HTA: Health technology assessment. RWD: Real-world data. RWE: Real-world evidence.

Therapy	HTA Body	RWD/RWE source & appraisal from HTA Accelerator	In line with use case
Imlygic®	G-BA	No information provided	n/a
	NICE	"RWE was used as supporting evidence" but no further information was provided	n/a
Yescarta®	G-BA	"RWE was used as supporting evidence" but no further information was provided	n/a
	NICE	"RWE was used as supporting evidence" but no further information was provided	n/a
Zolgensma®	G-BA	No information provided	n/a
	NICE	Thompson et al., 2017	✓

A7

RWD/RWE usage by NICE & G-BA

A7.1. Characterising RWD/RWE usage

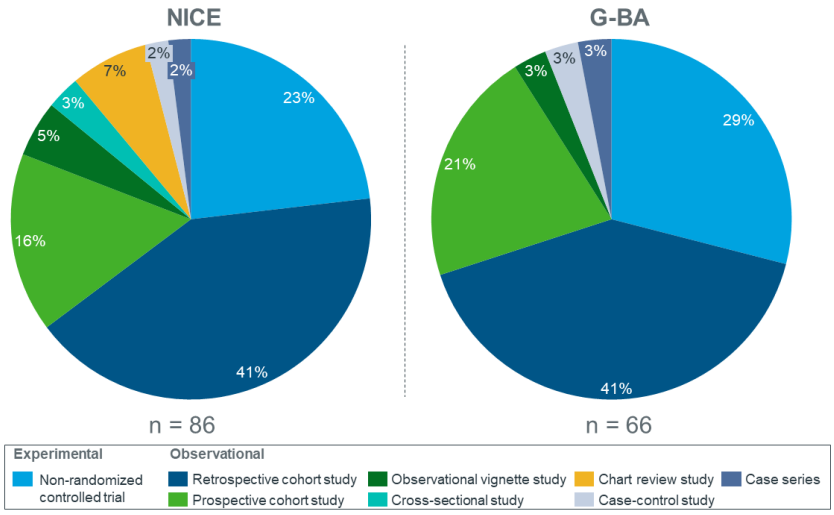


Figure A7.1. RWE usage by NICE and G-BA. RWE: Real-world evidence.

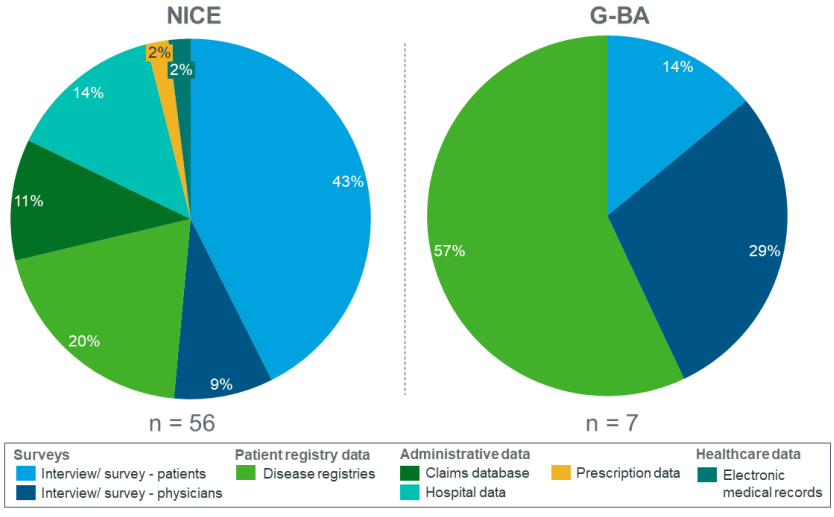


Figure A7.2. RWD usage by NICE and G-BA. RWD: Real-world data.

A7.2. Areas supported by RWD/RWE

Table A7.1. Comparison of areas supported by RWD/RWE in G-BA and NICE gene therapy HTAs. *Only applicable for RWD/RWE usage of NICE Libmeldy and Zynteglo. OS: Overall survival. PFS: Progression free survival. QoL: Quality of life.

Area supported	Germany (G-BA)				England (NICE)				
	Accepted	Not accepted	Not identified	Total	Accepted	Not accepted	Not identified	Other*	Total
Burden of illness	0	0	7	7	1	1	13	9	24
Epidemiology	0	0	1	1	0	1	1	2	4
Patient characteristics	0	0	4	4	0	0	2	5	7
Treatment pathway	0	0	0	0	0	0	2	0	2
External comparator	1	30	3	34	7	11	9	4	31
Effectiveness - comparator	2	0	0	2	1	0	1	0	2
Effectiveness - intervention	12	6	0	18	13	3	0	7	23
Extrapolation of OS - comparator	0	0	0	0	3	0	1	0	4
Extrapolation of OS - intervention	0	0	0	0	1	3	1	0	5
Extrapolation of PFS - comparator	0	0	0	0	0	0	0	1	1
Safety - comparator	0	0	0	0	0	0	1	0	1
Safety - intervention	3	2	0	5	2	1	1	2	6
Validation of surrogate endpoints	0	0	2	2	0	0	0	0	0
QoL - comparator	0	1	0	1	0	5	0	0	5
QoL - intervention	0	0	0	0	0	0	1	0	1
Costs	0	0	0	0	1	0	7	1	9
Disutility	0	0	0	0	1	0	2	0	3
Health resource transition probability	0	0	0	0	0	0	3	0	3
Resource utilisation	0	0	0	0	0	2	8	2	12
Utility	0	0	0	0	5	0	3	6	14
Total	18	39	17	74	35	27	56	39	157

Table A7.2. Overlapping and unique RWD/RWE sources per area supported. Excluding Strimvelis. OS: Overall survival. PFS: Progression free survival. QoL: Quality of life. RWD: Real-world data. RWE: Real-world evidence.

Area supported	Overlapping RWD/RWE sources	G-BA unique	NICE unique
Burden of illness	4	1	19
Epidemiology	4	1	0
Patient characteristics	2	1	5
Treatment pathway	1	0	1
External comparator	16	15	14
Effectiveness - comparator	2	0	0
Effectiveness - intervention	14	4	4
Extrapolation of OS - comparator	3	0	1
Extrapolation of OS - intervention	2	0	3
Extrapolation of PFS - comparator	0	0	1
Safety - intervention	5	0	0
Validation of surrogate endpoints	0	2	0
QoL - comparator	1	0	4
QoL - intervention	1	0	0
Costs	0	0	8
Disutility	0	0	3
Health resource transition probability	0	0	3
Resource utilisation	0	0	12
Utility	0	0	13
Total	51	24	95

A8

RWD/RWE usage in illustrative use cases

A8.1. Imlygic®

Table A8.1. Imlygic RWD/RWE usage G-BA HTA report. RWD: Real-world data. RWE: Real-world evidence.

RWD/RWE source	RWD/RWE type	Area supported	Rationale for inclusion/exclusion
-	-	-	-

Table A8.2. Imlygic RWD/RWE usage NICE HTA report. Rationales directly derived from NICE (2016). RWD: Real-world data. RWE: Real-world evidence.

RWD/RWE source	RWD/RWE type	Area supported	Rationale for inclusion/exclusion
AJCC registry	Disease registry	Extrapolation of OS - intervention	"The committee concluded that, because of the lack of suitable effectiveness inputs in the economic model, it had not been presented with a plausible incremental cost-effectiveness ratio (ICER) for talimogene laherparepvec compared with ipilimumab."
Mortality data from life tables	Hospital data	Extrapolation of OS - intervention	"The committee concluded that, because of the lack of suitable effectiveness inputs in the economic model, it had not been presented with a plausible incremental cost-effectiveness ratio (ICER) for talimogene laherparepvec compared with ipilimumab."
Mols et al., 2010	Interview/ survey - patients	Disutility	Not identified
Linker, 2013	Electronic medical records	Costs	Not identified
		Resource utilisation	Not identified
MELODY	Retrospective cohort study	Resource utilisation	Not identified

A8.2. Yescarta®

Table A8.3. Yescarta RWD/RWE usage G-BA HTA report.Rationales derived from Gemeinsamer Bundesausschuss (2016). RWD: Real-world data. RWE: Real-world evidence.

RWD/RWE source	RWD/RWE type	Area supported	Rationale for inclusion/exclusion
Aurer et al., 2002	Retrospective cohort study	External comparator	Partly include only subpopulations from the identified studies, for which, however, no information on patient characteristics is available, therefore an assessment of comparability with the patients of the ZUMA-1 study is not possible.
Eyre et al., 2016	Interview/ survey - physician	External comparator	Relevant differences of the patient characteristics in comparison to the ZUMA-1 study (e.g. with regard to the age of the patients) were found
Pan et al., 2002	Retrospective cohort study	External comparator	Only sub-populations were selected for indirect comparison. No patient characteristics are available for these specifically selected sub-populations; comparability with the ZUMA-1 study can therefore not be assessed.
Armand et al., 2008	Retrospective cohort study	External comparator	Information on relevant patient characteristics of the specifically selected comparison populations is equally missing
Avivi et al., 2014	Retrospective cohort study	External comparator	The comparability of the patient populations is not given, for example because of significant differences in the age of the patients
Bacher et al., 2012	Retrospective cohort study	External comparator	Partly include only subpopulations from the identified studies, for which, however, no information on patient characteristics is available, therefore an assessment of comparability with the patients of the ZUMA-1 study is not possible.
Fenske et al., 2016	Retrospective cohort study	External comparator	Partly include only subpopulations from the identified studies, for which, however, no information on patient characteristics is available, therefore an assessment of comparability with the patients of the ZUMA-1 study is not possible.
Ghobadi et al., 2015	Retrospective cohort study	External comparator	The comparability of the patient populations is not given, for example because of significant differences in the age of the patients
Heinzelmann et al., 2018	Retrospective cohort study	External comparator	Partly include only subpopulations from the identified studies, for which, however, no information on patient characteristics is available, therefore an assessment of comparability with the patients of the ZUMA-1 study is not possible.
Lazarus et al., 2010	Retrospective cohort study	External comparator	The comparability of the patient populations is not given, for example because of significant differences in the age of the patients

continues on next page

Rigacci et al., 2012	Retrospective cohort study	External comparator	The comparability of the patient populations is not given, for example because of significant differences in the age of the patients
van Kampen et al., 2011	Retrospective cohort study	External comparator	The comparability of the patient populations is not given, for example because of significant differences in the age of the patients
Zentrum für Krebsregisterdaten (ZfKD)	Disease registry	Epidemiology	Not identified
SCHOLAR-1	Retrospective cohort study	External comparator	Despite the uncertainties and possible differences between the patient populations, the present indirect historical comparison with the SCHOLAR-1 study is considered sufficiently valid for the assessment of the extent of the additional benefit, taking into account the inconclusively assessable prognostic significance of the ECOG status, the IPI value, and the disease stage for the further course of therapy in the present treatment situation as well as the advanced, predominantly deterministic disease state of the patient population examined here.
NCI 09-C-0082	Non-randomised controlled trial	Effectiveness - intervention	The NCI 09-C-0082 supportive study is an open, single-arm phase I dose-finding study. In the study, the manufacturing process of Axi-Cel was varied, and various doses of lymphocyte-depleting chemotherapy, most of which do not conform to regulatory requirements, were investigated. The study is therefore not used for the benefit assessment.

Table A8.4. Yescarta RWD/RWE usage NICE HTA report. Rationales directly derived from NICE (2019). RWD: Real-world data. RWE: Real-world evidence.

RWD/RWE source	RWD/RWE type	Area supported	Rationale for inclusion/exclusion
RWE cohort from an Hospital database	Hospital data	External comparator	The committee acknowledged that survival outcomes were very similar using the CORAL and SCHOLAR-1 cohorts. It noted the limited data in the small Oxford audit dataset and agreed not to consider it further.
Eyre et al., 2016	Interview/ survey - physician	Treatment pathway	Not identified
SCHOLAR-1	Retrospective cohort study	Effectiveness - comparator	The committee agreed that there were limitations to all of the potential data sources for the comparator arm but that using patient-level data from the updated adjustments to the SCHOLAR-1 data was most appropriate.
Maurer et al., 2014	Retrospective cohort study	Health state transition probability	Not identified
Nagle et al., 2013	Disease registry	Treatment pathway	Not identified
Kansara et al., 2014	Hospital data	Burden of illness	Not identified

continues on next page

Kochenderfer et al., 2017 | NCI 09-C-0082	Non-randomised controlled trial	Effectiveness - intervention	Not identified

A8.3. Zolgensma®

Table A8.5. Zolgensma RWD/RWE usage G-BA HTA report.*: For orphan drugs, according to the G-BA's Regulation, it is to be taken into account that the information on the the extent of the additional benefit must be based on the marketing authorization and the studies that justifying the approval. Rationales derived from Gemeinsamer Bundesausschuss (2021)

RWD/RWE source	RWD/RWE type	Area supported	Rationale for inclusion/exclusion
LT-001	Prospective cohort study	Safety - intervention	Data from the LT-001 study was submitted with the marketing authorization application in October 2018 and is therefore part of the basis for the approval of Onasemnogen-Abeparvovec*.
		Effectiveness - intervention	Data from the LT-001 study was submitted with the marketing authorization application in October 2018 and is therefore part of the basis for the approval of Onasemnogen-Abeparvovec.
LT-002	Prospective cohort study	Safety - intervention	Ongoing study, no data available and no data included in regulatory submission.
		Effectiveness - intervention	Ongoing study, no data available and no data included in regulatory submission.
Gregoretti et al., 2013	Chart review study	Burden of illness	Not identified
NeuroNext	Prospective cohort study	External comparator	Not identified
PNCR	Hospital data	External comparator	Not identified
CL-101, START	Non-randomised controlled trial	Effectiveness - intervention	The observed results from the ongoing study program are overall in good good agreement with the results of the completed studies CL-303 and CL101, on the basis of which substantial additional benefit can be inferred.
CL-102 STRONG	Non-randomised controlled trial	Effectiveness - intervention	SMA type 2, intrathecal use (off-label).
NCT01839656 (CS3A)	Non-randomised controlled trial	Effectiveness - comparator	Divergent intervention.
Bach et al., 2002	Retrospective cohort study	Patient characteristics	Not identified
SHINE (CS11)	Non-randomised controlled trial	Effectiveness - comparator	Divergent intervention.
Pane et al., 2018	Retrospective cohort study	Burden of illness	Not identified
Swoboda et al., 2005	Prospective cohort study	Burden of illness	Not identified
De Sanctis et al., 2016	Retrospective cohort study	External comparator	Not identified

Table A8.6. Zolgensma RWD/RWE usage NICE HTA report.Rationales directly derived from NICE (2021d). RWD: Real-world data. RWE: Real-world evidence.

RWD/RWE source	RWD/RWE type	Area supported	Rationale for inclusion/exclusion
LT-001	Prospective cohort study	Safety - intervention	The committee concluded that, compared with best supportive care, there are substantial clinical benefits with onasemnogene abeparvovec for people with type 1 SMA. However, it pointed out that, because follow up was short in START and STR1VE-US, the expected long-term outcomes remain uncertain.
		Effectiveness - intervention	The committee concluded that, compared with best supportive care, there are substantial clinical benefits with onasemnogene abeparvovec for people with type 1 SMA. However, it pointed out that, because follow up was short in START and STR1VE-US, the expected long-term outcomes remain uncertain.
LT-002	Prospective cohort study	Safety - intervention	Ongoing study, no data available and no data included in regulatory submission.
		Effectiveness - intervention	Ongoing study, no data available and no data included in regulatory submission.
NeuroNext	Prospective cohort study	External comparator	The committee concluded that NeuroNext was the most appropriate source to estimate outcomes for best supportive car.
		Extrapolation of OS - comparator	The committee concluded that NeuroNext was the most appropriate source to estimate outcomes for best supportive care.
Gregoretti et al., 2013	Chart review study	Extrapolation of OS - comparator	The ERG and committee considered that the company's approach to estimating long-term outcomes was appropriate, but that there was a lack of long-term data to inform these assumptions.
PNCR	Hospital data	External comparator	Not identified
Prescription cost analysis	Prescription data	Resource utilisation	Not identified
Strauss et al., 2018	Prospective cohort study	QoL- comparator	Excluded from cost-effectiveness analysis as HRQoL is not reported by motor function status or SMA type, but by SMN2 copy number only.
Zerres et al., 1997	Retrospective cohort study	Extrapolation of OS - intervention	Not identified
Kissel et al., 2001	Prospective cohort study	QoL- comparator	Excluded from cost-effectiveness analysis as Did not include all health states (included SMA type 3 patients, aged 3–17 years) and used PedsQL, which would require use of mapping that is associated with methodological limitations.

continues on next page

Klug et al., 2016	Cross-sectional study	Resource utilisation	Excluded from cost-effectiveness analysis as the study used PedsQL, which would require use of mapping that is associated with methodological limitations.
Lopez-Bastida et al., 2017	Cross-sectional study	Resource utilisation	Excluded from cost-effectiveness analysis as it was deemed more appropriate to use the UK parent-proxy cohort only
Zuluaga et al., 2017	Observational vignette study	QoL- comparator	Excluded from cost-effectiveness analysis as the study is a duplicate; used method reported in Lloyd et al., 2017
Thompson et al., 2017	Observational vignette study	Utility	"Included in scenario analyses that used various alternative health-state utility sources. The committee considered that there was uncertainty around the health-state utilities used in the model and that they had major effect on estimates of cost effectiveness. However, it concluded that they appeared to be the most appropriate to use in decision making."
Lloyd et al., 2017	Observational vignette study	Utility	"Included in scenario analyses that used various alternative health-state utility sources. The committe considered that there was uncertainty around the health-state utilities used in the model and that they had major effect on estimates of cost effectiveness. However, it concluded that they appeared to be the most appropriate to use in decision making."
Tilford et al., 2005	Interview/ survey - patients	Disutility	Not identified
RESTORE registry	Disease registry	Effectiveness - intervention	"RESTORE is a prospective, long-term registry initiated by AveXis, of patients who have been diagnosed with SMA.. The current data available from the registry are limited to ** patients and outcome data presented in the CS appendix were limited to survival data reporting ********** are still alive as of 31 January 2020 data cut. The ERG does not discuss these data further as data are not available for other outcomes of relevance to the NICE decision problem."
Bladen et al., 2014	Disease registry	Health state transition probability	Not identified
De Sanctis et al., 2016	Retrospective cohort study	External comparator	Not identified
SMA UK Patient and Caregiver survey	Interview/ survey - patients	Costs Not identified	
UK HCRU Survey	Interview/ survey - physician	Costs	Not identified
UK life table data	Hospital data	Health state transition probability	Not identified
Noyes et al., 2006	Interview/ survey - patients	Resource utilisation	Not identified
Alanizi et al., 2018	Retrospective cohort study	Burden of illness	Not identified
Nusinersen UK early access programme (EAP)	Hospital data	Costs	Not identified
Bach et al., 2002	Retrospective cohort study	Patient characteristics	Not identified
NCT01839656 (CS3A)	Non-randomised controlled trial	QoL - intervention	Not identified
SHINE (CS11)	Non-randomised controlled trial	Effectiveness - comparator	Not identified

continues on next page

Swoboda et al., 2005	Prospective cohort study	Burden of illness	Not identified
CL-102 STRONG	Non-randomised controlled trial	Effectiveness - intervention	Onasemnogene abeparvovec was administered via intrathecal administration, which is not relevant to the NICE decision problem and thus this study is not discussed further.
CL-101, START	Non-randomised controlled trial	Effectiveness - intervention	Therefore, the committee considered that the results from the START and STR1VE-US were generalisable to people with type 1 SMA with up to 3 copies of SMN2 gene. However, it recognised that no evidence was presented for babies with type 1 SMA who were older than 6 months at treatment administration and this was a key limitation.