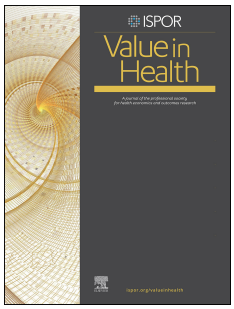

Evidence quality and health technology assessment outcomes in re-appraisals of drugs for rare diseases in Germany

Lea A. Wiedmann, MSc, John A. Cairns, MPhil, Ellen Nolte, PhD

PII: S1098-3015(24)02795-5

DOI: https://doi.org/10.1016/j.jval.2024.07.012

Reference: JVAL 4096

To appear in: *Value in Health*

Received Date: 15 December 2023

Revised Date: 8 June 2024

Accepted Date: 8 July 2024

Please cite this article as: Wiedmann LA, Cairns JA, Nolte E, Evidence quality and health technology assessment outcomes in re-appraisals of drugs for rare diseases in Germany, *Value in Health* (2024), doi: https://doi.org/10.1016/j.jval.2024.07.012.

Target Journal: VIH

Title: Evidence quality and health technology assessment outcomes in re-appraisals of drugs for rare diseases in Germany

Authors: Lea A. Wiedmann, MSc[1], John A. Cairns, MPhil[1], Ellen Nolte, PhD[1]

[1] Department of Health Services Research and Policy, London School of Hygiene & Tropical Medicine, London, UK

Corresponding Author information:
Lea A. Wiedmann, MSc
Department of Health Services Research and Policy
London School of Hygiene & Tropical Medicine
15-17 Tavistock Place
London WC1H 9SH, United Kingdom
Email: lea.wiedmann@lshtm.ac.uk.
Phone: +44 (0) 20 7636 8636

Précis: Clinical evidence quality in initial appraisals and re-appraisals of rare disease treatments (RDTs) in Germany is limited and high clinical benefit ratings are uncommon

Word count: 4,097
Number of Pages: 20
Number of Figures: 1
Number of Tables: 5
Supplementary material:
Pages: 25
Figures: 2
Tables: 25

Author Contributions:
Concept and design: Lea A. Wiedmann, John A. Cairns, Ellen Nolte
Acquisition of data: Lea A. Wiedmann
Analysis and interpretation of data: Lea A. Wiedmann, John A. Cairns
Drafting of the manuscript: Lea A. Wiedmann
Critical revision of the paper for important intellectual content: Lea A. Wiedmann, John A. Cairns, Ellen Nolte
Statistical Analysis: Lea A. Wiedmann
Provision of study materials or patients: n/a
Obtaining funding: n/a
Administrative, technical, or logistic support: n/a
Supervision: John A. Cairns, Ellen Nolte
Other (if applicable, please specify): n/a

Funding/Support: Lea A. Wiedmann is supported by the Economic and Social Research Council (ESRC) (grant number: ES/P000592/1).

Role of the Funder/Sponsor: The funder was not involved in any aspect of the study conception and conduct, or the decision to submit the paper for publication.

Acknowledgements: We gratefully acknowledge the valuable input from Orlagh Carroll and the very helpful discussions with colleagues at the London School of Hygiene and Tropical Medicine, especially Andrew Briggs, Luke Vale, Jiyeon Kang, Olimpia Lamberti, Silvia Moler-Zapata, and David Lugo-Palacios.

Highlights:

- Planned re-appraisal can facilitate access to innovative health technologies because health technology assessment (HTA) agencies can accept greater uncertainty at the initial appraisal. There is limited evidence on how re-appraisal practices impact HTA outcomes.
- In Germany, for the majority of re-appraisals that had used a limited appraisal process and followed a time-limited initial decision, maturity of survival data and the clinical benefit rating (CBR) did not change despite additional evidence being available. For the majority of re-appraisals that had used a regular appraisal process because the RDT exceeded the revenue threshold, no additional evidence was available, maturity of survival data did not change, and the CBR decreased.
- Reasons for conducting re-appraisals of RDTs in Germany vary and the process followed for re-appraisal determines which CBRs are possible. Consideration of re-appraisal practices and outcomes for RDTs in Germany can enhance understanding of the opportunities and challenges of conducting re-appraisals for RDTs.

Evidence quality and health technology assessment outcomes in re-appraisals of drugs for rare diseases in Germany

Objective

Evidence on re-appraisals of health technologies in Germany is limited, and for rare disease treatments (RDTs) the Federal Joint Committee follows different processes (limited or regular) depending on whether an annual revenue threshold has been exceeded. Our objective is to better understand (re)-appraisal processes and their outcomes for RDTs in Germany.

Methods

We analysed appraisal documents of 55 RDT indications for which an initial appraisal and a re-appraisal were conducted between 2011 and 2023. We extracted information for the type of evidence, the risk of bias, the availability of additional evidence, and the change in the maturity of survival data as proxies for evidence quality. Specifically, we reviewed the reasons for conducting re-appraisals; examined how evidence quality and the clinical benefit rating (CBR) differed between initial appraisals and re-appraisals; and explored the association between evidence quality and (i) the CBR and (ii) the change in the CBR following re-appraisal.

Results

Most re-appraisals were conducted because the annual revenue threshold was exceeded, or the initial appraisal resolution was time-limited. Almost all initial appraisals used the limited process, while the majority of re-appraisals used the regular process. The CBR increased in only

nine and decreased in 21 of 55 re-appraisals. There was some evidence that re-appraisals with an accepted randomised controlled trial were significantly more likely to achieve a higher CBR.

Conclusions

Findings confirmed that reasons and processes for conducting re-appraisals of RDTs in Germany differ. Further, high CBRs in re-appraisals were not common and evidence quality in initial appraisals and re-appraisals was limited.

Keywords: Germany, rare disease treatments, Federal Joint Committee, health technology assessment, AMNOG

Introduction

Regulatory approvals of rare disease treatments (RDTs) are frequently based on fewer and lower quality clinical data compared to non-RDTs[1-3]. Evidence of their clinical benefit is typically limited or incomplete due to small patient populations, short trial durations, a lack of active comparators and uncertainty about standards of care, and tools to measure patient outcomes[4-7]. This means that health technology assessment (HTA) agencies who inform or are responsible for funding decisions will have to balance substantial uncertainties in appraising the effectiveness of treatments[8-10] with the need to meet the health and care needs of people with rare diseases.

Planned re-appraisal can facilitate access to innovative health technologies because greater uncertainty can be accepted at the initial appraisal[11]. Re-appraisals are typically conducted when additional evidence has become available[12], although processes vary[13]. As such, re-appraisal provides an opportunity to identify cases where an initial positive recommendation should be rescinded in the light of additional evidence. It has been suggested that differences in re-appraisal practices may explain variation in HTA outcomes across countries[14] but evidence on how re-appraisals impact HTA outcomes remains limited. Recent work examining oncology drugs that had received conditional approval for funding in England, found that re-appraisal of these drugs for routine use in the National Health Service (NHS) had largely led to positive recommendations despite many remaining uncertainties in the evidence base[15].

This study examined re-appraisal processes for RDTs in Germany. We document the evidence quality of RDTs and the extent to which re-appraisals led to changes in clinical benefit ratings (CBRs). Our analysis is timely in light of cost-containment measures introduced by the 2022 health

reform (GKV-FinStG), which reduced the annual revenue threshold by which RDTs are being considered for regular appraisal[16]; this change will make re-appraisals of RDTs more likely[17]. Specifically, we analysed the quality of evidence and the CBR for indications for which both an initial appraisal and a re-appraisal are available. We reviewed the reasons for conducting re-appraisals, examined how evidence quality and the CBR differed between initial appraisals and re-appraisals, and explored the association between evidence quality and (i) the CBR in re-appraisals and (ii) the change in the CBR following re-appraisal.

Appraisal of health technologies in Germany

All newly authorised health technologies can enter the market immediately and require an early benefit assessment (EBA)[18]. The process of EBA and price negotiation is sometimes referred to as the 'AMNOG process', following the title of the legislation introducing it in 2011[19]. EBAs are generally performed when a new health technology or indication for an already licensed technology enters the German market[20]. The EBA is a two-stage appraisal process whereby the independent Institute for Quality and Efficiency in Health Care (IQWiG) usually conducts an initial advisory assessment of patient-relevant outcomes of the technology based on the clinical evidence submitted by the manufacturer[21]. In the benefit dossier, the manufacturer usually has to demonstrate the clinical benefit of the technology over an appropriate comparator therapy (ACT) as determined by the Federal Joint Committee (GBA), the highest decision-making body in the German statutory system[18]. This is followed by a final appraisal by the GBA of the added clinical benefit of the technology over the ACT. The CBR provided in the appraisal resolution forms the basis for price negotiations between the national payer (National Association of Statutory Health Insurance Funds, GKV-Spitzenverband) and the manufacturer[22, 23]. Price negotiations take account

of the CBR of the technology in question and the nature of the appropriate comparator, such as whether the ACT is patent protected or within the data exclusivity period[16].

In Germany, RDT appraisals currently follow one of two routes (figure 1). RDTs will typically undergo what is referred to as 'limited appraisal'. In this case, added clinical benefit is assumed by virtue of the RDT in question having received marketing authorisation at the European level[24]. The GBA does not define an ACT (as it would do in a regular appraisal) and only quantifies the *extent of added clinical benefit* on the basis of (comparative) evidence provided by the manufacturer using four categories (non-quantifiable added benefit, minor, considerable, major)[25-28]. However, RDTs that have exceeded the annual revenue threshold of €30 million across indications (until December 2022: €50 million) undergo a 'regular appraisal', which requires manufacturers to submit a new dossier demonstrating the benefit of the RDT compared to an ACT specified by the GBA[29]. The regular appraisal also considers two additional categories of the extent of added clinical benefit (no added benefit; less benefit than the ACT)[30].

Frequently, technologies, including RDTs, are re-appraised, for example, when an initial appraisal resolution was time-limited, new scientific evidence has become available and either the GBA or the manufacturer request a re-appraisal, a routine practice data collection demanded by the GBA was completed, health technologies exceed the insignificance threshold of €1 million per year, or RDTs exceed the annual revenue threshold[31].

Methods

This study reviewed the reasons for conducting re-appraisals, examined evidence quality and CBRs in initial appraisals and re-appraisals, and analysed changes in CBRs following re-appraisal. We assessed RDT evidence quality based on various extracted data that we used as proxies for evidence quality: change in the maturity of survival data, the type of evidence, the risk of bias, and the availability of additional evidence. Our primary analysis sought to explore the association between evidence quality and the CBR in re-appraisals and the change in the CBR following re-appraisal. We used descriptive statistics and regression analyses in this assessment (see details below). Descriptive statistics included an overview of the characteristics of both initial appraisals and re-appraisals. Further, we presented characteristics of re-appraisals conducted for time-limited initial appraisals and those that exceeded the revenue threshold. In addition, we presented cross-tabulations for (1) the type of re-appraisal by reasons for re-appraisal and (2) the change in the CBR by the availability of additional evidence.

We also conducted a targeted review of the evidence base of the subset of re-appraisals in which the CBR either increased or decreased following re-appraisal (n=30). The aim of this review was to understand the type of the evidence submitted for these indications and why the GBA gave its CBR, including why submitted evidence was sometimes not used. This provided an additional opportunity to identify criteria that were important for decisions on the CBR in re-appraisals. Table 1 provides an overview of the identification variables, and variables extracted for the primary statistical analysis and the targeted review of the evidence base.

Finally, in a supplementary analysis we provided descriptive statistics of the summary assessment of the clinical benefit for each outcome category (mortality, morbidity, including adverse events, and quality of life) by the GBA for those re-appraisals for which additional evidence was available (n=32). This was done to explore potential changes in the clinical benefit in each outcome category between initial appraisals and re-appraisals (see supplementary table S19 for additional details).

Appraisals were identified from the publicly available GBA appraisal database[32], using its rare disease filter to narrow searches to appraisals completed between January 2011 and September 2023. We selected only RDT appraisals for which there were initial appraisals and re-appraisals, allowing a comparison. This comprised 55 indications (with 110 CBRs for initial appraisals and re-appraisals) and 34 different RDTs (see supplementary figure S1 for further details on the appraisal selection process).

The principal sources of data are publicly available technology appraisal documents. Appraisal documents comprise (i) the evidence submission by the manufacturer *(Dossier)*, (ii) the benefit assessment conducted by GBA or IQWiG *(Nutzenbewertung)*, (iii) the justification document (*Tragende Gründe*), and (iv) the appraisal resolution (*Beschlusstext*) (see supplement 2 for additional descriptions).

We treated the CBR in re-appraisals and the change in the CBR following re-appraisal as dependent variables, and the change in the maturity of survival data, the type of evidence, the risk of bias, and the availability of additional evidence as independent variables. We report descriptive statistics for all variables for both initial appraisals and re-appraisals. Data extraction was done by the first

author and the categorisation of variables for which we did not directly rely on the information as reported in the appraisal documents was discussed and agreed with a second author.

We explored whether evidence quality as measured by several proxy variables could explain the CBR in re-appraisals or the change in the CBR in re-appraisals when compared to initial appraisals. We conducted univariable and multivariable ordinal logistic regression analyses[33] with the CBR as the dependent variable (1)

$$\text{logit}(P(Y \geq j)) = \alpha_j + \beta_{jMAT}MAT + \beta_{jTOE}TOE + \beta_{jROB}ROB + \beta_{jAAE}AAE \quad (1)$$

for a given category j, where Y represents the CBR in re-appraisals (ordinal outcome variable with three categories: no added clinical benefit/less benefit than ACT (reference category), non-quantifiable/minor, and considerable/major). As none of the re-appraisals in our dataset was assigned a benefit less than the ACT, all observations in the reference category were re-appraisals with the CBR of no added benefit (see supplementary table S1 for further descriptions of the CBR categories). Due to the small number of observations in the categories major (n=2) and minor added benefit (n=5), we have grouped the observations into three categories.

In addition, we conducted univariable and multivariable binary logistic regression analyses[33] with the change in the CBR as dependent variable (2)

$$\text{logit}(P(Y = 1)) = \alpha_0 + \beta_{MAT}MAT + \beta_{TOE}TOE + \beta_{ROB}ROB + \beta_{AAE}AAE \quad (2)$$

where Y represents the change in the CBR following re-appraisal (binary outcome variable with categories: no change/increase in the CBR following re-appraisal ($Y = 1$) versus a decrease in the CBR following re-appraisal ($Y = 0$)). Due to the small number of observations for which we recorded an increase in the CBR (n=9), we have grouped the observations into two categories.

In both equations, α is the intercept and β denotes the coefficients. MAT denotes the change in the maturity of survival data, measured as the change in the maturity of survival based on the proportion of deaths in the intervention arm of the main study; TOE denotes the type of evidence, measured as the availability of at least one RCT accepted by the GBA; ROB denotes the risk of bias in relation to the quality of the main study (the ROB is affected by the study type, design, and conduct, and the availability of information[34]), and is measured in this study based on the assessment of the GBA or IQWiG; AAE denotes the availability of additional evidence, measured by the availability of a new study, a new data cut of a known study, a new meta-analysis, or a new indirect treatment comparison at re-appraisal compared to initial appraisal (further detailed in table 1 and supplementary table S7).

We assessed the relationship between all variables which are considered proxies for evidence quality using Pearson's Chi-squared tests. Because of its strong positive relationship with the TOE variable, we excluded the ROB variable from the final models. As a result, we estimated the final ordinal logistic regression model and the binary logistic regression model with MAT, TOE, and AAE as independent variables. We chose to include TOE as an independent variable to assess the impact of at least one RCT accepted by the GBA on the CBR. We reported clustered standard errors at RDT level because our dataset included several RDTs for which different indications were

appraised. We tested the parallel regression assumption, i.e. the relationship between predictors and the outcome is consistent across all outcome levels, in the ordinal logistic regression models using a Brant test[35]. We performed all analyses and visualisations of the data using R version 4.3.1[36] (see supplement 5 for a list of R packages used).

Results

Descriptive statistics

Almost half of the appraisals were for oncological conditions (47%) (table 2). The two most common reasons for conducting re-appraisals were exceeding the revenue threshold (49%) or time-limited initial appraisals (38%). For most initial appraisals (85%) the CBR was either minor or non-quantifiable. Conversely, a third of re-appraisals (33%) found no added clinical benefit. In these cases, all initial appraisals followed the limited appraisal process. More than half of re-appraisals (62%) followed the regular appraisals route. Compared to initial appraisals, the accepted evidence at re-appraisal included at least one randomised controlled trial (RCT) less often (55% versus 75%). Compared to re-appraisals, more initial appraisals had main studies with a low risk of bias rating (53% versus 38%). Additional evidence was submitted for more than half of re-appraisals (58%), and there was no change in the maturity of survival data for the majority of re-appraisals (75%). The CBR increased only for nine of 55 (16%) indications and decreased for 21 of 55 (38%) indications. The CBR did not change for 20 of 55 (36%) indications, despite additional evidence being available (see supplementary table S10).

Considering the subset of re-appraisals of time-limited initial appraisals, all but one re-appraisal (95%) were conducted using the limited appraisal process (see supplementary table S8-S9 for additional descriptive statistics). Only eight of 21 (38%) re-appraisals were rated as providing a minor, considerable or major benefit. Almost all re-appraisals drew on additional evidence but only in a minority was there a change in the maturity of survival data or an increase in the CBR (43% and 19% respectively) compared to the initial appraisal. Among re-appraisals conducted because

the revenue threshold was exceeded, only nine of 27 (33%) achieved a minor, considerable or major benefit; the CBR decreased in more than half (59%); in most no additional evidence was available (67%) and the maturity of survival data did not change (89%). Limited re-appraisals were mostly conducted because an initial appraisal resolution was time-limited (20 of 21) and regular re-appraisals because of exceeding the revenue threshold (27 of 34) or the removal of the RD status (5 of 34).

Lastly, in our supplementary analysis we found no change in the summary assessment of the clinical benefit in the majority of re-appraisals for which additional evidence was available when compared to initial appraisals (see supplementary tables S20-S24 for additional descriptive statistics). In those appraisals where a change was recorded, there was no clear trend in the direction of results (positive or negative).

Statistical analyses

Table 3 reports the results of the multivariable ordinal logistic regression model for re-appraisals with the CBR as dependent variable. It shows that re-appraisals for which accepted RCT evidence was available were significantly more likely (p = 0.001) to achieve a higher CBR (adjusted odds ratio (AOR) = 34.82, 95% CI 5.10, 237.83). There was no significant association between a higher CBR and re-appraisals that had more mature survival data (AOR = 4.90, 95% CI 0.72, 33.23, p = 0.110) or for which additional evidence was available (AOR = 5.46, 95% CI 0.94, 31.55, p = 0.064). Results of the Brant test confirmed that the parallel regression assumption holds (see supplementary tables S11).

Table 4 reports AORs for the multivariable binary logistic regression model for re-appraisals with the change in the CBR as dependent variable. Re-appraisals for which accepted RCT evidence compared to other types of evidence (AOR = 11.81, 95% CI 2.53, 55.13) and additional evidence compared to no additional evidence (AOR = 18.75, 95% CI 2.05, 171.77) was available were significantly associated (p = 0.002 and p = 0.009, respectively) with no change or an increase in the CBR. There was no significant association between re-appraisals where the maturity of the survival data had increased and no change or an increase in the CBR (AOR = 4.44, 95% CI 0.30, 66.27, p = 0.279).

Targeted review of evidence quality

Table 5 provides an overview of the changes in the CBR following re-appraisal. In nine re-appraisals, the CBR was higher than in the initial appraisal. In these re-appraisals, the submitted data included new studies or later data cuts for previously submitted trials (see supplementary table S15), and we recorded a change in the maturity of survival data in five re-appraisals.

In 21 re-appraisals, the CBR decreased compared to the initial appraisal. In three of these 21 re-appraisals the CBR decreased from a minor to a non-quantifiable benefit (see supplementary table S18). This was because additional data from a new study was not used in the re-appraisal because the dosage of the intervention did not correspond to the summary of product characteristics (cabozantinib), or because the evidence for the specific patient group was limited (two ivacaftor indications). In addition, in 18 of these 21 re-appraisals the CBR decreased to no added clinical benefit, which was mainly driven by appraisals that were assigned a minor or non-quantifiable added clinical benefit at initial appraisal where their added clinical benefit was assumed by virtue

of their RDT status (see supplementary tables S16-S17). In 15 of these 18 re-appraisals, no added clinical benefit was assigned because none of the submitted evidence was accepted by the GBA. For example, evidence was not accepted when no direct comparative study with the ACT was submitted but only unadjusted indirect comparisons, the direct evidence did not include the ACT, study duration was too short, no data was submitted, or the evidence dossier submitted by the manufacturer was incomplete and did not meet formal requirements.

Discussion

This study found that re-appraisals of RDTs in Germany were mostly conducted because the RDT under consideration exceeded the annual revenue threshold or the initial appraisal was time-limited. Our results indicated that for the majority of re-appraisals that had used a limited appraisal process and followed a time-limited initial decision, maturity of survival data and the CBR did not change despite additional evidence being available. Our supplementary analysis also found that, based on the summary assessment of outcome categories, few re-appraisals of time-limited initial appraisals demonstrated improvement in mortality, morbidity or quality of life. These findings raise the question about whether and to what extent these re-appraisals provide additional insights beyond the initial appraisal. For most regular re-appraisals that were conducted because RDTs exceeded the annual revenue threshold, no additional evidence was available and maturity of survival data did not change. In addition, there was a decrease in the CBR in more than half of these re-appraisals, which corresponds to findings from IQWiG[37]. Our study added to their findings by expanding the range of re-appraisals, including those for which time-limited initial appraisals were issued, by examining additional proxies of evidence quality, including risk of bias and change in the maturity of survival data, by exploring associations between proxies of evidence quality and the extent and change in the CBR, and by additionally conducting a targeted review of evidence quality in selected appraisals.

Overall, we found that in more than half of re-appraisals additional evidence, mostly in the form of new studies or later data cuts of already submitted studies, was considered for decision-making. Similar observations were reported for re-appraisals of oncology drugs in England[15]. We found some evidence that the availability of additional evidence was associated with no change or an

increase in the CBR, which is unsurprising given that there was no change in the CBR for the majority of re-appraisals for which additional evidence was available.

Our study showed some evidence that accepted RCT evidence was associated with a higher CBR in re-appraisals. The decrease in accepted RCTs in re-appraisals compared to initial appraisals indicates that either RCT evidence was not submitted, or for those where RCT evidence was available, this evidence was not accepted for decision-making because of an inadequate dossier or study design. Our findings confirmed that GBA/IQWiG places high value on high-quality RCTs to inform decision-making, as is the case for many other HTA agencies[38-41].

We also found that only approximately one quarter of appraisals had more mature survival data at re-appraisal compared to the initial appraisal. This indicates that immature survival data is common in RDT appraisals. In our analysis, a change in the maturity of survival data was also not significantly associated with a higher CBR or with no change or an increase in the CBR. While overall survival is considered a patient-relevant outcome that is evaluated in each appraisal[42], some RDTs provide more value by improving morbidity or quality of life. However, without a survival benefit, drugs are unlikely to obtain a major CBR, reflecting IQWiG's ranking of outcomes according to their relevance[34] (see supplement 1 for additional information on the methodological approach). The RDT provided substantial survival benefits in all three appraisals assigned major added benefit.

We further noted that in the majority of re-appraisals where the CBR decreased compared to the initial appraisal, the added clinical benefit was not proven because none of the submitted evidence

was accepted by the GBA. Our targeted review found that one reason for this was non-submission of direct comparative analyses with the ACT. The choice of comparator by the manufacturer can thus have major effects on the CBR if there is a mismatch between the ACT and the comparator.

An analysis of downstream consequences of RDT re-appraisals was beyond the scope of this paper which reviewed the evidence quality and CBRs in re-appraisals. Following re-appraisal and a new CBR, price re-negotiations can take place which can change reimbursement for the RDT in question. A recent IQWiG analysis found that the majority of RDT re-appraisals conducted under the regular appraisal process resulted in a price reduction[43]. However, it is important to recognise that drug prices are shaped by many other factors, too, such as the number of patients eligible for the technology, the therapeutic area, whether technologies are indicated for paediatric care, costs of comparators, or prices paid in other European countries[28]. Future work should look at changes in price and uptake of RDTs following re-appraisal, including a disaggregation by reason for re-appraisal.

Our findings are relevant to the long-standing debate about the German EBA process for RDTs[17, 37, 43-47], which, as noted above, assumes an added benefit of RDTs by virtue of the RDT in question having received marketing authorisation at European level. RDTs are only required to undergo the regular appraisal process when they exceed a specified annual revenue threshold. As regular appraisal frequently fails to demonstrate an added clinical benefit for the RDT, the assumed clinical benefit has been described as 'fictitious added benefit'[37]. Kranz et al.[46] argue that superior therapeutic benefit should only be granted to RDTs based on robust comparative clinical evidence. In this context, the implementation of the European Union (EU) HTA Regulation (2021/2282) is

notable as it foresees comparing RDTs with the standard of care applicable in individual member states[48]. However, generating the required comparative clinical efficacy data will likely remain challenging, particularly if RDTs are being approved based on premature evidence, including single arm or phase 2 studies. While requiring comparative clinical evidence might lead to more robust and higher quality evidence for new RDTs, potential implications in terms of fewer product launches or delayed approvals could counterbalance the benefit this requirement, albeit to an unknown degree. In Germany recent reform reduced the annual revenue threshold for RDTs from €50 million to €30 million[16]. This will increase the number of regular appraisals for RDTs[17]. A reduced threshold may also result in a higher number of re-appraisals using similar evidence as the initial appraisal. Further research should explore the impact of the lower revenue threshold on the number, evidence quality and CBR in RDT re-appraisals.

The main strength of this study was that it systematically explored issues of re-appraisals of RDTs in Germany, thereby contributing to an emerging field of analyses of HTA re-appraisals. However, there was limited scope for comparison, given the relatively small number of studies evaluating HTA re-appraisals for medicines, particularly RDTs. The main limitation of this study was the small number of eligible re-appraisals. Conclusions about the association between proxies of evidence quality and (i) the CBR in re-appraisals and (ii) the change in the CBR following re-appraisal were uncertain, as reflected in the large confidence intervals. These should be interpreted with some level of caution. The limited number of observations also influenced the coding decisions for some variables used in our statistical analyses. For example, we were only able to consider three categories of benefit (instead of five) and two categories of change in the CBR (instead of three). A minor limitation was that we excluded risk of bias as independent variable

from the models due to multicollinearity; however, we did not deem this to introduce bias because much of the relevant information was contained in type of evidence variable.

Conclusion

This study has highlighted the importance of distinguishing between the reasons for re-appraisal (particularly exceeding the revenue threshold and time-limited initial appraisals) and appraisal processes (limited versus regular) for RDTs in Germany. Overall, our findings confirmed that high CBRs for RDTs in re-appraisals were not common in the German EBA process and that the quality of evidence submitted for both initial appraisals and re-appraisals of RDTs was limited.

References

1. Pregelj L, Hwang TJ, Hine DC, et al. Precision Medicines Have Faster Approvals Based On Fewer And Smaller Trials Than Other Medicines. *Health Affairs*. 2018;37(5):724-31. https://doi.org/10.1377/hlthaff.2017.1580
2. Kesselheim AS, Myers JA, Avorn J. Characteristics of clinical trials to support approval of orphan vs nonorphan drugs for cancer. *JAMA*. 2011;305(22):2320-6.
3. Michaeli DT, Michaeli CT. 1696MO Partial orphan cancer drugs: FDA approval, clinical benefit, trials, epidemiology, price, beneficiaries, and spending. *Ann Oncol*. 2023;34:S927.
4. Nicod E, Annemans L, Bucsics A, et al. HTA programme response to the challenges of dealing with orphan medicinal products: Process evaluation in selected European countries. *Health Policy*. 2019;123(2):140-51. https://doi.org/10.1016/j.healthpol.2017.03.009
5. Annemans L, Makady A. TRUST4RD: tool for reducing uncertainties in the evidence generation for specialised treatments for rare diseases. *Orphanet J Rare Dis*. 2020;15:1-9.
6. Nicod E, Meregaglia M, Whittal A, et al. Consideration of quality of life in the health technology assessments of rare disease treatments. *Eur J Health Econ*. 2022;23(4):645-69. https://doi.org/10.1007/s10198-021-01387-w
7. Nestler-Parr S, Korchagina D, Toumi M, et al. Challenges in Research and Health Technology Assessment of Rare Disease Technologies: Report of the ISPOR Rare Disease Special Interest Group. *Value Health*. 2018;21(5):493-500. https://doi.org/10.1016/j.jval.2018.03.004
8. Vreman RA, Naci H, Goettsch WG, et al. Decision Making Under Uncertainty: Comparing Regulatory and Health Technology Assessment Reviews of Medicines in the United States and Europe. *Clin Pharmacol Ther*. 2020;108(2):350-57. https://doi.org/10.1002/cpt.1835
9. Bloem LT, Vreman RA, Peeters NWL, et al. Associations between uncertainties identified by the European Medicines Agency and national decision making on reimbursement by HTA agencies. *Clin Transl Sci*. 2021;14(4):1566-77. https://doi.org/10.1111/cts.13027
10. Vreman RA, Bouvy JC, Bloem LT, et al. Weighing of Evidence by Health Technology Assessment Bodies: Retrospective Study of Reimbursement Recommendations for Conditionally Approved Drugs. *Clin Pharmacol Ther*. 2019;105(3):684-91. https://doi.org/10.1002/cpt.1251
11. Farmer C, Barnish MS, Trigg LA, et al. An evaluation of managed access agreements in England based on stakeholder experience. *Int J Technol Assess Health Care*. 2023:1-30. https://doi.org/10.1017/S0266462323000478
12. Facey KM, Espin J, Kent E, et al. Implementing Outcomes-Based Managed Entry Agreements for Rare Disease Treatments: Nusinersen and Tisagenlecleucel. *Pharmacoecon*. 2021;39(9):1021-44. https://doi.org/10.1007/s40273-021-01050-5
13. Mendell A, Vannabouathong C, Le K, Dyrda P, Severn M. *Overview of Health Technology Assessment Processes for Time-Limited Recommendations*. Ottawa (ON): Canadian Agency for Drugs and Technologies in Health (CADTH), 2023. https://www.ncbi.nlm.nih.gov/books/NBK594388/. Published May 2023. Accessed March 7, 2024.
14. Vreman RA, Mantel-Teeuwisse AK, Hövels AM, Leufkens HGM, Goettsch WG. Differences in Health Technology Assessment Recommendations Among European Jurisdictions: The Role of Practice Variations. *Value Health*. 2020;23(1):10-16.

15. Kang J, Cairns J. "Don't Think Twice, It's All Right": Using Additional Data to Reduce Uncertainty Regarding Oncologic Drugs Provided Through Managed Access Agreements in England. *Pharmacoecon*. 2023;7(1):77-91.
16. *Gesetz zur finanziellen Stabilisierung der gesetzlichen Krankenversicherung (GKV-Finanzstabilisierungsgesetz) (07.11.2022)*. Accessed November 6, 2023. http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl122s1990.pdf
17. Greiner W, Witte J. AMNOG Report 2023: Das GKV-Finanzstabilisierungsgesetz und seine Auswirkungen. In: Storm A, ed. *Beiträge zur Gesundheitsökonomie und Versorgungsforschung (Band 43)*. Heidelberg: medhochzwei Verlag GmbH, 2023. https://www.dak.de/dak/download/report-2610274.pdf. Accessed December 9, 2023.
18. Gemeinsamer Bundesausschuss (GBA). The benefit assessment of medicinal products in accordance with the German Social Code, Book Five (SGB V), section 35a. Accessed November 06, 2023. https://www.g-ba.de/english/benefitassessment/
19. *Gesetz zur Neuordnung des Arzneimittelmarktes in der gesetzlichen Krankenversicherung (Arzneimittelmarktneuordnungsgesetz – AMNOG) (22.12.2010)*. Accessed April 26, 2024. http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl110s2262.pdf
20. Gemeinsamer Bundesausschuss (GBA). *Verfahrensordnung des Gemeinsamen Bundesauschusses, 5. Kapitel: Bewertung des Nutzens und der Kosten von Arzneimitteln nach §§ 35a und 35b SGB V, § 8 Abs. 1 Nr. 1-2*. GBA, 2023. https://www.g-ba.de/downloads/62-492-3280/VerfO_2023-07-20_iK_2023-11-08.pdf. Published November 8, 2023. Accessed November 18, 2023.
21. Ruof J, Schwartz FW, Schulenburg JM, Dintsios CM. Early benefit assessment (EBA) in Germany: analysing decisions 18 months after introducing the new AMNOG legislation. *Eur J Health Econ*. 2014;15(6):577-89.
22. Stern AD, Pietrulla F, Herr A, Kesselheim AS, Sarpatwari A. The Impact Of Price Regulation On The Availability Of New Drugs In Germany. *Health Aff (Millwood)*. 2019;38(7):1182-87.
23. Haas A, Tebinka-Olbrich A, Erdmann D, et al. Ergebnisse des AMNOG-Erstattungsbetragsverfahrens. In: Schröder H, Thürmann P, Telschow C, et al., eds. *Arzneimittel-Kompass 2022: Qualität der Arzneimittelversorgung*. Berlin, Heidelberg: Springer, 2022. https://doi.org/10.1007/978-3-662-66041-6_19. Accessed November 18, 2023.
24. *Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung, § 35a Bewertung des Nutzens von Arzneimitteln mit neuen Wirkstoffen, Verordnungsermächtigung, Abs. 1 , Satz 3, Nr. 2-3*. Accessed November 6, 2023. https://www.sozialgesetzbuch-sgb.de/sgbv/35a.html
25. Gemeinsamer Bundesausschuss (GBA). *Verfahrensordnung des Gemeinsamen Bundesauschusses, 5. Kapitel: Bewertung des Nutzens und der Kosten von Arzneimitteln nach §§ 35a und 35b SGB V, § 12 Abs. 1 Nr. 1*. GBA, 2023. https://www.g-ba.de/downloads/62-492-3280/VerfO_2023-07-20_iK_2023-11-08.pdf. Published November 8, 2023. Accessed November 18, 2023.
26. Gemeinsamer Bundesausschuss (GBA). *Verfahrensordnung des Gemeinsamen Bundesauschusses, 5. Kapitel: Bewertung des Nutzens und der Kosten von Arzneimitteln nach §§ 35a und 35b SGB V, § 5 Abs. 7 Nr. 1-4*. GBA, 2023. https://www.g-ba.de/downloads/62-492-

3280/VerfO_2023-07-20_iK_2023-11-08.pdf. Published November 8, 2023. Accessed November 18, 2023.
27. Schlander M, Dintsios CM, Gandjour A. Budgetary Impact and Cost Drivers of Drugs for Rare and Ultrarare Diseases. *Value Health*. 2018;21(5):525-31.
28. Worm F, Dintsios CM. Determinants of Orphan Drug Prices in Germany. *Pharmacoecon*. 2020;38(4):397-411.
29. Gemeinsamer Bundesausschuss (GBA). *Verfahrensordnung des Gemeinsamen Bundesauschusses, 5. Kapitel: Bewertung des Nutzens und der Kosten von Arzneimitteln nach §§ 35a und 35b SGB V, § 12 Abs. 1 Nr. 2*. GBA, 2023. https://www.g-ba.de/downloads/62-492-3280/VerfO_2023-07-20_iK_2023-11-08.pdf. Published November 8, 2023. Accessed November 18, 2023.
30. Gemeinsamer Bundesausschuss (GBA). *Verfahrensordnung des Gemeinsamen Bundesauschusses, 5. Kapitel: Bewertung des Nutzens und der Kosten von Arzneimitteln nach §§ 35a und 35b SGB V, § 5 Abs. 7 Nr. 5-6*. GBA, 2023. https://www.g-ba.de/downloads/62-492-3280/VerfO_2023-07-20_iK_2023-11-08.pdf. Published November 8, 2023. Accessed November 18, 2023.
31. Gemeinsamer Bundesausschuss (GBA). *Verfahrensordnung des Gemeinsamen Bundesauschusses, 5. Kapitel: Bewertung des Nutzens und der Kosten von Arzneimitteln nach §§ 35a und 35b SGB V, § 8 Abs. 1 Nr. 3-8*. GBA, 2023. https://www.g-ba.de/downloads/62-492-3280/VerfO_2023-07-20_iK_2023-11-08.pdf. Published November 8, 2023. Accessed November 18, 2023.
32. Gemeinsamer Bundesausschuss (GBA). Nutzenbewertung von Arzneimitteln. Verfahren nach § 35a SGB V (AMNOG). Accessed November 23, 2023. https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/
33. Harrell FEJ. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Cham, Switzerland: Springer International Publishing AG, 2015.
34. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). *Allgemeine Methoden (Version 7.0)*. Köln: IQWiG, 2023. https://www.iqwig.de/methoden/allgemeine-methoden_version-7-0.pdf. Published September 19, 2023. Accessed March 7, 2024.
35. Brant R. Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. *Biometrics*. 1990;46(4):1171-78.
36. R Core Team. R: A Language and Environment for Statistical Computing. Accessed December 9, 2023. https://www.R-project.org/
37. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). *Evidenz zu Orphan Drugs*. Köln: IQWiG, 2021. https://www.iqwig.de/download/ga21-01_evidenz-zu-orphan-drugs_arbeitspapier_v1-0.pdf. Published December 23, 2021. Accessed December 9, 2023.
38. Hogervorst MA, Pontén J, Vreman RA, Mantel-Teeuwisse AK, Goettsch WG. Real World Data in Health Technology Assessment of Complex Health Technologies. *Front Pharmacol*. 2022;13:837302.
39. National Institute for Health and Care Excellence (NICE). *NICE health technology evaluations: the manual.* NICE, 2022. https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741. Published January 31, 2022. Accessed December 9, 2023.

40. Canadian Agency for Drugs and Technologies in Health (CADTH). *Procedures for CADTH Reimbursement Reviews*. CADTH, 2023. https://www.cadth.ca/sites/default/files/Drug_Review_Process/CADTH%20Drug%20Reimbursement%20Review%20Procedures.pdf. Published November 30, 2023. Accessed December 9, 2023.
41. Makady A, Ham RT, de Boer A, et al. Policies for Use of Real-World Data in Health Technology Assessment (HTA): A Comparative Study of Six HTA Agencies. *Value Health*. 2017;20(4):520-32. https://doi.org/10.1016/j.jval.2016.12.003
42. *Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (Arzneimittel-Nutzenbewertungsverordnung - AM-NutzenV), § 2 Satz 3*. Accessed December 9, 2023. https://www.gesetze-im-internet.de/am-nutzenv/__2.html
43. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). *Preis- und Kostenentwicklung von Orphan Drugs*. 2024. https://www.iqwig.de/download/ga22-01_preisund-kostenentwicklung-von-orphan-drugs_arbeitspapier_v1-0.pdf. Published January 17, 2024. Accessed March 1, 2024.
44. Batram M, Witte J, Greiner W, Gensorowsky D. AMNOG-Report 2022: Orphan Drugs - Erstattungs- und Versorgungsherausforderungen. In: Storm A, ed. *Beiträge zur Gesundheitsökonomie und Versorgungsforschung (Band 38)*. Heidelberg: medhochzwei Verlag GmbH, 2022. https://www.dak.de/dak/download/report-2524570.pdf. Accessed December 9, 2023.
45. Greiner W, Witte J, Gensorowsky D, Pauge S. AMNOG-Report 2020: 10 Jahre AMNOG - Rückblick und Ausblick. Heidelberg: medhochzwei Verlag GmbH, 2020. https://www.dak.de/dak/download/report-2335144.pdf. Accessed December 9, 2023.
46. Kranz P, McGauran N, Banzi R, et al. Reforming EU and national orphan drug regulations to improve outcomes for patients with rare diseases. *BMJ*. 2023;381:e072796. https://doi.org/10.1136/bmj-2022-072796
47. Hecken J. Aussagekräftige Studien sind selbst bei Orphan Drugs möglich. *Ärzte Zeitung*. January 19, 2022. Accessed March 1, 2024. https://www.aerztezeitung.de/Politik/Aussagekraeftige-Studien-sind-selbst-bei-Orphan-Drugs-moeglich-426101.html.
48. *Regulation (EU) 2021/2282 of the European Parliament and of the Council of 15 December 2021 on health technology assessment and amending Directive 2011/24/EU*. Accessed November 06, 2023. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R2282

Table 1

Table 1: Overview of the extracted data

Variable	Description	Categories
Identification variables		
Brand name	Brand name of the RDT as recorded in the GBA database.	Narrative description
RDT molecule name	RDT molecule name as recorded in the GBA database.	Narrative description
Indication	Indication as recorded in the GBA database.	Narrative description
Therapeutic area	Therapeutic indication as recorded in the GBA database.	1. Oncological diseases 2. Metabolic diseases 3. Other diseases
Reason for appraisal	The reason for appraisal as recorded in the justification document of the appraisal.	1. New (first) indication 2. New (subsequent) indication for a RDT already on the market 3. Exceeded the €50 million revenue threshold 4. Time-limited initial appraisal 5. Manufacturer's request (new scientific evidence) 6. Removal of RDT status
Clinical benefit rating (CBR)	The clinical benefit rating (CBR) issued by the GBA in the appraisal resolution.	1. Major 2. Considerable 3. Minor 4. Non-quantifiable added benefit 5. No added benefit 6. Less benefit than the ACT
Type of appraisal process	The type of appraisal process conducted.	0. Limited 1. Regular
Primary statistical analysis (n=55)		

Clinical benefit rating (CBR)	The CBR for each appraisal as determined by the GBA.	1. No added benefit/less benefit than ACT 2. Non-quantifiable added benefit/minor 3. Considerable/major
Change in the CBR	Changes in the CBR when comparing re-appraisals to initial appraisals.	0. Decrease 1. No change/increase
Change in the maturity of survival data (MAT)	The change in the maturity of survival was based on the proportion of deaths in the intervention arm of the main study. We categorised re-appraisals as having more mature survival data when there was at least a 5 percentage points increase in the proportion of deaths compared to initial appraisals, which we considered to be an adequate cut-off to reflect meaningful changes in on our dataset; 'no change' includes re-appraisals where no mortality data was submitted or evidence was not accepted.	0. No change (immature survival data) 1. Change (more mature survival data)
Type of evidence (TOE)	The type of evidence accepted by the GBA in the appraisal; 'other' refers to single arm or observational studies, if evidence was not accepted, or no data was submitted.	0. Other 1. At least one RCT
Risk of bias (ROB)	The risk of bias in the main study as determined by the GBA/IQWiG; 'other' refers to a high/unclear risk of bias, or if evidence was not accepted.	0. Other 1. Low
Availability of additional evidence (AAE)	We categorised a re-appraisal to contain 'additional evidence' when a new study, a new data cut of a known study, a new meta-analysis, or a new indirect treatment comparison was available at re-appraisal compared to the initial appraisal. We categorised re-appraisals in which evidence was not accepted as providing no additional evidence.	0. No additional evidence 1. Additional evidence
Targeted review of the evidence base (n=30)		
RDT and indication	RDT molecule name and indication as recorded in the GBA database for appraisals in which the CBR either increased or decreased following re-appraisal.	Narrative description

Evidence base	Brief description of the evidence base, including study type and name.	Narrative description
Reasoning GBA	Brief description of the reasons by the GBA for the CBR in appraisals in which the CBR decreased following re-appraisal.	Narrative description
Type of additional evidence	Brief description of the type of additional evidence in appraisals in which the CBR increased following re-appraisal.	Narrative description

ACT = appropriate comparator therapy; CBR = clinical benefit rating; GBA = Federal Joint Committee; IQWiG = Institute for Quality and Efficiency in Health Care; RCT = randomised controlled trial

Table 2

Table 2: Characteristics of analysed initial appraisals (n=55) and re-appraisals (n=55)

	Number of initial appraisals (%)	**Number of re-appraisals (%)**
Total	55 (100.0)	55 (100.0)
Therapeutic area		
Oncological	26 (47.3)	
Metabolic	16 (29.1)	
Other	13 (23.6)	
Reason for appraisal		
First indication	44 (80.0)	-
Subsequent indication	11 (20.0)	-
Exceeded revenue threshold	-	27 (49.1)
Time-limited initial appraisal	-	21 (38.2)
Manufacturer's request (new scientific evidence)	-	2 (3.6)
Removal of RD status	-	5 (9.1)
Clinical benefit rating (CBR)		
Major	1 (1.8)	2 (3.6)
Considerable	7 (12.7)	13 (23.6)
Minor	19 (34.5)	5 (9.1)
Non-quantifiable added benefit	28 (50.9)	17 (30.9)
No added benefit	-	18 (32.7)
Less benefit than ACT	-	0 (0)
Type of appraisal process		
Regular	2 (3.6)	34 (61.8)
Limited	53 (96.4)	21 (38.2)
Type of evidence		

Other	14 (25.5)	25 (45.5)
At least one RCT	41 (74.5)	30 (54.5)
Risk of bias		
Other	26 (47.3)	34 (61.8)
Low	29 (52.7)	21 (38.2)
Availability of additional evidence		
No additional evidence	-	23 (41.8)
Additional evidence	-	32 (58.2)
Change in the maturity of survival data		
No change (immature data)	-	41 (74.5)
Change (more mature data)	-	14 (25.5)
Change in the CBR		
Decrease	-	21 (38.2)
No change	-	25 (45.5)
Increase	-	9 (16.4)

CBR = clinical benefit rating; RD = rare disease; RCT = randomised controlled trial

Table 3

Table 3: Multivariable ordinal logistic regression model exploring the association between evidence quality and the clinical benefit rating (CBR) in re-appraisals (n=55)

Evidence quality	**AOR**	**95% CI***	**SE***	**p-value**
Change in the maturity of survival data				
No change[a]	-	-	-	-
Change	4.90	0.72, 33.23	0.98	0.110
Type of evidence				
Other[a]	-	-	-	-
RCT	34.82	5.10, 237.83	0.98	0.001
Availability of additional evidence				
No additional evidence[a]	-	-	-	-
Additional evidence	5.46	0.94, 31.55	0.90	0.064

[a] = reference level; AOR = adjusted odds ratio; CI = confidence interval; RCT = randomised controlled trial; RDT = rare disease treatment; SE = Standard Error

* 95% confidence intervals and standard errors account for clustering at RDT level

Table 4

Table 4: Multivariable binary logistic regression model exploring the association between evidence quality and the change in the clinical benefit rating (CBR) in re-appraisals (n=55)

Evidence quality	**AOR**	**95% CI***	**SE***	**p-value**
Change in the maturity of survival data				
No change[a]	-	-	-	-
Change	4.44	0.30, 66.27	1.38	0.279
Type of evidence				
Other[a]	-	-	-	-
RCT	11.81	2.53, 55.13	0.79	0.002
Availability of additional evidence				
No additional evidence[a]	-	-	-	-
Additional evidence	18.75	2.05, 171.77	1.13	0.009

[a] = reference level; AOR = adjusted odds ratio; CI = confidence Interval; RCT = randomised controlled trial; RDT = rare disease treatment; SE = standard error

* 95% confidence intervals and standard errors account for clustering at RDT level

Table 5

Table 5: Re-appraisals with changes in the clinical benefit rating

		Clinical benefit rating following re-appraisal			
		Major	Considerable	Non-quantifiable	No added benefit
Clinical benefit rating at initial appraisal	Considerable	*none*		*none*	2
	Minor	*none*	4	3	7
	Non-quantifiable	1	4		9

green = increase in the CBR; blue = decrease in the CBR