Publication series
INTERDISCIPLINARY PLATFORM ON BENEFIT ASSESSMENT

Volume 9
October 2019

Contextual evidence – Strategies for targeted therapy

Springer Medizin

PUBLICATION SERIES INTERDISCIPLINARY PLATFORM ON BENEFIT ASSESSMENT

VOLUME 1
Four years of AMNOG – Discourse and impulses

VOLUME 2
Clinical studies – Which endpoints count?

VOLUME 3
Adaptive Pathways – Opportunities and risks

VOLUME 4
AMNOG 2.0 – Information problems

VOLUME 5
Evidence gaps – What does registry data offer?

VOLUME 6
Physician information via software – Ways and goals

VOLUME 7
Physician information via software – Orientation or control?

VOLUME 8
European Benefit Assessment – Opportunities and risks

VOLUME 9
Contextual evidence – Strategies for targeted therapy

ALL VOLUMES AVAILABLE ONLINE AT:
HTTPS://WWW.AERZTEZEITUNG.DE/POLITIK_GESELLSCHAFT/GP_SPECIALS/PLATTFORM_ZUR_NUTZENBEWERTUNG/
OR AT HTTPS://TINYURL.COM/ZGFEFLK

Volume 9
October 2019

Contents

Goals of the plattform

Since the introduction of AMNOG in 2011, Germany has a well-established and widely accepted „adaptive system" for the assessment of the patient-relevant additional benefit (Health Technology Assessment, HTA). The assessment of the additional benefit by the Federal Joint Committee (G-BA) is the result of expert work based on a law (AMNOG) and procedural and methodical regulations.

The active players on the side of the G-BA and the health insurance funds are classified as scientists, hospital physicians and office-based statutory health insurance physicians, the Medical Service of the Health Funds and employees of the insurance fund administration, but also as patient representatives, however, they act on the basis of their own interests. Value dossiers for new pharmaceuticals, likewise qualified and interest-based, are submitted to the G-BA by the pharmaceutical companies, which serve as the basis for the assessment of the additional benefit.

Because the supply of pharmaceuticals to the population is significantly influenced by the assessment of the additional benefit, it makes sense to provide critical and careful support for the assessment process with a focus on identifying possible faults and counteracting imbalances. The Interdisciplinary Platform on Benefit Assessment set itself the task of supporting the benefit assessment within a small group of experts with the following objectives:

- Discussing the procedures for the assessment of the additional benefit, including in relation to approval of pharmaceuticals,
- Working towards international standards of evidence-based medicine and of health economy being adhered to as well as applied and further developed,
- Determining whether and to what extent patient-relevant additional benefits, in particular in the areas of mortality, morbidity and quality of life, are identified and which methodological problems occur during the process,
- dentifying possible undesirable developments, in particular with regard to supplying patients with new active substances,
- Enabling and holding a constructive dialogue with all players involved in the benefit assessment procedure, e. g. on the further development of the legal framework conditions of AMNOG.

Moreover, the European perspective in HTA of innovative pharmaceuticals was reinforced by the European Commission's proposal for a Regulation on HTA in 2018. Monitoring the conflict between the well-established national assessment and the intended European HTA harmonisation is also a central concern of the platform. The Interdisciplinary Platform would like to make a contribution to ensuring that new active substances are transparently and fairly assessed. According to the Advisory Council, an interdisciplinary dialogue about the results of the assessment and the applied benefit assessment methods is essential. Furthermore, in the benefit assessment process it sees a good opportunity to inform the prescribing physicians of the expected additional benefits of new pharmaceuticals for patients earlier than it was previously the case.

The Interdisciplinary Platform is a result of the discussion process between clinicians and experts. The mutual desire to pool specialist knowledge in the form of interdisciplinary seminars is supported by an open consortium of sponsors. These include AbbVie Deutschland GmbH & Co. KG, DAK Gesundheit, MSD Sharp & Dohme GmbH, Novo Nordisk Pharma GmbH, Roche Pharma AG, Association of Research-Based Pharmaceutical Companies (vfa e.V.), and Xcenda GmbH.

The Advisory Council of the Interdisciplinary Platform on Benefit Assessment

Contextual evidence: From mega trial to more targeted treatment methods

By Professor Jörg Ruof

The precision of medicinal treatment interventions has improved considerably due to pathophysiological knowledge advances, availability of various biomarkers and new, partly genetic treatment options for many diseases. Although this trend leads to an increased complexity of medical care, there is no doubt that affected patients benefit from it.

However, the comparative assessment of the additional benefit according to Section 35a of the 5th German Social Codebook (Sozialgesetzbuch V, SGB V) is associated with various challenges, such as:

- The number of patients available for a specific clinical study decreases and only comprises the predefined subpopulation and the respective line of treatment;
- Comparative treatments are still partly based on a generic approval status and do not or only insufficiently reflect the current gain in pathophysiological knowledge, respectively;
- The evidence available at the time of approval might allow for approval-associated evaluation of the benefit risk profile, but the data will not or not yet, respectively, be sufficient for the quantification of an additional benefit.

The Ordinance on the Benefit Assessment of Pharmaceuticals (AM-NutzenV) basically takes account of this trend. The ordinance states (Section 5, Paragraph 3: „If it is not possible or inappropriate to conduct or request studies of the highest level of evidence, proof of the best available evidence level must be provided". The ordinance does not specify how to handle this best available evidence leaving it up to the G-BA or self-governing bodies, to decide about the additional benefit, its quantificability, or potential time-limitation of the decision.

The publication series of the Interdisciplinary Platform on Benefit Assessment presents a discussion paper from

different perspectives on the increasing contextualisation of the concept of evidence which is currently mainly based on the „gold standard" of randomised comparative trials (RCTs):

- Holger Schünemann presents the examples of two controversially discussed medicinal innovations, i. e. the GRADE methodology that was also developed in McMaster. It provides an „Evidence to Decision" (EtD) framework making value decisions that are always required in regulatory or HTA processes comprehensible and transparent.
- The report by the Paul-Ehrlich-Institut (Elena Wolff-Holz & Klaus Cichutek) first illustrates the key parameters of the German and European approval process with appropriate variants. Subsequently, various examples are presented in which – among other things – single arm studies (ceritinib, axicabtagene ciloleucel), intra-patient comparisons (emicizumab), and complex study designs (canakinumab) made up a significant proportion of regulatory decision-making.
- From the IQWiG's perspective, Thomas Kaiser describes procedures in which evidence from single arm studies was used for benefit assessment. Here, especially innovative procedures for the treatment of hepatitis C, but also from the field of oncology are mentioned. In conclusion, the use of single arm studies should be limited to exceptional cases during benefit assessment.
- The abstracts of Bernhard Wörmann (haematology and oncology) and Dirk Müller-Wieland (diabetology) illustrate the treatment-relevant perspective of clinicians. The quick identification of effective medicinal products, consideration of biomarkers and the selection of patient-oriented endpoints are the key challenges for future benefit assessment in oncology. Key factors in diabetology include availability and interpretation of cardiovascular safety studies – among others – at subgroup level as well as high-quality determination of the patient perspective.
- Subsequently, readers gain an insight into the present state of discussion and current European legislative process from Giovanni Tafuri, the Scientific Director of EUnetHTA. The focus here is on the development of a sustainable model for the cooperation of the 83 national and regional European HTA institutions.
- Michael Hennrich then presents the „crux of the matter" from the political perspective. Thus, the revised version of Section 5, Paragraph 3b authorises the G-BA to request post-marketing data collection for certain pharmaceuticals. Previous experiences with orphan diseases and pharmaceuticals for advanced therapies in the Medicinal Products Act (AMG) as well as experiences with cancer registries are outlined, e.g. the German RABBIT registry (Register for the long-term observation of therapy with biologics in adult patients with rheumatoid arthritis) or CRISP Platform (Clinical Research Platform Into Molecular Testing, Treatment and Outcome of Non-Small Cell Lung Carcinoma Patients) for lung cancer. The discussion also raises the question of how future data collection can be better coordinated at European level. At present, Michael Hennrich only sees further need for political adaptation at the level of prescription of pharmaceuticals.

Readers should especially take note of the variety of examples in all articles on the basis of which a decision can be taken whether the current legal and methodological frameworks are appropriate or require adaptation.

Contact:
joerg.ruof@r-connect.org

Contextual evidence: Which evidence is required for which research question?

By Holger J. Schünemann, MD, PhD Professor, Departments of Health Research Methods, Evidence and Impact and of Medicine Director, McGRADE and Michael G. DeGroote Cochrane Canada Centres, McMaster University Health Sciences Centre

To achieve this goal in the context of regulatory, technology assessment and guideline decisions, structured decision frameworks should be utilized that are interchangeable, yet adjustable to the type of decision-making. Decisions depend on the perspective that is taken, e.g. that of the individual, the population, or the health system. Using structured processes that can be understood and contextualized, enhances transparency and create efficiencies on a micro and macro level.

The organizers of the conference suggested one example focusing on treatment of hemophilia A and I utilized a second one focusing on multi-drug resistant tuberculosis (MDR-TB) to introduce the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Evidence to Decision (EtD) frameworks. It is evident that judgments are required to inform decisions about the criteria in the EtDs. The most stringent methodological approaches do not make these judgements disappear, but they can make them transparent.

The two examples focused on two relatively new drug technologies, the use of emicizumab in patients with hemophilia A and inhibitors and the use of bedaquiline in MDR-TB. Based on existing health technology assessments of the German Gemeinsamer Bundesausschuss and the Institute and World Health Organization guidelines, I describe which evidence is required for which criterion in the EtD. For the two examples described here, despite concerns about the evidence that demonstrates the effects of the intervention on the outcomes, using the EtD, decision makers will be confident that they, at the time of the decision, make that decision with confidence.

Introduction

Decision-makers in healthcare should strive do no net harm[1]. To achieve this goal in the context of regulatory, technology assessment and guideline decisions, structured decision frameworks should be utilized that are interchangeable, yet adjustable to the type of decision-making. Decisions depend on the perspective that is taken, e.g. that of the individual, the population, or the health system. Using structured processes that can be understood and contextualized, enhances transparency and create efficiencies on a micro and macro level.

Many criteria have been proposed to create comprehensive frameworks. GRADE Evidence to Decision (EtD) frameworks are increasingly utilized to accomplish the goals of comprehensiveness, transparency and usability in different context and jurisdictions[2-6]. The EtD consist of four sections: the question and background, the assessment, the conclusions and the presentation (Figure 1).

The question and background section include the detailed description of the population, intervention, comparison and outcome (PICO) and possible subgroups as well as eventual conflicts of interest of the group asking and answering the question. The assessment describes the criteria that drive the decision to be made. Groups or organizations assess up to 11 criteria (the criteria are flexibly chosen on the basis of the decision-maker's needs or requirements) that include a section on research evidence, ideally from systematic reviews or health technology assessment (HTA), additional considerations and the judgments that are made (Figure 2).

The conclusion section begins with a decision or recommendation that follows from reviewing, and if desired weighting, the disaggregated judgments on the criteria. Depending on the framework used and perspective taken these may be clinical guideline recommendations, cover-

Interactive Evidence-to-Decision approaches

Question
- Details – PICO subgroups
- Background and conflicts of interest

Assessment
- Criteria
- Judgements
- Research evidence (HTA and systematic reviews)
- Additional considerations

Conclusions
- Type of decision – recommendation
- Justification
- Implementation considerations – monitoring and evaluation
- Research considerations

Presentation
- Guideline group meetings & informing coverage decisions
- Database of decision framework
- Interactive Decision Aids (iDeAs), Apps

Perspectives
- Clinical – individual
- Clinical – population
- Health systems & Public health
- Health systems & Politics

Type of decisions
- Recommendation
- Policy
- Coverage

Use
- Group decision making
- In person / Online

Source: gradepro.org

Prof. Dr. Dr. Holger Schünemann *works as a professor in the departments of Health Research Methods, Evidence, and Impact – which he led from 2009 to 2019 – and Medicine of the McMaster University. After his studies in human medicine at Hanover Medical School (Ph.D. in 1994), he completed various postgraduate degrees in Pulmonary Biology, Epidemiology, Internal Medicine and Preventive Medicine / Public Health at the University of Buffalo, State of New York. He is Co-chair of the GRADE working group and Director of the Cochrane Canada and McMaster GRADE centres. Moreover, he operates a practice for internal medicine.*

Criteria, evidence base and additional considerations that impact strength and direction of the GRADE recommendation

1. Problem: Is the problem a priority?
2. Desirable effects: How substantial are the desirable anticipated effects?
3. Undesirable Effects: How substantial are the undesirable anticipated effects?
4. Certainty of evidence: What is the overall certainty of the evidence of effects?
5. Values: Is there important uncertainty about or variability in how much people value the main outcomes?
6. Balance of effects: Does the balance between desirable and undesirable effects favour the intervention or the comparison?
7. Resources required: How large are the resource requirements (costs)?
8. Certainty of evidence of required resources: What is the certainty of the evidence of resource requirements (costs)?
9. Cost effectiveness: Does the cost-effectiveness of the intervention favour the intervention or the comparison?
10. Equity: What would be the impact on health equity?
11. Acceptability: Is the intervention acceptable to key stakeholders?
12. Feasibility: Is the intervention feasible to implement?

Source: GRADE EtD framework

Figure 2. Criteria, research evidence, judgments and additional considerations that influence the strength and direction in the GRADE Evidence to Decision Frameworks.

age decisions or health system decisions. Finally, in GRADE's official tool and app GRADEpro allows for different types of presentation and dissemination approaches (Figure 1). To make decisions transparent, decision-makers should describe the perspective they take (choosing from an individual or population clinical, health system and public health, or health system and policy perspective) and the type of decision (recommendation, policy or coverage). While many of these concepts are used by decision-makers in Europe, including Germany, there still is a lack of comprehensive structures and collaboration to accomplish common goals and allow for exchange between organizations.

My goal for this presentation was - based on examples that created debate among decision-makers - to describe, what type of research is required to achieve confidence in the results of that research and how to achieve confidence in a decision, even if the research does not permit judgments of high confidence in intervention effects. Furthermore, I aimed at describing that the use of such decision frameworks can facilitate interagency exchange and use. To illustrate these concepts, in this article I discuss which

and how criteria in the EtD framework and the domains within those criteria influence that confidence.

2. Examples

The organizers of the conference suggested one example and I utilized a second one that we previously used to introduce the EtD framework[5]. The GRADE EtDs are a result of a European Commission supported 5-year project (the DECIDE project) and have since been used by numerous organizations to make decisions and recommendations in healthcare[7].

Hemophilia A

The example I was asked to address relates to hemophilia A. Hemophilia A is an inherited, serious bleeding disorder in which the affected person's blood does not clot properly because of congenital factor VIII deficiency. This, in turn, can lead to uncontrolled bleeding which occurs with minor trauma or even spontaneously. Hemophilia A can severely impact on the person's quality of life. It is a rare condition that affects approximately 1 in 5,000 individuals and about 1 in 10,000 are affected severely. In hemophilia, the term inhibitor refers to an autoantibody that typically forms in response to infused factor VIII. Inhibitors are most common in individuals with very low baseline factor VIII levels. Emicizumab is a recombinant, humanized, bispecific, therapeutic monoclonal antibody designed to replace the hemostatic function of FVIIIa (in the human body by bridging activated factor IX (FIXa) and factor X (FX))[8, 9]. One key clinical PICO question that addresses if there is net health benefit to answer in this context would be:

Multi drug resistant tuberculosis

I used a second example to illustrate some of the key issues in this article. Tuberculosis (TB) is among the oldest diseases known to mankind, but it remains one of the top ten causes of death globally, as well as the leading infectious disease killer. Multi drug resistant tuberculosis (MDR) describes TB that is difficult to treat because of the organism's resistance to key antibiotics and it affects more than 600,000 people and kills more than 240,000 each year according to WHO estimates. Treatment of MDR typically involves a regimen of several antibiotics to achieve synergy that can achieve cure in patients. Bedaquiline is a fairly new antibiotic agent that the World Health Organization has recommended in the treatment of MDR[10]. Thus, a critical question is if in MDR-TB patients, should bedaquiline should be added to a background MDR-TB treatment regimen based on WHO-recommendations? The following represents a PICO health question that we addressed in a WHO panel making recommendations to evaluate if there is net health benefit exists:

Decisions are required and question do not disappear even in the context of evidence that leaves decision makers with uncertainty. I will describe what type of research is required to achieve confidence in the results of that research and how to achieve confidence in a decision, even if the research does not permit judgments of high confidence in intervention effects and other types of required research evidence. This results from following structured processes and understanding and acknowledging the gaps in research evidence.

Population: **In people with factor VIII deficiency** with inhibitors, what is the impact of
Intervention: **emicizumab** compared with
Comparison: **no emicizumab** on
Outcomes: **bleeding outcomes, adverse effects of treatment, quality of life**

3. Certainty in the evidence versus confidence in the recommendation or decision

What is certainty of evidence?

Certainty that an estimate of association or effect is correct or, better, that a true effect lies on one side of a specified threshold or within a chosen range[11].

What is confidence in a decision?

Confidence in the decision will arise when a decision-maker feels that the best available evidence – whether of high or very low certainty – has been compiled through systematic reviews or HTA and considered in the context of a transparent and comprehensive approach for making judgments about that evidence. Whatever that decision is, e.g. to cover or not cover an intervention or to make a conditional recommendation for an intervention in the face of uncertainty of the evidence about the health effects, a decision-makers may find their work easily trustworthy and defendable even if the research evidence has gaps or shortcomings.

In the hemophilia A example, the evidence came primarily from two randomized trials, one in patients with inhibitors[8], and two recent representative HTAs that I chose based on a pragmatic survey of the literature[12, 13]. There was no formal rating of the certainty of the body of evidence but an assessment of risk of bias only in the reports which I will return to below. I will describe a hypothetical scenario for a decision that could be made with confidence based on the evidence reviewed.

Population: **In multidrug-resistant tuberculosis (MDR-TB) patients, what is the impact of**
Intervention: **bedaquiline plus background MDR-TB treatment compared with the**
Comparison: **background MDR-TB treatment alone**
on
Outcomes: **cure by 120 weeks, adverse drug reactions (clinical and biological serious adverse events), mortality, time to culture conversion, culture conversion at 24 weeks, acquired resistance to fluoroquinolone and injectable drugs**

In the WHO MDR-TB guidelines the key data came from a single randomized trial with data from up to 160 participants. The evidence was judged to be at very low certainty (see section 4) for the estimates on cure, mortality and adverse effects as the main outcomes. Despite this uncertainty in the estimates on patient important outcomes, the panel could feel confident that it made the right decision by conditionally recommending the intervention and stipulating under which circumstances the intervention was a viable option. This difference between the (very low) certainty in the evidence and (high) confidence in the decision is a result of considering all aspects that inform the recommendation in an open and transparent way and then emphasizing that uncertainty about the effects of bedaquiline. Furthermore, acknowledging that further research might increase the certainty in the evidence and result in a different recommendation enhanced that confidence in the decision. I will explain the details and rationale below.

4. Criteria in an evidence to decision framework

Figure 2 lists the criteria in the GRADE EtD framework from which those conducting assessments and making decisions can choose.

4.1. Is the problem a priority?

The judgment if the problem is a priority is determined by the importance and frequency of the health care issue that

is addressed (burden of disease, prevalence, cost, baseline risk, policy issues and others).

Evidence required and examples

When prevalence and baseline risk are the drivers of importance the evidence required to address if a problem is a priority usually comes from cohort studies. For guideline development groups, surveys asking if there is variation in practice, new important evidence, unclear existing guidance or lack of recommendations are often utilized to address the importance of a problem. While high prevalence of a condition or high number of people affected increases the likelihood that an intervention that addresses the problem should be a priority, for rare diseases other reasons for priorities may arise. For example, introduction or approval of new interventions of presumed important efficacy or high drug cost which address rare conditions for which there previously were few options create priorities for patients, clinicians and policy-makers. Similarly, mandates of drug agencies and regulators create priorities. Thus, the type of research evidence required here differs and narrative descriptions in the additional considerations section describing the priority will often be used. For the examples 1 and 2, new evidence emerging for interventions with high potential treatment success for a rare disease and the availability and approval of a new antibiotic for a deadly condition based on observational studies and trials were reasons for priorities, respectively.

4.2. and 4.3. How substantial are the intervention effects on the desirable and undesirable health effects (benefits and harms and burden)?

This requires an evaluation of the magnitude of the absolute effects of the intervention on both the desirable and undesirable health effects. This requires consideration of all critical and sometimes the important outcomes for a decision and determining if these effects are trivial, small, moderate or large.

Evidence required and examples

For example 1, we require intervention studies, ideally properly conducted randomized controlled trials comparing emicizumab with no emicizumab that evaluate the impact on the outcomes of interest: bleeding rates, adverse effects and quality of life. The observed effects were reported as 87% relative reduction in bleeding episodes, an important improvement in quality of life and an increase in adverse effects[8, 12].

For example 2, randomized trials that add bedaquiline to an existing WHO regimen would be the type of studies that directly permit an estimate of the absolute estimates of effect in representative patients on all outcomes of interest. The results suggested a large effect (over 20%) on cure rates that was balanced (in the 2013 guidelines) by a potentially large mortality increase (up to 10%) and adverse effects.

4.4. Certainty of the evidence about the desirable and undesirable health effects

The certainty of the evidence can be evaluated for each of the criteria that include research evidence in an EtD but plays an eminent role in the evaluation of the effects of the intervention on the health effects. The eight GRADE certainty domains are assessed individually and then in relation to each other to obtain an overall view of the certainty of the evidence, ideally based on systematic reviews related to the PICO question. The eight domains are: risk of bias, inconsistency, indirectness, imprecision, publication bias, strong association, dose-response relation, plausible residual confounding.

4.4.1. Assessing Risk of Bias (Limitations in study design and execution/conduct) across studies

Both randomized controlled trials (RCTs) and observational (non-randomized) studies may have additional risk of misleading results if they are flawed in their design or conduct. The signaling question for this domain is „Are the research studies well done?". This is often referred to problems with ‚validity" or ‚internal validity" which GRADE labels as ‚study limitations" or ‚risk of bias." Various tools are used to assess this risk of bias in randomized controlled trials and observational studies but most use similar items, e.g. bias due to confounding, bias in selection of participants into the study, bias in classification of interventions, bias due to deviations from intended interventions, bias due to missing data, bias in measurement of outcomes, and bias in selection of the reported result.[14] GRADE suggests doing this by study and outcome and then across studies by outcome for an overall rating of the risk of bias [15].

4.4.2. Assessing inconsistency of the results

The signaling question for this domain is „Are the results consistent across studies when they should be consistent because they address similar PICO questions?". Widely differing estimates of the treatment effect (i.e. heterogeneity or variability in results) across studies suggest true differences in underlying treatment effect. When heterogeneity exists, but investigators fail to identify a plausible explanation, the certainty of evidence should be rated down by one or two levels, depending on the magnitude of the inconsistency in the results. Inconsistency may arise from differences in populations (e.g. drugs may have larger relative effects in sicker populations), interventions & comparators (e.g. larger effects with higher drug doses, slightly altered interventions or different comparators), or outcomes (e.g. diminishing treatment effect with time).

4.4.3. Assessing indirectness of the results

The signaling question for this domain is „How directly do the results relate to our PICO question?". Evidence is always somewhat indirect. Indirectness describes how directly the identified evidence relates to the research question. There are two types of related indirectness (also called generalizability, transferability, external validity, relevance, applicability, translatability). First, evidence may come from an indirect population, intervention, comparator, or outcome. Second, indirect comparison which occurs when a comparisons of intervention A versus B is not available, but studies compared A with C and B with C. Such studies allow indirect comparisons of the magnitude of effect of A versus B. This evidence is of lower certainty than head-to-head comparisons of A and B would provide.

4.4.4. Assessing imprecision of the results

The signaling question for this domain is „Is the effect size precise - due to random error?". Imprecision describes the degree of random error that may influence the assessment of the results and interpretation. Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect or association. Guiding principles for calling results imprecise exist. When considering the certainty of evidence, the issue is whether the confidence interval around the estimate of effect or association is sufficiently narrow.

4.4.5. Assessing risk of publication bias

The signaling question is „Are these all of the studies that have been conducted for this outcome?". Publication bias is a systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to the selective publication of studies (publication bias). That is, investiga-

tors fail to report studies they have undertaken (typically those that show no effect) or journals are less likely to accept studies that show no effect for publication. This often occurs if there is for-profit interest and only small studies exist that show positive effects.

4.4.6. Assessing if there is a strong association

The signaling question is „Is the effect large or very large and (relatively) unbiased?" When methodologically well-done observational studies (at low risk of bias on the relevant tool) yield large or very large and consistent estimates of the magnitude of a treatment or exposure effect, we may be confident about the results. In those situations, the weak study design is unlikely to explain all of the apparent benefit or harm, even though observational studies are likely to provide an overestimate of the true effect. The larger the magnitude of effect, the stronger becomes the evidence. GRADE suggests relative risks (RR) of >2 or <0.5 with no plausible important confounding and very large RR >5 or <0.2 with no plausible important confounding as thresholds for evaluation.

4.4.7. Assessing dose-response relation

The signaling question is „Are there relations between exposure dose and effects that make us more confident?". The presence of a dose-response gradient may increase our confidence in the findings of observational studies and thereby increase the certainty of evidence. Only studies with no threats to validity (not downgraded for any reason) should make us more confident.

4.4.8. Assess effects of plausible residual confounding

The signaling question for this domain is „Do the results despite worst case scenario predictors still allow strong conclusions about the effect?" On occasion, all plausible confounding from observational studies may be working to reduce the demonstrated effect or increase the effect if no effect was observed. For example, if only sicker patients receive an experimental intervention or exposure, yet they still fare better, it is likely that the actual intervention or exposure effect may be larger than the data suggest.

4.4.9. Overall certainty of the evidence

The certainty of the evidence is then initially categorized by outcome into high, moderate, low and very low after evaluating the evidence on all domains based on rating down the certainty within and combined for domains by one to three levels depending on the instruments used[15, 16]. For decisions, the lowest certainty of the evidence of all critical outcomes, whether benefit or harm, is then used to express an overall certainty of the evidence for the decision.

4.4.10 Evidence required and examples

For both examples, availability of all randomized controlled trials with low risk of bias that are consistent, precise and direct would provide high certainty of the evidence in treatment effects. For the hemophilia A example, in the two HTA reports there was concern about risk of bias and precision of the estimates. This would leave us with low or very certainty of the evidence given the small number of patients enrolled. In rare circumstances, we have high confidence in treatment effects if they are very large and precise. This can occur in situations where prior to the intervention all patients traditionally suffered from severe consequences including death and after the introduction of the intervention such outcomes are absent or nearly absent. The effect of emicizumab on bleeding outcomes at this point does not fulfill the guiding principle for high certainty based on large effects because the effect is based on a relatively small number of patients and events. However,

properly conducted non-randomized studies, if precise and free of other important limitations that would show effect sizes similar to those observed in the small randomized trial may increase our certainty in the future – at least for the outcome bleeding rate.

For the MDR-TB example, when the WHO panel assessed the evidence in 2013, the certainty of the evidence was rate at low and very low for most outcomes - primarily because of concerns about imprecision and indirectness.

4.5. Values and preferences or the importance of outcomes

This criterion describes how important health outcomes are to those affected, how variable they are and if there is uncertainty about this. This relative importance of outcomes is equivalent to values and preferences that people assign to health outcomes. High variability of which value those affected assign to the outcomes or more uncertainty about how important the outcomes are led to more uncertainty in the evidence.

Evidence required and examples

The type of evidence required would be studies that directly evaluate the decisions about use of emicizumab and bedaquiline in the respective patient populations based on informed decisions and an explanation of the expected outcomes[17,18]. These types of informed choice studies provide information about underlying values and preferences. Alternatively, observational studies in these patient populations that directly elicit the relative importance of the outcomes could provide high certainty of the evidence. Such studies exist but more evidence in this field is usually required [19, 20] and was explicitly used in the ICER report[13]. In situations where this evidence is not available, multiple strategies for obtaining the evidence or using decision panels' input in additional considerations can provide transparency although rarely would lead to high certainty of the evidence. In the WHO MDR-TB guideline on bedaquiline the guideline group „felt that there were potentially large variations in patient values and preferences for each outcome. Most members felt that patients would place high value on survival but that it was less clear that patients would value microbiological culture conversion in the same way."

4.6. Balance of the desirable and undesirable health effects

The greater the net benefit or net harm the more likely is a strong recommendation for or against the option or intervention. This evaluation requires balancing the desirable and undesirable health effects considering the sum of them, the variability of how important they are and the certainty of the evidence about the intervention's health effect and the relative importance of the outcomes.

Evidence required and examples

The evidence for this criterion stems from evaluating the health benefits and harms, the values and the certainty of the evidence. Thus, the balance of the benefits and harms is a combination of the numerical (absolute) effect measures and the corresponding relevant importance of the outcomes. Although rarely done the overall certainty of the evidence should, thus, also be a result of certainty of evidence assessment for the intervention effects and the underlying values and preferences. Given the certainty described above, for both examples the certainty would be low to very low for both examples as for some of the critical outcomes (bleeding and rare adverse effects in example 1 and mortality and cure in example 2), the evidence was rated at low to very low, respectively.

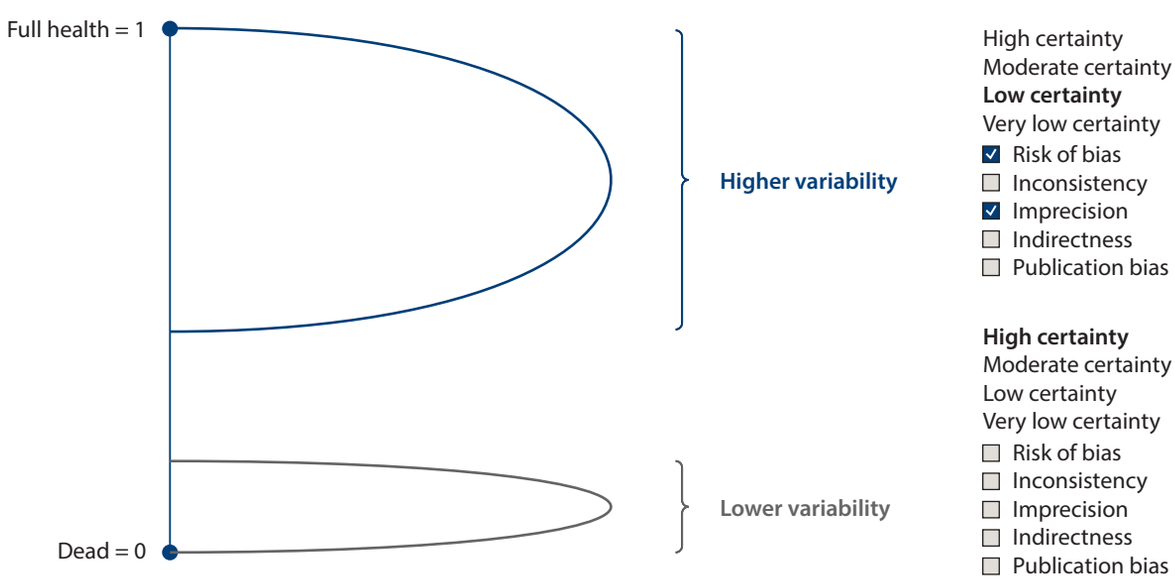
4.7. – 4.9. Resource implications

This describes how resource intense an option is, if it is cost-effective and if there is incremental benefit by looking at cost-effectiveness. The more advantageous or clearly disadvantageous these resource implications are the more likely is a strong recommendation.

Evidence required and examples

For the three criteria, first observational studies or, better, trials that measure cost in the actual intervention studies would be required[21-23]. The certainty of the evidence assessment follows general GRADE principles that also apply to resource use and cost[22, 23]. For cost-effectiveness studies, appropriate, credible and complete models should be developed that use the effect estimates of the systematic reviews used in the EtDs, the corresponding estimates of the relative importance of the outcomes. The HTA report by Institute for Clinical and Economic Review (ICER)[13] conduced cost-effectiveness analyses. The overall certainty in that model would usually not be higher than that for its input variables derived from the evidence about the effect estimates and relative importance of the outcomes. However, the conclusions by ICER were that the „results from our cost-effectiveness analysis show emicizumab to be a dominant strategy". For the bedaquiline example, a cost-effectiveness was conducted. The guideline group concluded that „while the cost-effectiveness modelling showed overall benefit, there were concerns about the simplifying as-

Variability and certainty of evidence about the relative significance of endpoints



Source: Professor Holger Schünemann

Figure 3. Variability and certainty in the evidence about the relative importance of outcomes.

sumptions used (e.g. no accounting for the difference in serious adverse events, no accounting for effect on transmission, uncertainty about application of trial outcomes – including deaths – to routine programmatic conditions, etc.). The guideline development group also felt that cost effectiveness would not necessarily translate into affordability or country readiness to pay given the potentially high cost of bedaquiline."

4.10. Equity

Interventions should similarly benefit all populations who are the target of a decision or recommendation. Inequities can result if the effects differ for population groups, in particular those who are disadvantaged or cause disadvantages in some groups of the population. The greater the likelihood to reduce inequities or increase equity and the more accessible an option is, the more likely is a strong recommendation.

Evidence required and examples

GRADE suggests that if health inequity is determined to be a concern by stakeholders, there may be five methods for explicitly assessing impact of interventions on health equity:[1] include health equity as an outcome;[2] consider patient-important outcomes relevant to health equity;[3] assess differences in the relative effect size of the treatment; [4] assess differences in baseline risk and the differing impacts on absolute effects; and[5] assess indirectness of evidence to disadvantaged populations and/or settings[24]. Often this type of evidence can come from observational studies evaluating the importance of outcomes or randomized and observational studies evaluating effect modification. The impact of the interventions on equity in both examples is unclear. While the relevant health outcomes appear addressed in both scenarios, some questions remain. For example, rare disease may both facilitate access for all patients because of smaller number of patients to be treated but could also make access more challenging if only patients with access to specialized centers or those with coverage will receive it. The WHO guideline group felt „that effects on equity of bedaquiline addition to WHO-recommended MDR-TB treatment was difficult to assess, due to the uncertainly of affordability and country willingness to pay, as well as the difference in opinion on the balance of benefits and harms discussed above". This could mean that not all patients receive the treatment if recommended[25].

4.11. Acceptability

Acceptability deals with if there are key stakeholders that would likely not accept the distribution of the benefits and harms, cost, the values assigned. Other considerations affecting acceptability are if the intervention would adversely affect people's autonomy or if key would disapprove of the intervention morally for reasons (such as in regard to ethical principles such as no maleficence, beneficence, or justice). The greater the acceptability of an option to all or most stakeholders, the more likely is a strong recommendation.

Evidence required and examples

The type of evidence required are intervention studies measuring acceptability or observational studies of the interventions' use in practice. Surveys of relevant stakeholders (patients, providers, policy makers and others) may be informative and provide research evidence that can be used. Studies of patient adherence may provide indirect evidence that interventions are acceptable. The GRADE domains for certainty of the evidence should apply to this criterion but GRADE has not yet provided explicit guidance for its assessment. Evidence is being obtained for emicizu-

mab currently through surveys. While the WHO guideline group did not consider this criterion explicitly in the 2013 guideline, it had obtained some evidence in its 2017 update stating „the expert group felt that effects on equity of bedaquiline addition to WHO-recommended MDR-TB treatment was difficult to assess, due to the uncertainly of affordability and country willingness to pay, as well as the difference in opinion on the balance of benefits and harms discussed above". They described evidence from Vietnam and Belarus[26].

4.12. Feasibility

Feasibility relates to if the intervention or option is sustainable or if there are important barriers that are likely to limit the feasibility of implementing the intervention (or option) or require consideration when implementing it. The greater the acceptability of an option to all or most stakeholders, the more likely is a strong recommendation.

Evidence required and examples

Similar to the evidence about acceptability, evidence to assess feasibility is can be derived from observational studies and surveys of the interventions use. Additional considerations include logical conclusions (e.g., if a new intervention is not approved, it may not be feasible to use). For many drug interventions feasibility considerations may be derived from indirect evidence of similar interventions or these logical considerations. This evidence will have to be obtained but it is likely that emicizumab is feasible to use. The WHO guideline group evaluating bedaquiline felt in its update of the 2013 guidelines, based on observations and reports in national tuberculosis programs that reported feasibility, that it was probably feasible to implement the use of bedaquiline but the group was not unanimous in its judgment[26]

5. Resolution of the examples

Example 1 – hemophilia A

I reviewed two HTA reports that described two key studies evaluating emicizumab for hemophilia A in patients with inhibitors[12, 13, 27]. The conclusions differed. The Gemeinsamer Bundesausschuss (GBA) in Germany concluded that there is a suggestion for a non-quantifiable additional net benefit of emicizumab for patients in whom alternatives are not available. ICER concluded that „the findings of our analysis suggest that emicizumab prophylaxis provides gains in quality-adjusted life years at substantially lower costs over a lifetime horizon, with these findings remaining robust across multiple sensitivity and scenario analyses"[13].

A guideline's panel recommendation might be a conditional recommendation for the use of emicizumab in hemophilia A patients with inhibitors based on very low to low certainty of the evidence of the effects. This would primarily be a result of the lack of higher certainty evidence, acceptability to some stakeholders and cost in various settings.

Example 2 – multi-drug resistant tuberculosis

The WHO guideline group made a conditional recommendation for the use of bedaquiline in 2013 and 2017[25, 26]. The primary reasons were initial concerns about adverse effects in the randomized trials which were somewhat reduced based on an individual patient data meta-analysis in the 2017 update[28]. Cost and feasibility where other reasons that were detailed in the guideline reports. For an interactive EtD framework for the bedaquiline example see here: http://bit.ly/2GPgze7 (see tab with interactive summary of findings). Presentations similar to those for the bedaquiline example would facilitate understanding of the information in HTA reports, in particular for the summary of findings ta-

ble that some journals use to present findings of systematic reviews[29, 30].

6. Implications for practice and decision-making

Judgments are required to inform decisions about the criteria in the EtDs. The most stringent methodological approaches do not make these judgements disappear, but they can make them transparent. For the two examples described here, despite concerns about the evidence that demonstrates the effects of the intervention on the outcomes, using the EtD decision makers would or can be confident that they, at the time of the decision, make the decision with confidence. Given the gaps in research that would be described in the EtD framework, it is clear that decisions may change but not the confidence in them which can be high if a decision maker feels all relevant research evidence, additional considerations and judgments have been transparently used and described although the type of decision (coverage or strong instead of conditional recommendation) may be different after review of new or additional evidence.

7. Conclusions

Recipients of health care interventions deserve a transparent approach to making decisions that result from appropriate health care questions. Using transparent frameworks allows being confident in the decision. The criteria in the EtDs require different types of evidence to answer the relevant question including observational evidence and evidence obtained through surveys and priority setting exercises. For effects of interventions, randomized controlled trials are the evidence that provides the highest certainty although exceptions exist, for example when well designed and executed observational studies or randomized trials with minor flaws show large or very large effects that are credible. But this evidence was not (yet) present here. The concern is that none of the hemophilia reports, while comprehensive and transparent, clearly described the decision using easily comprehensible EtD frameworks or tables such as the one used in the MDR-TB example and multiple new guideline recommendations such as the ones by the European Commission Breast Guidelines or the American Society of Hematology[31, 32]. Such frameworks can help understand differences in judgments and resolve or discus disagreements between organizations that make decisions.

References

[1] Schunemann HJ. Guidelines 2.0: do no net harm-the future of practice guideline development in asthma and other diseases. Current allergy and asthma reports. 2011;11(3):261-8.

[2] Moberg J, Oxman AD, Rosenbaum S, Schunemann HJ, Guyatt G, Flottorp S, et al. The GRADE Evidence to Decision (EtD) framework for health system and public health decisions. Health Res Policy Syst. 2018;16(1):45.

[3] Schunemann HJ, Wiercioch W, Brozek J, Etxeandia-Ikobaltzeta I, Mustafa RA, Manja V, et al. GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLOPMENT. J Clin Epidemiol. 2017;81:101-10.

[4] Schunemann HJ, Mustafa R, Brozek J, Santesso N, Alonso-Coello P, Guyatt G, et al. GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. J Clin Epidemiol. 2016;76:89-98.

[5] Alonso-Coello P, Schunemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ. 2016;353:i2016.

[6] Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ. 2016;353:i2089.

[7] Treweek S, Oxman AD, Alderson P, Bossuyt PM, Brandt L, Brozek J, et al. Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence (DECIDE): protocol and preliminary results. Implementation science : IS. 2013;8:6.

[8] Oldenburg J, Mahlangu JN, Kim B, Schmitt C, Callaghan MU, Young G, et al. Emicizumab Prophylaxis in Hemophilia A with Inhibitors. N Engl J Med. 2017;377(9):809-18.

[9] Mahlangu J, Oldenburg J, Paz-Priel I, Negrier C, Niggli M, Mancuso ME, et al. Emicizumab Prophylaxis in Patients Who Have Hemophilia A without Inhibitors. N Engl J Med. 2018;379(9):811-22.

[10] Organization WH. The use of bedaquiline in the treatment of multidrug-resistant tuberculosis. Interim policy guidance http://apps.who.int/iris/bitstream/10665/75146/1/9789241548441_eng.pdf . Accessed 8 Dec 2014. 2013.

[11] Hultcrantz M, Rind D, Akl EA, Treweek S, Mustafa RA, Iorio A, et al. The GRADE Working Group clarifies the construct of certainty of evidence. J Clin Epidemiol. 2017;87:4-13.

[12] GBA. Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Emicizumab 2018 [Available from: https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/355/.

[13] Institute-for-Clinical-and-Economic-Review-(ICER). Emicizumab for Hemophilia A with Inhibitors: Effectiveness and Value 2018 [Available from: https://icer-review.org/material/hemophilia-a-final-evidence-report/.

[14] Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919.

[15] Schunemann HJ, Cuello C, Akl EA, Mustafa RA, Meerpohl JJ, Thayer K, et al. GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. J Clin Epidemiol. 2019;111:105-14.

[16] Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-6.

[17] Zhang Y, Coello PA, Guyatt GH, Yepes-Nunez JJ, Akl EA, Hazlewood G, et al. GRADE guidelines: 20. Assessing the certainty of evidence in the importance of outcomes or values and preferences-inconsistency, imprecision, and other domains. J Clin Epidemiol. 2019;111:83-93.

[18] Zhang Y, Alonso-Coello P, Guyatt GH, Yepes-Nunez JJ, Akl EA, Hazlewood G, et al. GRADE Guidelines: 19. Assessing the certainty of evidence in the importance of outcomes or values and preferences-Risk of bias and indirectness. J Clin Epidemiol. 2019;111:94-104.

[19] Soucie JM, Grosse SD, Siddiqi AE, Byams V, Thierry J, Zack MM, et al. The effects of joint disease, inhibitors and other complications on health-related quality of life among males with severe haemophilia A in the United States. Haemophilia. 2017;23(4):e287-e93.

[20] Naraine VS, Risebrough NA, Oh P, Blanchette VS, Lee S, Stain AM, et al. Health-related quality-of-life treatments for severe haemophilia: utility measurements using the Standard Gamble technique. Haemophilia. 2002;8(2):112-20.

[21] Parmelli E, Amato L, Oxman AD, Alonso-Coello P, Brunetti M, Moberg J, et al. GRADE EVIDENCE TO DECISION (EtD) FRAMEWORK FOR COVERAGE DECISIONS. Int J Technol Assess Health Care. 2017:1-7.

[22] Brunetti M, Shemilt I, Pregno S, Vale L, Oxman AD, Lord J, et al. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. J Clin Epidemiol. 2013;66(2):140-50.

[23] Guyatt GH, Oxman AD, Kunz R, Jaeschke R, Helfand M, Liberati A, et al. Incorporating considerations of resources use into grading recommendations. BMJ. 2008;336(7654):1170-3.

[24] Welch VA, Akl EA, Pottie K, Ansari MT, Briel M, Christensen R, et al. GRADE equity guidelines 3: health equity considerations in rating the certainty of synthesized evidence. J Clin Epidemiol. 2017.

[25] World-Health-Organization. The use of bedaquiline in the treatment of multidrug-resistant tuberculosis. Interim policy guidance http://apps.who.int/iris/bitstream/10665/75146/1/9789241548441_eng.pdf . Accessed 8 Dec 2014. 2013.

[26] World-Health-Organization. Report of the Guideline Development Group Meeting on the use of bedaquiline in the treatment of multidrug-resistant tuberculosis . Report of the Guideline Development Group Meeting on the use of bedaquiline in the treatment of multidrug-resistant tuberculosis, A review of available evidence (2016). Geneva: World Health Organization. Licence: CC BY-NC-SA 3.0 IGO Accessed July 15, 2017. 2017 [

[27] Gesundheitswesen IfrQuWi. Emicizumab (Hämophilie A) – Nutzenbewertung gemäß § 35a SGB V [Available from: https://www.iqwig.de/en/projects-results/projects/drug-assessment/a18-20-emicizumab-haemophilia-a-benefit-assessment-according-to-35a-social-code-book-v.9393.html.

[28] Mbuagbaw L, Guglielmetti L, Hewison C, Bakare N, Bastard M, Caumes E, et al. Outcomes of Bedaquiline Treatment in Patients with Multidrug-Resistant Tuberculosis. Emerg Infect Dis. 2019;25(5):936-43.

[29] Schunemann HJ, Santesso N, Brozek JL. Interactive Summary of Findings tables: the way to present and understand results of systematic reviews. JBI Database System Rev Implement Rep. 2019;17(3):259-60.

[30] Chu DK, Wood RA, French S, Fiocchi A, Jordana M, Waserman S, et al. Oral immunotherapy for peanut allergy (PACE): a systematic review and meta-analysis of efficacy and safety. Lancet. 2019;393(10187):2222-32.

[31] Schunemann HJ, Lerda D, Dimitrova N, Alonso-Coello P, Grawingholt A, Quinn C, et al. Methods for Development of the European Commission Initiative on Breast Cancer Guidelines: Recommendations in the Era of Guideline Transparency. Ann Intern Med. 2019.

[32] Schunemann HJ, Cushman M, Burnett AE, Kahn SR, Beyer-Westendorf J, Spencer FA, et al. American Society of Hematology 2018 guidelines for management of venous thromboembolism: prophylaxis for hospitalized and nonhospitalized medical patients. Blood Adv. 2018;2(22):3198-225.

Creation and evaluation of evidence on the example of haematology and oncology

Professor Bernhard Wörmann | Medical Director of the German Society for Haematology and Medical Oncology (DGHO) and Division of Haematology, Oncology, and Tumour Immunology at the Charité Universitätsmedizin Berlin

Evaluation of the efficacy and benefit of new pharmaceuticals is performed on the basis of evidence-based medicine. The wave of new pharmaceuticals for haematology and oncology can improve the patients' prognosis and quality of life, but also requires a critical assessment of the instruments used for the creation and evaluation of evidence. Particular challenges include the conduction of high-quality randomised clinical trials (RCTs) in consideration of all patient-relevant endpoints also for rare diseases (orphan diseases), waiving of RCTs in cases where convincing data from single arm studies are available, critical assessment of test methods including the application of new methods for the analysis of biomarkers, as well as securing the relation of study endpoints in early phase I/II studies to the long-term treatment goal.

Introduction

The early benefit assessment of new pharmaceuticals according to the German Pharmaceutical Market Reorganisation Act (AMNOG) has become an established procedure.[1] Scientific-medical professional associations accompany the process both from a methodological and scientific point of view and from the medical perspective.[2, 3] On the basis of the large number of procedures alone, haematology and oncology adapted a pioneering role in the evaluation and further development of the process of early benefit assessment solely (see figure 1).

Moreover, oncology currently plays a key role in medicine with regard to the translation of pathophysiological knowledge of disease patterns into diagnostics and therapy. Today, cancer is no longer just a disease, but the collective term for a variety of different diseases based on biological classifications. This has also an impact on the patients' treatment as well as the design of clinical studies.

In the following section, the assessment of evidence of new pharmaceuticals will be summarised on the basis of current results of early benefit assessment, followed by an overview of existing forms of evidence generation within the scope of clinical studies. Finally, an outlook for evidence generation and evaluation will be provided on the example of pharmaceuticals that are currently under development.

Assessment of evidence during early benefit assessment

Together with the German Society for Haematology and Medical Oncology (DGHO), the members of the commission „Benefit assessment of new pharmaceuticals" of the Association of the Scientific Medical Societies in Germany

(AWMF) analysed the results of early benefit assessments of new pharmaceuticals that have been conducted between 2011 and 2018.[4] One special feature of the German AMNOG procedure is the evaluation of subgroups the Federal Joint Committee (G-BA) can specify prior to the procedure. As these subgroups have an impact on both pricing and potentially on the economic prescription of pharmaceuticals, the evaluation of the determinations is performed by subgroups. Figure 2 shows a comparison of the overall results of all completed procedures (figure 2 A) with the specific results in oncology (figure 2 B) and haematology (figure 2 C).

By the end of 2018, the G-BA completed 344 procedures. Taking into account the formation of subgroups/subpopulations, the G-BA took 684 decisions. In 61 percent of all subgroups/subpopulations, the additional benefit was „not proven". 14.2 percent of all positive decisions on the additional benefit were rated as „low", 11.7 percent as „considerable", 0.4 percent as „significant", and 12 percent as „not quantifiable". In 0.7 percent of the subgroups/subpopulations, the damage was rated higher than the benefit and classified as category „lower additional benefit". The category „not quantifiable" mainly contains pharmaceuticals with an orphan drug status.

For oncology products, the proportion of decisions with the determination „additional benefit not proven" is slightly under 50 percent and the proportion of subgroups with the determination „considerable additional benefit" makes up almost 20 percent. Haematology shows a completely different pattern. The largest relative proportion are procedures with the determination „non quantifiable additional benefit". A comprehensive description of the major differences between the specialist fields is subject of a separate publication.[4]

Creation of evidence: Quality of approval studies

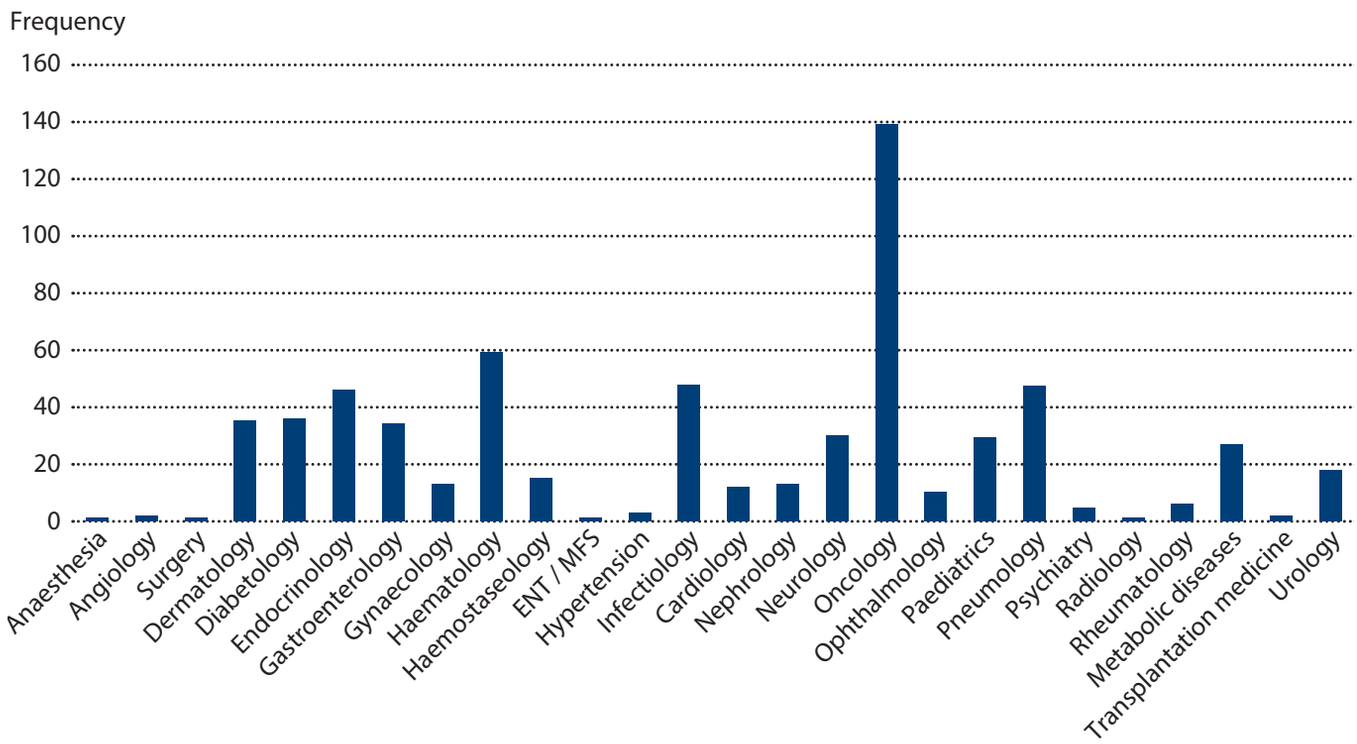
Approval study/ies and supplementary data are the basis of the early benefit assessment. Accepted forms of clinical studies include (in descending sequence of significance):[5]

- Meta-analyses
- > Two randomised clinical trials with concordant results
- One randomised clinical trial
- Non randomised intervention study
- Prospective observational study
- Retrospective observational study
- Case series / Case reports.

Randomised clinical trials (RCT) are the gold standard of evidence-based medicine, and it would be desirable to have more than one RCT on the same research question. Due to a heterogeneous patient collective and other unforeseen differences, studies with an identical design can produce different results. For financial and logistical reasons, studies with an identical design have become rather an exception than the rule in the testing of new pharmaceuticals in oncology.

Professor Bernhard Wörmann *works as a physician specialising in internal medicine, haematology, and internal oncology with additional qualification in palliative care. Since 2010, he has been Medical Director of the German Society for Haematology and Medical Oncology (DGHO). He works as a physician at the Virchow Campus of the Charité Universitätsmedizin Berlin, Germany.*

Representation of specialist fields in the early benefit assessment of new pharmaceuticals



Source: Professor Bernhard Wörmann

Figure 1: Haematology and oncology adopted a pioneering role in the evaluation and further development of the process of early benefit assessment solely due to the large number of procedures.

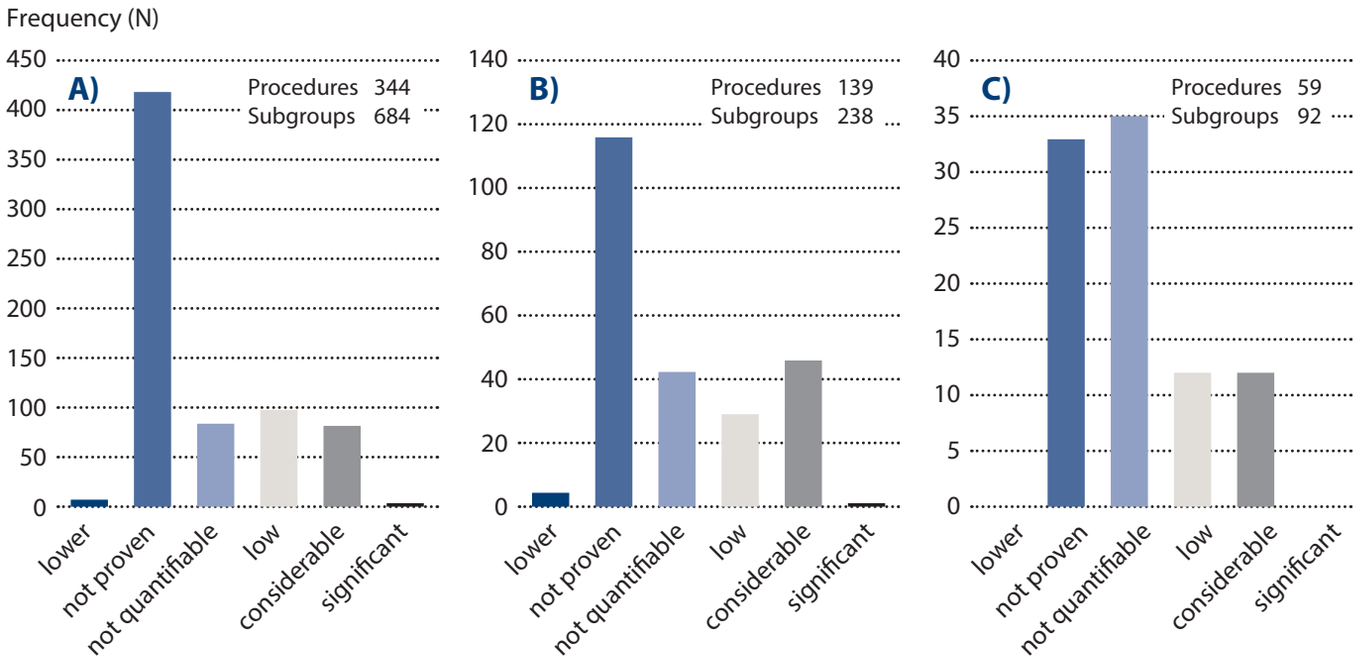
The proportion of RCTs varies significantly in haematology and oncology. While it makes up some 80 percent in oncology, it only constitutes 66 percent in haematology. The main cause is the higher rate of pharmaceuticals with an orphan drug status in haematology. Remarkable is, however, that pharmaceuticals with an orphan drug status also make up 32 percent in oncology. This shows that the conduction of randomised clinical trials is often even possible for pharmaceuticals for rare diseases.

Determination of the patient's quality of life

Approval study/ies and additional data are the basis of the early benefit assessment. Patient-relevant endpoints of pharmacotherapy are based on the four key substantive terms:

- Mortality / Lethality
- Morbidity
- Adverse events
- Quaity of Life

Determination of the additional benefit of new pharmaceuticals in all medical fields (A), in oncology (B) and haematology (C)



Source: Professor Bernhard Wörmann

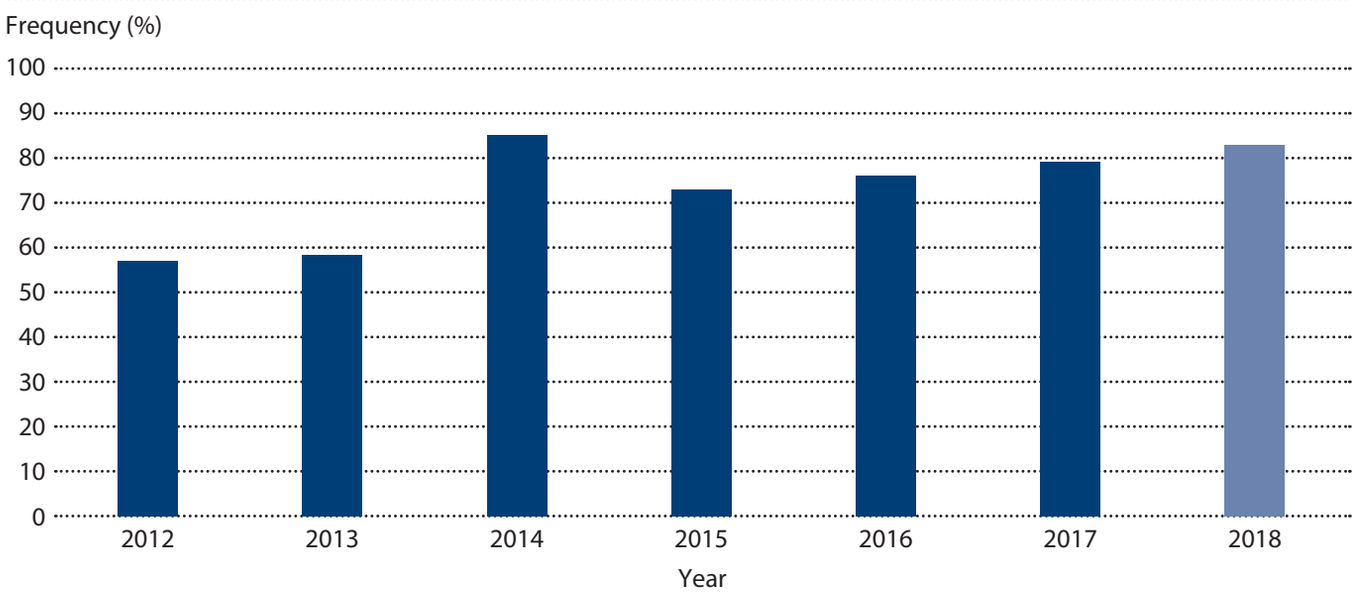
Figure 2: Taking account of all specialist fields, the additional benefit was „not proven" in 61 percent of all subgroups/ subpopulations. For oncology products, the proportion of decisions with the determination is below 50 percent.

Endpoints of mortality include overall survival or – if morbidity parameters are also included – disease-free or event-free survival. Endpoints of morbidity are very diverse. Besides a disease-free, event-free, and progression-free survival, it comprises remission and relapse rate, symptoms, avoidance of stressful symptoms, as well as imaging, laboratory, and cytological/histological parameters.

In the past, the patient's quality of life was not captured in clinical studies as a standard feature. This has changed in recent years. At the end of the 1960s of last century, the term „Quality of life" was adopted from the USA. Quality of life is highly subjective. In medicine, the assessment of the patient's quality of life reflects the attempt to make the subjective perception of health and disease measurable. In the meantime, some of the many regional and national questionnaires have gained acceptance, were validated and are now used worldwide in international studies.

Since the implementation of the AMNOG process, the proportion of procedures with standardised determination of the quality of life has increased and is approximately 80

Determination of the quality of life in oncological procedures



Source: Professor Bernhard Wörmann

Figure 3: Since the implementation of the AMNOG process, the proportion of procedures with standardised determination of the quality of life has increased and is approximately 80 percent in oncology.

percent in oncology (see figure 3). Unfortunately, the determination of the quality of life in non-randomised studies remains an exception.

Comparison of the results of phase-II and III studies

Even if phase-III studies are generally accepted as the gold standard for the determination of the (additional) benefit of new pharmaceuticals, the ethical question as to whether randomisation is legitimate arises at the individual patient level for highly effective new pharmaceuticals. Patients do not participate in studies primarily for altruistic reasons. They have high expectations on the new pharmaceutical hoping they will not be randomised to the placebo arm.

In the past years, there were three examples in oncology in which the results of an initial phase-II study leading to approval were compared with those of a confirmatory phase-III study. Table 1 shows a summary of the results.

In the studies on blinatumomab and osimertinib, the results of the phase-III study confirm the results of the phase-II-study, while results on olaratumab diverge. This might be attributable to the fact that in the study on blinatumomab in patients with relapsed/refractory acute lymphoblastic leukaemia (ALL) a clinically rather homogeneous patient collective was evaluated as was the case with osimertinib in patients with EGFR T790M + non-small cell lung cancer (NSCLC). By contrast, more than 20 different

histological subtypes were included in the studies on olaratumab for the treatment of advanced soft-tissue sarcoma. The question is whether confirmatory phase-III studies are feasible and necessary in homogeneous patient collectives.

Challenges: Quick identification of effective pharmaceuticals

The abundance of substances for the possible use in oncology and haematology is almost unmanageable. Figure 4 provides an overview of the pharmaceuticals that are currently available at the MD Anderson Cancer Center (Houston, Texas, USA) for clinical studies in patients with acute myeloid leukaemia (AML).[12]

Biomarker

The term „Biomarker" is a generic term for very different parameters used for the classification and stratification of patient groups, mainly in oncology. Biomarkers that are directly related to the mechanism of action of the pharmaceutical, e.g. kinase inhibitors, are particularly relevant. In recent years we have, however, experienced that an individual biomarker, even a so called driver mutation, has different effects within the context of different malignancies. Thus, remission rates of BRAF inhibitors such as dabrafenib, encorafenib and vemurafenib in patients with a BRAF V600E mutation range from <10 percent with colorectal carcinoma[15] and 50 percent with melanoma[16] up to >90 percent with hairy cell leukaemia (HZL).[17]

Moreover, previously accepted standards of DNA-based analyses by means of sequencing, FISH etc. might have to be complemented for some aberrations. Recently, it has been shown that e.g. ALK translocations usually lead to protein expression, although this is not the case in all patients. In these cases, identification of patients who cannot benefit from an ALK-targeted therapy is facilitated by means of RNA based ALK analysis.[18]

Particularly difficult are biomarkers that are not directly associated with the pathomechanism of the targeted therapy, but serve as surrogate parameters. These include e.g. for the immune checkpoint inhibitor analyses of PD-L1 or mutation load as tumour mutational burden (TMB).

Patient-oriented endpoints

The choice of the right endpoint presents a major challenge in the identification of new pharmaceuticals. Many targeted pharmaceuticals were first analysed in patients within the scope of phase-I/II studies alone or as part of basket trials with the rate of remission as short-term endpoint in many cases. In aggressive malignancies, such as acute leukaemia, an extension of the survival is only possible, if remission has been achieved, usually complete remission. In case of indolent malignancies, this correlation is less obvious. As a very short remission, even if it is confirmed histologically/cytologically and/or by medical imaging, only provides a questionable benefit, remission is not taken as primary endpoint in the vast majority of studies. This is also the case for the determinations in early benefit assessments of oncology products.

Moreover, endpoints such as improvement of the patient's quality of life are usually also not decisive in early phase I/II studies. Thus, clinical research on substances that might be of long-term benefit for patients might be terminated at an early stage.

Studies on new pharmaceuticals for acute myeloid leukaemia[12]

Newly Diagnosed
- Fludarabine + Ara-C + G-CSF + Gemtuzumab
- ATRA and arsenic +/- Gemtuzumab
- Nivolumab + Ida + Ara-C
- Cladribine + Ida + Ara-C
- FLAG-Ida + Venetoclax
- CPX-351 + Venetoclax
- BP1001 + LD Ara-C
- Pracinostat + AZA
- Chemo +/- Uproleselan
- Venetoclax + Cladribine + LDAC + LDAC w/ AZA
- Cladribine + LD Ara-C alternating DAC
- Nivolumab + AZA
- Dexrazoxane
- MCLA-117
- AML123
- DCLL9718S + AZA
- PCM-075 + LDAC
- PDR001/MBG453 + Decitabine
- SL401 + AZA
- Venetoclax + Decitabine
- Ulocuplumab + LD Ara-C

FLT3 Mutated Only
- ASP2215 vs ASP2215 + AZA vs AZA
- AC220 + AZA or LD Ara-C
- Quinzartinib + DAC

IDH Mutated only
FT-2102
AG120
AG221
Enasidenib + AZA

Frontline AML Post-Hypomethylating Therapy for MDS
- Nivolumab + AZA
- MCLA-117
- DCLL9718S + AZA
- BP1001 + LD Ara-C
- Ulocuplumab + LD Ara-C

Secondary Leukemia
- Ruxolitinib + DAC

Salvage
FLT3 Positive at relapse
- AlloSCT Initial Salvage Therapy
- AC220 + LD Ara-C
- Quinzartinib + DAC
- DS3032b + Quizartinib
- SKI-G-801
- Venetoclax + Gilteritinib
- CPX-351 + Venetoclax
- Palbociclib
- CD33-CAR-T
- Venetoclax + Decitabine
- Cladribine + Ida + Ara-C + Sorafenib
- SY-1425
- MCLA-117
- AMG 330
- XmAb 14045
- FLAG-Ida + Venetoclax
- AMG673
- DCLL9718S + AZA
- PDR001/MBG453 + Decitabine
- Daratumumab
- DS-3032b + LD Ara-C

RAS Mutated
- AlloSCT Initial Salvage Therapy
- FLAG-Ida + Venetoclax
- Venetoclax + Decitabine
- CD33-CAR-T
- SY-1425
- MCLA-117
- AMG 330
- AMG 673
- DCLL9718S + AZA
- PDR001/MBG453 + Decitabine
- Daratumumab
- CPX-351 + Venetoclax
- DS-3032b + LD Ara-C

AML with MLL gene at 11q23 Translocations
- SY-1425
- DS-3032b + LD Ara-C

IDH Mutated
- FT-2102
- AG120
- Enasidenib + AZA
- Venetoclax + Decitabine
- CD33-CAR-T
- SY-1425
- AMG 330
- XmAb 14045
- PDR001/MBG453 + Decitabine
- Daratumumab
- AG221
- DS-3032b + LD Ara-C

All others (Regardless of mutation status)
- Alvocidib (Flavopiridol)/Ara-C/Mitoxantrone (FLAM)
- CPX-351 + Venetoclax
- AlloSCT Initial Salvage Therapy
- CD33-CAR-T
- FLAG-Ida + Venetoclax
- Pilot Study CPX-351 w/ Gemtuzumab
- Cladribine + Ida + Ara-C + Sorafenib
- Idasanutlin + Ara-C vs Ara-C + Placebo
- SY-1425
- Nivolumab + 5-AZA
- LY2606368 + Ara-C + Flu
- Venetoclax + Decitabine
- MCLA-117
- DS-3032B
- AMG 330
- XmAb 14045
- AML123
- ADCT-301
- DCLL9718S + AZA
- OX40
- GO + Glasdegib
- AMG673
- Daratumumab
- IACS-010759
- SL401 + AZA
- MGD006
- PDR001/MBG453 + Decitabine
- DS-3032b + LD Ara-C

Quelle: [12]

Figure 4: The abundance of substances for use in oncology and haematology is almost unmanageable as illustrated by the example of acute myeloid leukaemia.

Comparison of the results of phase-II vs III studies

Study	Patients	Control	New therapy	N[1]	RR[2]	PFÜ/RFÜ[3] (HR)[4]	OS[5]
MT103-211 [6]	ALL[6], Ph-, refractoriness or relapse within 12 months	–	Blinatumomab	189	43	–	6.1
TOWER [7]	Ph-, refractoriness or relapse within 12 months, second or third relapse	Standard chemotherapy	Blinatumomab	405	16 vs. 34[7] p < 0.001	4.6 vs. 6.3	4.0 vs. 7.7 0.,71[8] p = 0.01
Jänne, 2015 [8] Dossier	NSCLC, EGFR T790M, after EGFR-TKI therapy	–	Osimertinib	400	66.1	9.7	85 %[10]
Mok, 2016 [9] Dossier	EGFR T790M, after EGFR-TKI therapy	Cisplatin/ carboplatin + pemetrexed	Osimertinib	419	31 vs. 71[7] p < 0.001	4.4 vs. 10.1 0.37[8] p < 0.001	n.b. vs n.b.[9] n. s. [11]
Tap, 2016 [10]	Soft-tissue sarcoma, no pretreatment with anthracyclines	Doxorubicin	Olaratumab	143	11.9 vs 18.2 n. s.	4.1 vs 6.6 0.67 p = 0.06	14.7 vs 26.5 0.46 p = 0.0003
ANNOUNCE, 2019 [11]	No pretreatment with anthracyclines	Doxorubicin	Olaratumab	460	–	6.8 vs 5.4 1.23 p = 0.042	19.7 vs 20.4 0.95 n. s.

[1] N – number of patients; [2] RR – rate of remission; [3] PFS – progression-free survival in months, RFS – relapse-free survival in months; [4] HR – hazard ratio; [5] OS – overall survival in months or %; [6] ALL – acute lymphoblastic leukaemia; EGFR T790M – mutation of the gene for the EGF-receptor, NSCLC – non-small cell lung cancer, Ph – Philadelphia chromosome; 7 rate after 24 weeks; [7] control outcome, new therapy outcome; [8] hazard ratio for new therapy; [9] n.c. – not calculable; [10] survival rate in % at 9 months; [11] n.s. – not significant

Source: Professor Bernhard Wörman

Table 1: In the studies on blinatumomab and osimertinib, the results of the phase-III study confirm the results of the phase-II-study, while results on olaratumab diverge.

References

[1] German Pharmaceutical Market Reorganisation Act (AMNOG), Federal Law Gazette Part I dated 27 December 2010. http://www.bgbl.de/Xaver/start.xav?startbk=Bundesanzeiger_BGBl&bk=Bundesanzeiger_BGBl&start=//*[@attr_id=%27bgbl110s2262.pdf%27]#__Bundesanzeiger_BGBl__%2F%2F*%5B%40attr_id%3D'bgbl110s2262.pdf'%5D__1374326855765

[2] Early benefit assessment of new pharmaceuticals in Germany – Subgroup definition, analysis and list of criteria, 2016. https://www.dgho.de/informationen/gesundheitspolitische-schriftenreihe/band-8-fruehe-nutzenbewertung-subgruppen/dgho_schriftenreihe_Bd8-2016_final.pdf

[3] Early benefit assessment of new pharmaceuticals in Germany 2011-2017, April 2018. https://www.awmf.org/fileadmin/user_upload/Service/Publikationen/AWMF_AMNOG_210x297_36S_f_web_ok.pdf

[4] Early benefit assessment of new pharmaceuticals in Germany 2011-2018, Justice and sustainability, May 2019. https://www.awmf.org/fileadmin/user_upload/Service/Publikationen/AWMF_AMNOG_210x297_36S_f_web_ok.pdf

[5] Benefit assessment of new pharmaceuticals in oncology and haematology. The status quo. https://www.dgho.de/publikationen/schriftenreihen/fruehe-nutzenbewertung-/dgho_gpsr_arzneimittelnutzenbewertung.pdf

[6] Topp MS, Gökbuget N, Stein AS et al.: Safety and activity of blinatumomab for adult patients with relapsed or refractory B-precursor acute lymphoblastic leukaemia: a multicentre, single-arm, phase 2 study. Lancet Oncol 16:57-66, 2015. DOI: 10.1016/S1470-2045(14)71170-2

[7] Kantarjian H, Stein A, Gökbuget N et al.: Blinatumomab versus Chemotherapy for Advanced Acute Lymphoblastic Leukemia. N Engl J Med 376:836-847, 2017. DOI: 10.1056/NEJMoa1609783

[8] Jänne PA, Yang JC, Kim DW et al.: AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. N Engl J Med 372:1689-1699, 2015. DOI: 10.1056/NEJMoa1411817

[9] Mok TS, Wu YL, Ahn MJ et al.: Osimertinib or Platinum-Pemetrexed in EGFR T790M-positive lung cancer. N Engl J Med 376:629-640, 2017. DOI: 10.1056/NEJMoa1612674

[10] Tap WD, Jones RL, van Tine BA et al.: Olaratumab and doxorubicin versus doxorubicin alone for treatment of soft-tissue sarcoma: an open-label phase 1b and randomised phase 2 trial. Lancet 388:488-497, 2016. DOI: 10.1016/S0140-6736(16)30587-6

[11] Olaratumumab for treatment of soft-tissue sarcoma – negative Phase-III-trial. https://www.dgho.de/aktuelles/news/news/2019/olaratumab-beim-weichgewebssarkom-negative-phase-iii-studie

[12] https://www.mdanderson.org/research/departments-labs-institutes/ departments-divisions/leukemia/clinical-trials.html

[13] Woodcock J, LaVange LM: Master Protocols to Study Multiple Therapies, Multiple Diseases, or Both. N Engl J Med 377:62-70, 2017. DOI: 10.1056/NEJMra1510062

[14] Cecchini M, Rubin EH, Blumenhthal GM et al.: Challenges with Novel Clinical Trial Designs: Master Protocols. Clin Cancer Res Jan 29, 2019. DOI: 10.1158/1078-0432.CCR-18-3544

[15] Kopetz S, Desai J, Chan E et al.: Phase II Pilot Study of Vemurafenib in Patients With Metastatic BRAF-Mutated Colorectal Cancer. J Clin Oncol 33:4032-4038, 2015. DOI: 10.1200/JCO.2015.63.2497

[16] Sullivan R, LoRusso P, Boerner S, Dummer R: Achievements and challenges of molecular targeted therapy in melanoma. Am Soc Clin Oncol Educ Book 2015:177-186. DOI: 10.14694/EdBook_AM.2015.35.177

[17] Falini B, Martelli MP, Tiacci E: BRAF V600E mutation in hairy cell leukemia: from bench to bedside. Blood 128:1918-1927, 2016. DOI: 10.1182/blood-2016-07-418434

[18] Vollbrecht C, Lenze D, Hummel M et al.: RNA-based analysis of ALK fusions in non-small cell lung cancer cases showing IHC/FISH discordance. BMC Cancer 18:118, 2018. DOI: 10.1186/s12885-018-5070-6

Evidence in the benefit-risk assessment of biomedicines

By Elena Wolff-Holz and Professor Klaus Cichutek | Paul-Ehrlich-Institut, Langen

Marketing authorization (MA) of a medicine is based on rigorous benefit risk assessment of data providing high-quality evidence for a favorable benefit-risk balance.
The benefit-risk assessment for an MA is undertaken by independent and experienced regulatory experts from national or European national competent authorities in a variety of procedures. Submitted data characterize the contents and quality of the medicine, for biomedicines a detailed description of the manufacturing process including in-process controls is necessary because the quality of a biomedicine cannot be controlled by analysis of the end product only. In addition to non-clinical and pharmacological-toxicological results , data and knowledge from phase 1 to 3 clinical studies are presented.
A variety of types of MA procedures allow for flexibility to provide MA for approval of orphan medicines and/or in cases of high unmet medical need.

Marketing authorization procedures based on benefit risk assessment

Marketing authorization (MA) of a medicine is based on a rigorous benefit risk assessment of data providing high-quality evidence for a favourable benefit-risk balance. The benefit-risk assessment for an MA is undertaken by independent and experienced regulatory experts from national or European national competent authorities in a variety of procedures. Regulators perform a benefit risk assessment of new medicines based on evidence. Article 26 to Directive 2001/83/EC states that a marketing authorization (MA) shall be refused if either the benefit risk balance is not considered to be favorable, if therapeutic efficacy is insufficiently substantiated or if the medicine's qualitative and quantitative composition is not as declared. Similarly, in paragraph 25 of the German Medicinal Products Act (Arzneimittelgesetz (AMG)) it is stated that MA should be declined, if the demonstrated benefit risk balance was found to be unfavorable. Submitted data characterize the contents and quality of the medicine, for biomedicines a detailed description of the manufacturing process including in-process controls is expected because the quality of a biomedicine cannot be controlled by analysis of the end product only. In addition to non-clinical and pharmacological-toxicological data, data and knowledge from phase 1 to 3 clinical studies are presented.

For developers of medicines and for regulators, scientific guidelines (https://www.ema.europa.eu/en/human-regulatory/research-development) are important and it is here that the scientific evidence necessary for the targeted clinical indication and patient population is defined. Included are considerations on relevant clinical endpoints to measure therapeutic benefit and risks.. The Committee for Human Medicinal Products (CHMP) at the European Medici-

nes Agency (EMA) issues an opinion following scientific assessment of the quality, safety and efficacy data, which are presented in the MA application (MAA) dossier, by its member from the National Competent Authorities for medicines of the EU and European Economic Are (EEA) member states (NCAs). This is based on co-assessment of the Pharmacovigilance risk assessment Committee (PRAC) and, in case of an advanced therapy medicinal product (ATMP), of the Committee for Advanced Therapies (CAT). Only if the benefits of the investigational medicinal product (IMP) outweigh the associated risks, will an MAA be successful and result in MA.

In assessing the benefits and risks of a particular IMP, regulators are used to dealing with a more or less high degree of uncertainty. For example, it may be a big challenge to identify rare adverse drug reactions against a certain background incidence in a limited patient sample size, depending on the statistical assumptions underlying the specific clinical trial. For example a study of more than 160,000 patients would be required to detect a one in 1,000 incidence of a drug-induced adverse drug reaction, given a background incidence of six per 1,000 (Ref Eichler et al.).

During the benefit risk assessment, several aspects are taken into account: any possibility of a potential deficiency of the pharmaceutical quality of the medicinal product, the knowledge of the pharmacology and toxicology of the product and the totality of evidence of the efficacy and risks. Finally, the feasibility of risk minimization is assessed and a concrete risk management plan (RMP) may be mandated.

Dr. Elena Wolff-Holz *is a physician and Medical Assessor for Major Policy Issues at the Paul Ehrlich Institut focussing on autoimmune diseases, oncology, and biosimilars. Moreover, she works as a lecturer for pharmaceutical research at the Johann Wolfgang Goethe-University in Frankfurt/Main. She is a member of the Scientific Advice Working Party (SAWP) at the European Medicines Agency (EMA) and Chair of the Biosimilar Medicinal Products Working Party (BMWP) of the EMA. Before that, she held various managerial positions in research and clinical development functions in the biotech industry for 15 years.*

Prof. Dr. Klaus Cichutek *is the President of the Paul-Ehrlich-Institut, the Federal Institute for Vaccines and Biomedicines in Langen near Frankfurt/Main and extraordinary professor of Biochemistry at the Johann Wolfgang Goethe-University in Frankfurt/Main. Since 2012, he is a member of the Expert Advisory Panel on Biological Standardization (ECBS) and since 2014 of the Product Development for Vaccines Advisory Committee (PD-Vac) of the WHO. He is founding member of the German Center for Infection Research (DZIF).*

Variants of marketing authorization procedures

Standard MA **NO post-MA obligation to generate efficacy / safety data**	**Standard MA** **WITH post-MA obligation to generate efficacy / safety data**	**Conditional approval**	**Approval under exceptional circumstances**
Positive benefit risk	*Positive benefit risk*	*Positive benefit risk*	*Positive benefit risk*
Comprehensive data submitted with the dossier to obtain approval	**Comprehensive data in the approval dossier and data on issues from assessment submitted upon request during the evaluation phase**	***Insufficient data on E/S, but*** ■ **Unmet medical need** ■ **Benefit for public health from immediate availability > risk of uncertainty**	***Insufficient data***
Adaptation of content according to risk benefit analysis possible	Adaptation of content according to risk benefit analysis possible		
	MAH may be required to develop additional data through conditions imposed to the MA. Post MA-studies are ■ PASS (Post Authorization Safety Study) ■ PAES (Post Authorization Efficacy Study)	**Only possible, if** ■ Intended to treat seriously debilitating or life-threatening conditions ■ Orphan drugs ■ Emergency situations MAH expected to provide comprehensive data to switch to full MA	**Only possible, if** ■ Comprehensive data not possible due to rareness of condition ■ Limitations in scientific knowledge ■ Unethical to collect missing information MAH is NOT expected to be able to generate missing data
	MAH obliged to comply with ***obligations***	MAH obliged to comply wit ***special obligations***	MAH obliged to comply wit ***special obligations***
After 5 years renewal	After 5 years renewal	Valid for 1 year, renewable every year, switch to standard MA after 5 years	5 year validity with annual reassessment; switch to full MA not foreseen

Source: Internal compilation

There are several variants of MA procedures – national, national Europeanized such as mutual recognition and decentralized procedure and centralized procedure - and in all instances a positive benefit-risk balance must have been concluded before market access is granted.

In addition, there are a number of MA types. Standard MA has no further requirements, MA may include the obligation to generate more data post authorization (post authorization measures (PAM)), either regarding efficacy (post authorization efficacy study (PAES) and /or safety (post authorization safety study (PASS)). A comprehensive data set must be presented to the regulators for approval of a medicine and renewal of the MA must be obtained after five years with the possibility of an adaptation of content of the MA according to the updated benefit risk analysis.

Furthermore, there are two regulatory pathways for approval foreseen in which the applicant may submit data on efficacy and safety insufficient for regular MA procedures.

(1) Conditional MA for medicines addressing a high unmet medical need in which the benefit for public health from immediate availability exceeds the risk of uncertainty. This is only possible if it is intended to treat seriously debilitating or life threatening conditions or orphan products or in emergency situations. It is expected that comprehensive data will be provided at a later stage in order to switch from the conditional to a full MA. The MA holder (MAH) is obliged to comply with special obligations, the conditional approval is valid for one year and a switch to a standard MA is typically expected after five years.

A recent report on conditional MAs in the EU reveals that this pathway is rarely used but if granted the obligations are mostly fulfilled. Only 30 conditional MAs were granted in 10 years from 2006-2016 ((https://www.ema.europa.eu/en/ human-regulatory/marketing-authorisation/conditional-marketing-authorisation) for life-threatening conditions (24) , orphan medicines (14) and to address emergency situations (3) and listed by therapeutic area in oncology (19), Infectious disease (9), neurology (3) and ophthalmology (1). Post-authorization obligations were fulfilled in 70% of cases within the pre-specified timelines which substantiates that this regulatory tool works effectively. In 2017, EMA recommended 92 medicines for MA. Of these, 35 had a new active substance, i.e. one which had never previously been authorized in the EU (Reference: EMA Annual report, 2017) and of these three medicines (about 10%) received a recommendation for a conditional MA. Seven new medicines were recommended for MA following a review under accelerated assessment; this mechanism allows for a faster review of medicines of major therapeutic interest by EMA's scientific committees (within 150 days rather than up to 210 days). The Committee for Medicinal Products for Human Use (CHMP) issued negative opinions on six medicines in 2017 (6/35 or 15%). A negative opinion is given if the CHMP cannot conclude that the benefits of the medicine outweigh the risks.

(2) Another option in which the applicant may submit insufficient data on efficacy and safety for approval is the MA under exceptional circumstance . However, this is only possible if comprehensive data cannot be obtained due to the rareness of the condition or limitations in scientific knowledge or if it were unethical to collect missing information. In any case, it is not expected that the MAH will be able to generate missing data. Again the MAH is obliged to comply with special obligations but typically there is five-year validity period with an annual reassessment and a switch to full MA is not foreseen. MA under exceptional circumstances occurs rather rarely (only 32 MA under exceptional circumstances from 2002-2017) with reasons for MA under exceptional circumstances primarily being rarity of disease (78%), ethics (16%) or other (6%).

Examples for opportunities to create positive evidence from a regulatory perspective

Product	(Initial) Indication	Context	Basis of E/S	Outcome
Keytruda (pembrolizumab) Imfinzi (durvalumab)	Treatment of adv./ unresectable melanoma; adv./unresectable NSCLC	Define role of Biomarker	US: Fast track (single arm cohort) EU: 2 randomised studies	US: Accelerated approval EU: Standard MA = full approval
Zykadia (ceritinib) Alk inhibitor	2°L ALK-positive advanced NSCLC	Biomarker selected	Single arm study	Conditional MA with Annex II obligation
Yescarta (axicabtagene ciloleucel)	Pretreated R/R aggressive B-cell non-Hodgkin lymphoma (NHL)	Clear scientific context	Single arm study	Standard MA with obligation; prespecified rate; historical context
Hemlibra (emicizumab)	Hemophilia A	Clear scientific context	Randomized few patients; intrapatient control (NIS)	Standard MA with obligation; prespecified rate; historical context
Ilaris (Canakinumab)	Cryopyrin-associated periodic syndromes (CAPS; IL-1β overproduction	Rare, chronic disease, slowly progressive	Randomised study	MA under exceptional circumstance

Source: www.ema.europa.eu

During the last phase of the clinical development, double-blind randomised studies are the gold standard. However, these are not feasible in all cases.

Around 62% of the applicants who received a positive opinion for their medicine had received scientific advice from EMA during the development phase of their product. This procedure allows the regulatory experts of the EU national competent authorities in EMA´s Scientific Advice Working Party to provide early input on the kind of evidence that would be required for MA, and helps to reduce the risk of patients taking part in unnecessary or poorly designed clinical trials.

In summary, there is a high flexibility of MA decisions allowed within the regulatory framework. The comprehensiveness of data package will determine the type of MA and the Post-Authorisation Measures (PAM) are related to uncertainties remaining after the initial assessment of the benefit risk profile of the product.

When looking for the best way in developing a new medicinal product in its final stages of clinical development, clearly a double-blinded randomized controlled clinical trial (RCT) is the preferred option termed „gold standard". This may however not always be feasible and there are ot-

her possibilities to generate positive evidence from a regulatory perspective as will be discussed in the following examples of development of biological medicinal products.

Keytruda (pembrolizumab)

Whereas the US FDA has provisions for several accelerated pathways to MA, in the EU only two possibilities exist, namely conditional MA (discussed above) and accelerated assessment, which reduces the timeframe from 210 to 150 days, if the CHMP decides the product is of major interest for public health and therapeutic innovation. (Ref J. Martinalbo et al.)

The first-in-class checkpoint inhibitor, Keytruda, was approved by the US FDA based on data from two cohorts of a multi cohort study (Keynote 001, cohort B2, B3), which evolved out of one phase 1 trial conducted in patients with unresectable or metastatic melanoma. Keytruda was applied as first- or second-line therapy and in different strengths and different posologies. Unprecedented double digit response rates were achieved. Even though no prospective comparison to available standard treatment had been made, this unprecedented, extraordinary high response rate led to an early approval by US FDA. In contrast, the EU waited until further results of two large randomized controlled clinical trials, one in the ipilimumab (IPI)-naïve population (Keynote 006) and one in patients previously treated with IPI (Keynote 002) were analyzed and respective data were submitted and assessed, after which the full MA and label in all advanced melanoma patients was granted.

Imfinzi (durvalumab)

The European Commission (EC) is hesitant in providing MA for medicinal products based on positive results obtained in a subset of a larger patient population. Vice versa, if the overall trial results are positive and it can be singled out that only a clearly defined subgroup of patients may be put at unacceptable risk when taking the medication, then such subgroup may be excluded from the MA label. For example in the case of Imfinzi (durvalumab), the fourth approved PDL-1/PD1 checkpoint inhibitor, obtained MA fromFDA for unresectable, stage III non-small cell lung cancer (NSCLC) based on positive evidence for efficacy in the overall patient population of NSCLC patients studied. However, in patients with very low PD-L1 expression on the tumor cells (TC) with TC <1% a hazard ratio well above 1 indicated an increased risk and this subgroup was explicitely excluded in the MA granted by EC.

Zykadia (ceritinib)

In rare instances, data from single arm trials are accepted as sufficient evidence of clinical benefit. Ceritinib is an ALK-positive inhibitor primarily used for the treatment of metastatic NSCLC. Ceritinib received conditional MA based on two single arm clinical trials that demonstrated an overall response rate (ORR) of 40-57% (compared to prior evidence of about 25%), progression free survival (PFS) of 6-7 months (compared to prior evidence of about 3-4 months) and overall survival (OS) of 15-16 Months (compared to prior evidence of less than 12-months). Furthermore, no inacceptable risks were observed in more than 300 patients studied. At the time of the initial conditional MA in 2015, two specific obligations were imposed, one of which was the completion of an ongoing randomized controlled clinical trial comparing ceritinib to chemotherapy in the target indication and both specific obligations were fulfilled in 2017 allowing to switch to a full MA.

Yescarta (axicabtagene ciloleucel)

Yescarta is a CD19-directed autologous T cells-based immunotherapy. The medicine is classified as a gene therapy medicine. Its target group are patients with pretreated aggressive B-cell non-Hodgkin lymphoma (NHL). Yescarta was approved on the basis of one phase 1/2 open label, multicenter single arm trial (ZUMA-1) conducted in 108 patients. The primary endpoint was its ORR in 62% of patients. Secondary endpoints of the trial included duration of response (DOR), which lasted 14 months, overall survival (OS), which was more than 17 months, and severity of adverse events.

Nearly all patients experienced some degree of cytokine release syndrome (CRS), which can however be managed by infusions of tocilizumab, and neurologic toxicities, occasionally requiring intensified monitoring and occasional treatments, but not precluding a positive benefit risk balance.

The ORR was prespecified to be tested in the first (and a minimum of) 92 treated patients and was significantly higher than the prespecified rate of 20% ($P < 0.0001$) which during the regulatory review procedure needed to be substantiated by data from a new historic control cohort. Thus, a retrospective, patient-level, pooled analysis of outcomes in refractory aggressive NHL ($N = 636$) was conducted (Crump et al., 2017) to provide confirmation of the prespecified control response rate of 20% and provide a historical context for interpreting the ZUMA-1 results. Response and survival rate after treatment with available standard of care therapy was evaluated and found to be in the same range as the pre-specified rate and for all evidence parameters i.e. the ORR was 26% and the complete response rate (CR) was 7% with a median OS of 6.3 months which is well less than the therapeutic benefit observed with Yescarta treatment. Based on this positive benefit risk outcome, the applicant was obliged to conduct a non-interventional post-authorization safety study (PASS) in order to assess the safety and manageability of occurring adverse drug reactions (ADRs) in patients with B-lymphocyte malignancies treated with axicabtagene ciloleucel.

In summary, this case shows that data from a single arm study could suffice as evidence for MA as the clinical indication represents an unmet medical need and is a life-threatening condition, a clear up-front classification of the disease is possible and the mechanism of action is supported by a strong scientific rationale and/or preclinical data. Furthermore the markers of efficacy were accepted clinical endpoints and the medicine induced substantial and unprecedented responses rates compared to historical data. A positive benefit risk balance was decided and the applicant was obliged to conduct a non-interventional PASS study, in this case by using a registry.

Hemlibra (Emicizumab)

Hemlibra is a bispecific monoclonal antibody which bridges activated factor IX and factor X to substitute for the missing function of factor VIII (in its activated from)that is needed for effective haemostasis. Hemlibra is indicated for routine prophylaxis of bleeding episodes in patients with haemophilia A with factor VIII inhibitors. Emicizumab has no structural relationship or sequence homology to factor VIII and, as such, does not induce or enhance the development of direct inhibitors to factor VIII as fully humanized monoclonal antibody.

Hemlibra prophylaxis was evaluated in a randomised, multicentre, open-label clinical study in 109 adolescent and adult males with haemophilia A and detectable factor VIII inhibitors who had previously received either episodic (=on demand) or prophylactic treatment with bypassing agents (aPCC and rFVIIa).

Fifty-three patients previously treated with episodic (on-demand) bypassing agents were randomised in a 2:1 ratio to receive Hemlibra prophylaxis (n=35 patients) or no prophylaxis (n= 18 patients). The observed number of bleeds requiring treatment with coagulation factor over time was significantly reduced with an absolute bleeding rate (ABR) of 2.3% versus 23%, i.e a reduction. by 87% ($p<0.0001$). Other secondary endpoints pointed in the same direction with reduction the number of all bleeds (80% reduction; $p < 0.0001$), spontaneous bleeds (92%, < 0.0001) and similar results for joint bleeds and target joint bleeds, as well as assessing patients' health-related quality of life and health status.

Furthermore, twenty-four patients previously treated with prophylactic bypassing agents and then treated with hemlibra prophylaxis were compared to patients previously treated with episodic (on-demand) bypassing agents who had participated in a non-interventional study (NIS) prior to enrolment. The NIS in this case was an observational study with the main objective of capturing detailed clinical data on the bleeding episodes and haemophilia medication use of patients with haemophilia A outside of an interventional trial setting. In this intra-patient analysis, Hemlibra prophylaxis resulted in statistically significant ($p = 0.0003$) and clinically meaningful reduction (79%) in bleed rate for treated bleeds compared with previous bypassing agent prophylaxis collected in the NIS prior to enrolment .

In summary, even though only very few patients actually got treated with hemlibra, the available evidence was sufficient for a positive conclusion on a favourable benefit risk ratio, especially as hemophilia is accepted to be a rare, stable disease and an up-front classification of the patient group studied in clinical trial was possible. The mechanism of action of hemlibra/emicizumab is supported by a strong scientific rationale and/or preclinical data and the clinical endpoints used to establish efficacy (e.g. annualized bleeding rate etc.) are all accepted and the biomarker results were supportive. Hemlibra produced substantial unprecedented clinical responses compared with no treatment (randomized comparison) and also other prophylactic treatment (intrapatient comparison in non-interventional, observational study) thus leading to positive benefit risk for the full label.

Ilaris (Canakinumab)

Canakinumab is a human anti-human-IL-1β monoclonal antibody indicated for the treatment of Cryopyrin-Associated Periodic Syndrome (CAPS), in adults and children 4 years of age and older. CAPS refers to rare genetic syndromes generally caused by mutations in the NLRP-3 gene which result in excessive release of activated IL-1β that drives inflammation. Canakinumab binds to human IL-1β and neutralizes its activity by blocking its interaction with IL-1 receptors.

The efficacy and safety of ILARIS for the treatment of CAPS was demonstrated in a randomized clinical trial of Ilaris (N = 15 patiens) versus placebo (N = 16 patients) and consisting of three parts.

Part 1 was an 8-week open-label, single-dose period where all patients received ILARIS. Patients who achieved a complete clinical response and did not relapse by week 8 were randomized into Part 2, a 24-week randomized, double-blind, placebo-controlled withdrawal period. Patients who completed Part 2 or experienced a disease flare entered Part 3, a 16-week open-label active treatment phase.

In part 1, a complete clinical response was observed in 71% of patients one week following initiation of treatment and in 97% of patients by week 8. In the randomized withdrawal period, , the primary endpoint was defined as the

proportion of patients with a disease relapse/flare: a total of 81% of the patients randomized to placebo showed a flare as compared to none (0%) of the patients randomized to ILARIS. At the end of part 2, all 15 patients treated with ILARIS had absent or minimal disease activity and skin disease. This dramatic treatment benefit was observed with only few serious adverse events (AEs) reported. The most commonly reported adverse reactions in the CAPS patients were nasopharyngitis, diarrhea, influenza, headache, and nausea, which was considered acceptable compared to the benefit for the patients.

In summary, CAPS is an extremely rare and predictable disease. The mechanism of action of canakinumab is well understood and is supported by strong scientific rationale and preclinical data. The endpoints used to measure autoimmune activity are scientifically and clinically accepted and the treatment effects are overwhelmingly strong as measured by both, the primary endpoint and inflammatory markers. Even though this orphan disease is very rare, the regulators encouraged to conduct a randomized controlled trial with a creative clinical trial design in very few patients. Due to the evidence provided, the drug was approved with a MA under exceptional circumstances since it was considered unethical to conduct further studies with such strong treatment effect and no other available treatments for this disease.

Summary

For a marketing authorization (MA), the manufacture of a biomedicine has to be well controlled and consistent, the quality has to be high, the non-clinical data have to show proof-of-principle, pharmacology and little or no toxicity. Based on clinical data usually collected in phase 1, 2 and then phase 3 clinical trials regulators perform a rigorous benefit risk assessment. The preferred clinical evidence is obtained in a randomized controlled clinical trial, but there are exceptions, especially in rare clinical conditions and lack of feasibility to conduct many or large clinical trials. Single arm trials, historic controls and inter- patient plus intra-patient comparisons in creative clinical trial designs can make efficient use of data obtained in few patients.

Quite often biomarkers help to more accurately select patients who may benefit most or experience risks less often, thus paving the way to a positive benefit risk balance. This reduces the number of patients exposed to potentially inactive drug substances and results in reduced resource consumption. However, potential challenges remain such as uncertainty with regard to the robustness of results observed and the risk of overestimating beneficial effects. In some cases, further studies may be required after approval, sometimes as an obligation, termed PASS and PAES, to rmove any remaining uncertainties.

The available regulatory system is flexible enough and provides sufficient tools to enable a rigid scientific assessment and MA of innovative, high quality products, which are highly needed.

References

[1] Eichler,H.G.; Aronsson, B., Abadie, E., Salmonson, T.; 2010, New drug approval success rate in Europe in 2009;Nature Reviews Drug Discovery; volume 9, pages 355–356

[2] EMA Annual report, 2017
https://www.ema.europa.eu/en/documents/annual-report/2017-annual-report--european-medicines-agency_en.pdf

[3] Martinalbo, J. Bowen D, Camarero J, Chapelin M, Démolis P, Foggi P, Jonsson B, Llinares J, Moreau A, O'Connor D, Oliveira J, Vamvakas S, Pignatti F. (2016) Early market access of cancer drugs in the E; Ann. Oncol. 27: 96-105.

[4] Crump M, Neelapu SS, Farooq U, Van Den Neste E, Kuruvilla J, Westin J, Link BK, Hay A, Cerhan JR, Zhu L, Boussetta S, Feng L, Maurer M, Navale L, Wiezorek J, Go WY, Gisselbrecht C. (2017) Outcomes in refractory diffuse large B-cell lymphoma: results from the international SCHOLAR-1 study; Blood, 130(16):1800-1808

Suitability of single arm studies for the evaluation of the additional benefit

Thomas Kaiser | Institute for Quality and Efficiency in Healthcare (IQWiG), Cologne, Germany

Direct comparative randomised controlled trials (RCTs) generally provide the most appropriate evidence for comparative benefit assessments. However, single arm studies can also be submitted. Depending on the indication, these studies were submitted in varying numbers (hepatitis C: 78 percent of all research questions, oncology: 32 percent, and others: 13 percent). For the benefit assessment it is not decisive whether an individual single arm study is relevant, but whether the comparison of single arm studies on the intervention and comparative treatment is relevant. Suitability of such a comparison is thus largely dependent of the data on the comparative treatment. Moreover, an additional benefit can only be derived from such a comparison in case of large differences. Conducting single arm studies for the purpose of benefit assessment should be limited to these exceptional cases.

Background

The purpose of early benefit assessment is the evaluation of an additional benefit of a new active substance as compared to the previous standard treatment (appropriate comparative treatment).[1] Direct comparative randomised controlled trials (RCTs) are considered the gold standard for this purpose.[2] This standard of evidence-based medicine is also reflected in the regulation on early benefit assessment, i.e. Ordinance on the Benefit Assessment of Pharmaceuticals (AMNutzen-V).[3] However, if no direct comparative RCTs are available, the pharmaceutical company can also submit studies with a lower evidence level to furnish proof of the additional benefit of a new active ingredient.[3]

Besides non-randomised, direct comparative studies e.g. studies about a new active ingredient each of which is not aimed at comparing a new active ingredient with the appropriate comparative treatment (e.g. single arm studies or individual study arm from studies without comparator arm with the appropriate comparative treatment). These studies can and are in fact submitted in some dossiers of pharmaceutical companies for the evaluation of the additional benefit. The purpose of this article is to present the benefit assessment on the basis of these studies. The following points will be addressed:

- Conceptional and general methodological considerations on the suitability of single arm studies for the evaluation of the additional benefit
- Empirical assessment of the significance of single arm studies in previous early benefit assessments

The special topic „Extrapolation in children and adolescents“ for which single arm studies might also be presented in dossiers is not addressed in this article. Moreover, benefit assessments of pharmaceuticals for rare diseases

(orphan drugs) are also not subject of this article, as the additional benefit is considered to be proven upon market access regardless of the available evidence.[1]

Question of additional benefit assessment – The PICO process

The research question for a benefit assessment according to AMNOG is specified in Section 35a of the 5th German Social Codebook (Sozialgesetzbuch V, SGB V) and the corresponding Pharmaceutical Products Benefit Assessment Ordinance (AM-NutzenV):[1,3] Does the new active substance provide an additional benefit as compared to the appropriate comparative treatment for patients for whom the new active substance was approved with regard to patient-relevant endpoints? This question can be addressed using the PICO process which is common in evidence-based medicine: Population, intervention, comparator, outcome. Comparison includes comparing the results of intervention and comparator regarding defined endpoints within a predefined population constituting the treatment effect of the intervention as compared to the control.

Therefore, direct comparative RCTs are known to be the best suited evidence for treatment comparisons, because randomisation ensures fair starting conditions for both the intervention and control group: The mean of the intervention group is similar to the mean of the control so that potential differences between the treatment options are not entirely attributable to different starting conditions (e.g. more severely ill patients in the reference group). This is not the case in case of individual arms from various studies. In case of an evaluation on the basis of single arm studies should thus be verified whether the evaluated populations are sufficiently similar.

Moreover, it remains unclear whether there are at all single arm studies on the comparative treatment in which the relevant population was investigated. In fact, while single arm studies on the new active ingredient mostly justify their approval and inclusion criteria for these studies are thus congruent with the approved field of application, this is not per se the case in studies on the comparative treatment. When comparing individual arms of various studies, it is thus important for the respective pharmaceutical company that studies by other manufacturers or study groups are available on the comparative treatment in which the approval population of his new active ingredient has been investigated.

In the discussion about the suitability of single arm studies, there is normally less focus on the similarity of investigated endpoints between the respective studies . This does not only refer to which endpoints were evaluated in the studies to be compared, but also whether these endpoints have been operationalised in a different way and whether

Dr. Thomas Kaiser *is a physician and system developer. Following his time working as a programmer, he studied medicine in Cologne and worked in the field of internal medicine for several years. In 2002, he founded the Institute for Evidence-Based Medicine (IQWiG) in Cologne. Since 2004, the year in which the IQWiG was founded, he is engaged as Head of the Drug Assessment Department, and since 2011 as the joint Chairman together with Dr Beate Wieseler.*

the observation period is sufficiently similar. Here, direct comparative RCTs also offer clear advantages as quantity and operationalisation of endpoints apply both for the intervention and control arm (different observation periods between treatment arms can also be problematic in RCTs, see e.g.[4]). On the other hand, there is no guarantee that the studies on the comparative treatment evaluated the same endpoints as the studies on the new active ingredient, as these might have had another study objective. Even if equal endpoints (e.g. pain) have been evaluated, these might have been operationalised in such a different way (e.g. considerably different thresholds for „treatment success") that a comparative statement is not possible on the basis of these data.

In case of comparisons of individual arms of various studies, this is compounded by the fact that the available study data on the comparative treatment might remain unclear due to an insufficient publication quality, as important information on population and endpoints is missing (see e.g. benefit assessment of ceritinib for bronchial carcinoma[5] or nivolumab for Hodgkin lymphoma[6]).

All of this shows that a high-quality single arm study by the pharmaceutical company alone is not sufficient for the additional benefit assessment: The question is not whether an individual single arm study is suitable, but whether the comparison based on single arm studies on the intervention and comparative treatment is suitable. Suitability of single arm studies by the pharmaceutical company mainly depends on the available study data about the comparative treatment.

Non-adjusted comparisons and the „dramatic effect"

As mentioned above, similar starting conditions are created for the intervention and control group in a RCT through the randomisation and the evaluated groups have similar mean risks of death, morbidity, and adverse effects. This is ensured by similar distribution of known, but especially also unknown risk parameters. Such an equal distribution of risk parameters is not guaranteed if individual study arms are compared. Although there are methods for adjustment of the known and evaluated risk parameters (e.g.[7]).

These methods also depend on the fact that the relevant parameters have been determined at all in the studies and patient-individual data are available for adjustment (both is often not the case in studies of the comparative treatment). On the other hand, adjustments do not solve the underlying problem of potential inequality of distribution of unknown risk parameters. This assumption (cited as an example) from a publication on adjustment methods „...fundamental assumption of no unmeasured confounders"[7] is usually fundamentally wrong.

However, there are certain situations in which compari-

Box A: Early benefit assessment dossier – structure of result presentation in module 4 of the dossier

4.3.1 Results of randomised controlled studies

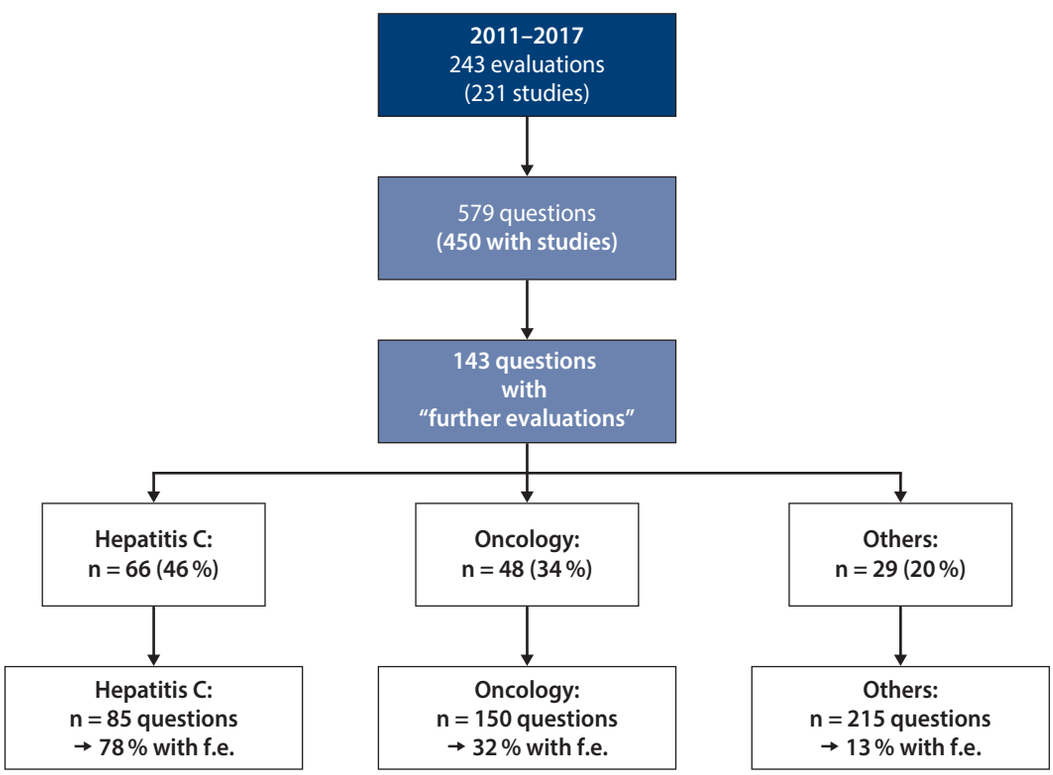
4.3.2 Further documents

- **4.3.2.1** Indirect comparisons on the basis of RCT
- **4.3.2.2** Non randomised intervention studies
- **4.3.2.3** Further evaluations

Quelle: Rule of Procedures of the Federal Joint Committee (G-BA)[10]

Single arm studies must be assigned to „Further evaluations" and are thus normally part of chapter 4.3.2.3 in module 4 of the dossier.

Frequency of "further evaluations" in early benefit assessment dossiers



Source: Internal analysis

Figure 1: The dominance of the indication hepatitis C becomes apparent when comparing the number of cases in which „Further evaluations" were submitted.

sons on the basis of single arm studies allow statements on the advantages or disadvantages of a certain treatment option despite the very high bias risk, i.e. if the observed differences are so significant that they cannot be explained by the bias alone. Glasziou et al. developed a recommendation for the results of simulation studies according to which an observed relative risk of 5 to 10 could not plausibly be explained by interfering variable influences.[8] The threshold also depends on the circumstances, e.g. quality of the evaluated studies, study design (e.g. comparison of the results of an unblinded with those of a blinded study), and consistency of the results in one endpoint category.[9]

Single arm studies in previous benefit assessments: Structure of the dossier for early benefit assessment

Direct comparative RCTs are the best suited evidence for

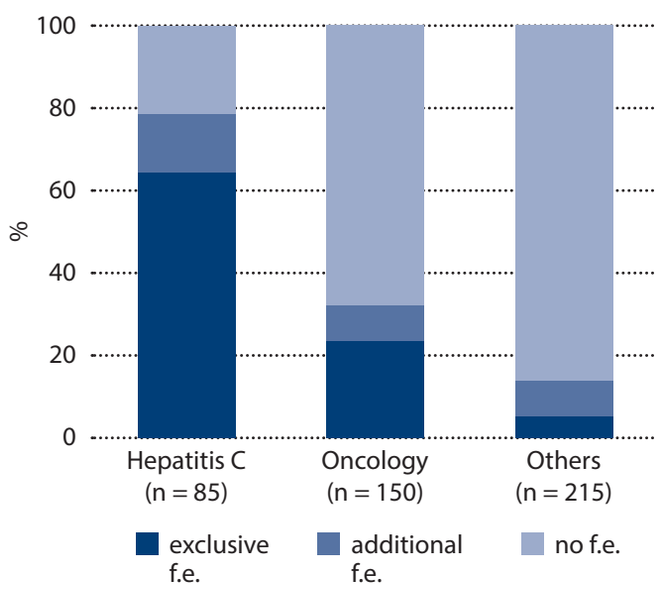
the evaluation of an additional benefit.[2] If these studies are available, the pharmaceutical company must present them in their dossier.[3] Moreover, it is up to the pharmaceutical company to submit other studies in the early benefit assessment dossier to furnish proof for the additional benefit of a new active ingredient. This is reflected in the structure of an early benefit assessment dossier. The structure of the dossier that has to be submitted by the pharmaceutical company is clearly specified in the Rule of Procedures of the Federal Joint Committee (G-BA).[10] Module 4 of the dossier shall contain the pharmaceutical company's processed data on the benefit assessment.Box A shows the structure of the results' section in module 4 of the dossier according to the G-BA's specifications. The single arm studies presented in this article must be assigned to „Further evaluations" and are thus normally part of chapter 4.3.2.3 in module 4 of the dossier. Thus, all assessments described below refer to the provision of evidence about the additional benefit in this chapter of the dossier. These studies might also be found in other chapters of the dossier, and expert statements without any underlying studies can occasionally be presented under „Further evaluations". As these are only individual cases according to our current experiences, this vagueness is accepted in the subsequent statements, as it can be expected that the basic statement will not be affected significantly hereby.

Submission of „Further evaluations" in previous dossiers

Figure 1 shows in how many benefit assessments „Further evaluations" have been submitted by pharmaceutical companies between 2011 and 2017. In total, this comprises 243 evaluations with 579 questions (one evaluation can comprise several questions, e.g. first and second line treatment). For 12 of 243 evaluations (5 percent) or 129 of 579 questions (22 percent), respectively, the pharmaceutical company did not submit any studies on the additional benefit in the dossier, neither direct comparative RCT nor „Further evaluations".

For 143 of 450 questions with studies (32 percent) „Further evaluations" were provided either as the only evidence or as additional evidence. In the majority of cases, these questions related to the indication hepatitis C (66 questions, 46 percent) or oncology (48 questions, 34 percent). All other indications only made up 29 of the questions with „Further evaluations" (20 percent). Preference of providing

Indication-specific frequency of "further evaluations" in early benefit assessment dossiers



Source: Internal analysis

Figure 2: In the vast majority of research questions „Further evaluations" were submitted as the only evidence in the indication hepatitis C.

„Further evaluations“ for the indication hepatitis C becomes even more apparent if all questions in the respective indications are considered. Figure 2 shows this.

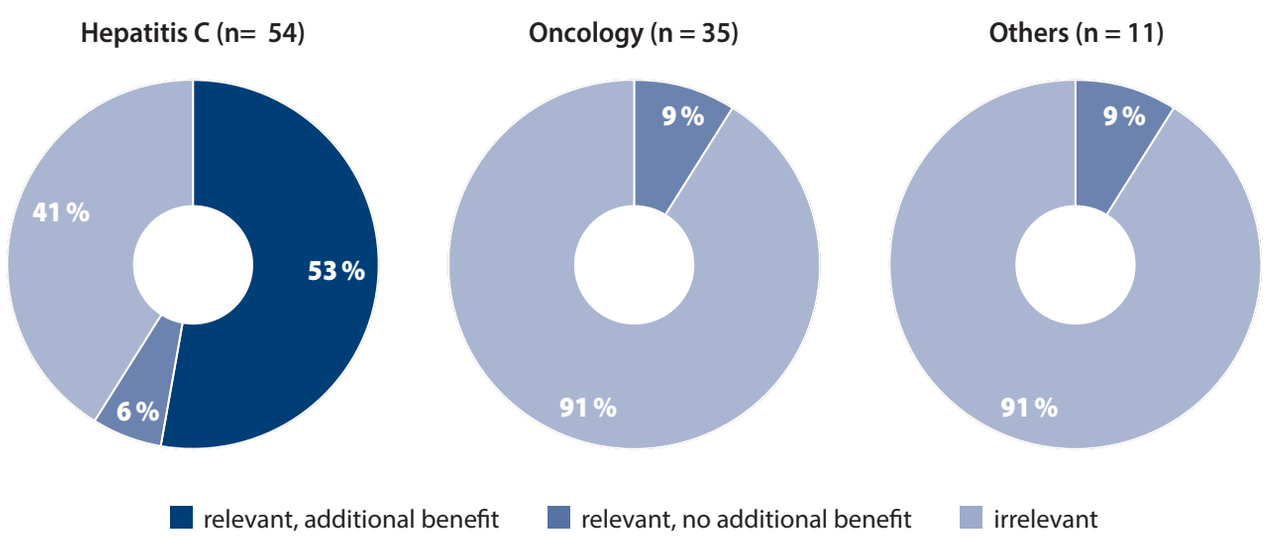
In the period under review, a total of 85 questions were evaluated in the indication hepatitis C. In 66 – and thus in the majority of questions (78 percent) – „Further evaluations“ were submitted, mostly as the only evidence (54 questions) and partly as additional evidence (12 questions). Contrary to this, „Further evaluations“ were less important in the indication oncology, whilst still 25 percent of all questions „Further evaluations“ were the only evidence presented by the pharmaceutical company. For all other indications, „Further evaluations“ were only provided in exceptional cases.

Results from „Further examinations“

Figure 3 shows whether the G-BA considered „Further evaluations“ as provided by the pharmaceutical company relevant in its decisions and if yes, whether an additional benefit was derived for this new active ingredient. Only those studies were relevant for which the G-BA presented the results in the decision on the benefit assessment. Only those questions were evaluated for which „Further evaluations“ were the only evidence provided by pharmaceutical companies.

The G-BA considered some 50 percent of „Further evaluations“ provided by pharmaceutical companies in the indication hepatitis C as relevant and derived an additional benefit for the new active ingredient from the relevant

Indication-specific relevance of “further evaluations” in the decisions of the G-BA



Source: Internal analysis

Figure 3: In the indication hepatitis C, the presented „Further evaluations“ were relevant in about half of the cases from the G-BA’s perspective. In other indications, this rate is significantly lower.

„Further evaluations“ in most cases. By contrast, in the indication oncology as well as in the other indications „Further evaluations“ were mostly irrelevant. In the three cases in the indication oncology in which „Further evaluations“ were relevant according to the G-BA's assessment, an additional benefit was derived for the new active ingredient. In all other indications, „Further evaluations“ did not result in an additional benefit in any of the cases. Figure 4 shows the extent and confidence level the G-BA determined across all indications in cases in which an additional benefit was derived from „Further evaluations“ (a total of 32 cases).

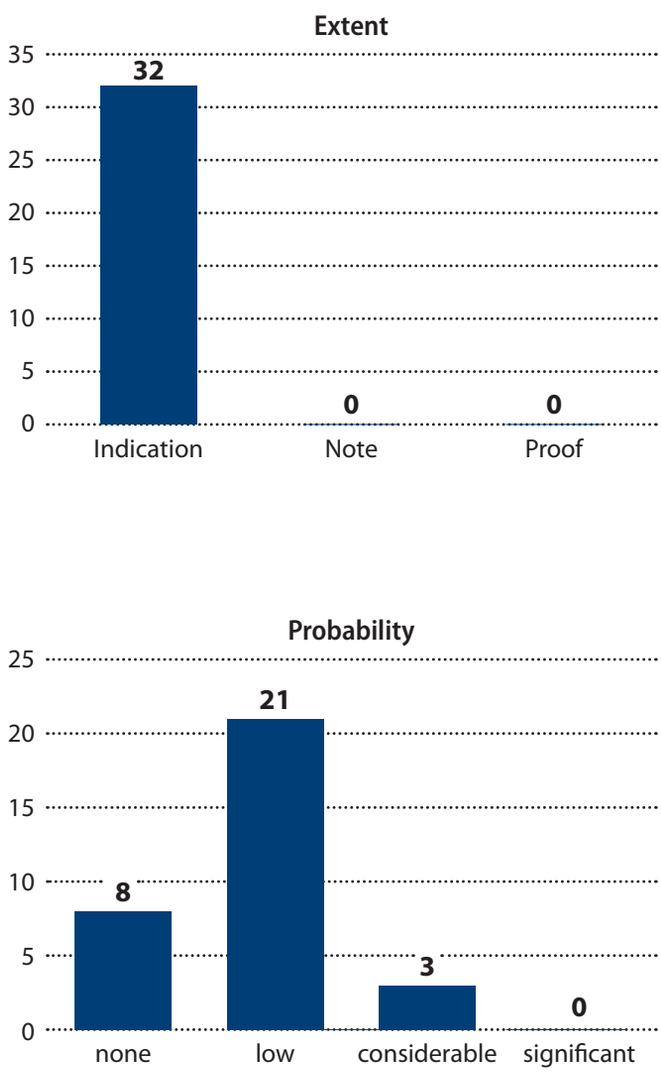
In the majority of the cases (24, 75 percent), the additional benefit was quantified by the G-BA on the basis of these studies and a low additional benefit was determined in most cases. A non quantifiable additional benefit was determined in 25 percent of the cases. In all 32 cases, the confidence level was low („indication“).

Classification of results from decisions on benefit assessment and conclusions

Results from previous early benefit assessments show that – depending on the indication – single arm studies (as „Further evaluations“) are either submitted in many cases or only in exceptional cases. In the indication hepatitis C, single arm studies were often the only evidence submitted during the past years. Due to the fact that previous treatment options were insufficient in some instances (e.g. in the treatment of children and adolescents, patients with decompensated liver cirrhosis or HIV-co-infected patients, but also due to the high adverse event rates of previous standard treatments) these studies were sufficient in half of the cases to derive an additional benefit, however, with a consistently low reliability of the results.

On the other hand, the single arm studies presented in the indication oncology were rarely suitable to derive any

Extent and probability of the additional benefit in the assessments on the basis of “further evaluations”



Source: Internal analysis

Figure 4: In 24 of 32 cases, the additional benefit was quantified by the G-BA on the basis of „Further evaluations“, i.e. in most cases as a „low“ additional benefit.

statements on the additional benefit. This underlines the criticism of the increasing trend of accelerated approvals without sufficient generation of evidence evidence,[11] As the studies that were conducted for this purpose are often not suitable to inform a appropriate treatment decision. Only in exceptional cases the opposite is the case, e.g. the active ingredient vismodegib for the treatment of basal cell carcinoma. Due to the fact that there were no satisfactory treatment options for the application field of vismodegib and spontaneous remissions were not reasonably to be expected, the G-BA derived an additional benefit on the basis of a single arm study.[12]

The overall conclusion is that single arm studies are not per se unsuitable for the evaluation of the additional benefit. However, their suitability is highly dependent on the context, i.e. whether sufficiently solid and published knowledge about the comparative treatment is available (e.g. from single arm studies or about the natural course of the disease). But if no major effects can be expected through the new treatment, e.g. because positive treatment results are achieved with the previous standard treatment, an additional benefit cannot be derived from single arm studies, even if they are generally suitable for the evaluation. Comparative assessments on the basis of these studies are potentially subject to major bias so that small treatment effects cannot be derived from these evaluations with sufficient certainty.

In conclusion, from a methodological and research economical perspective, but in particular in view of the fact that knowledge generation should primarily improve patient care, the conduction of single arm studies should be limited to the few exceptional cases in which very large (dramatic) treatment effects can reasonably be expected (also on the basis of sufficiently solid knowledge about the comparative treatment) from a new active ingredient.

References

[1] SGB V Manual (2017). 5th German Social Codebook – Health Insurance. Altötting: KKF-Verlag; 2017.

[2] Sackett DL (1979), Bias in analytic research. J Chronic Dis 32(1-2): 51-63.

[3] Federal Ministry of Health (2014). Pharmaceutical Products Benefit Assessment Ordinance according to § 35a Paragraph 1 SGB V for reimbursement agreements according to § 130b SGB V (Ordinance on the Benefit Assessment of Pharmaceuticals, AM-NutzenV); Date: 27.03.2014. URL: http://www.gesetze-im-internet.de/bundesrecht/am-nutzenv/gesamt.pdf.

[4] Unkel S, Amiri M, Benda N, Beyersmann J, Knoerzer D, Kupas K, Langer F, Leverkus F, Loos A, Ose C, Proctor T, Schmoor C, Schwenke C, Skipka G, Unnebrink K, Voss F, Friede T (2019). On estimands and the analysis of adverse events in the presence of varying follow-up times within the benefit assessment of therapies. Pharm Stat 18(2): 166-183.

[5] Institute for Quality and Efficiency in Healthcare (IQWiG) (2015). Ceritinib: Benefit assessment according to § 35a SGB V; dossier assessment dated 29/09/2015; Order A15-24 (IQWiG reports; volume 329). URL: https://www.iqwig.de/download/A15-24_Ceritinib_Nutzenbewertung_35a_SGB-V.pdf.

[6] Institute for Quality and Efficiency in Healthcare (IQWiG) (2017). Nivolumab: Benefit assessment according to § 35a SGB V; dossier assessment dated 30/03/2017; Order A16-76 (IQWiG reports; volume 501). URL: https://www.iqwig.de/download/A16-76_Nivolumab_Nutzenbewertung-35a-SGB-V_V1-0.pdf.

[7] Neugebauer R, Schmittdiel JA, Zhu Z, Rassen JA, Seeger JD, Schneeweiss S (2015). High-dimensional propensity score algorithm in comparative effectiveness research with time-varying interventions. Stat Med 34(5): 753-781.

[8] Glasziou PP, Chalmers I, Rawlins M, McCulloch P (2007). When are randomised trials unnecessary? Picking signal from noise. BMJ 334(7589): 349-351.

[9] Institute for Quality and Efficiency in Healthcare (IQWiG) (2017). General Methods Version 5.0. URL: https://www.iqwig.de/download/Allgemeine-Methoden_Version-5-0.pdf.

[10] Federal Joint Committee (G-BA) (2019). Annex II to the 5th Chapter of the Rule of Procedures of the Federal Joint Committee (G-BA): Format and structure of the dossier, documents to be submitted, specifications for technical standards. URL: https://www.g-ba.de/richtlinien/anlage/167/

[11] Banzi R, Gerardi C, Bertele V, Garattini S (2015). Approvals of drugs with uncertain benefit–risk profiles in Europe. Eur J Intern Med 26(8): 572-584.

[12] Federal Joint Committee (G-BA) (2016). Decision on the amendment of the German Pharmaceutical Products Directive (AM-RL): Annex XII - Decisions on the benefit assessment of pharmaceuticals with new active substances according to § 35a SGB V – vismodegib. Decision dated 04/08/2016. URL: https://www.g-ba.de/downloads/39-261-2672/2016-08-04_AM-RL-XII_Vismodegib_D-213_BAnz.pdf.

Creation and evaluation of evidence in endemic diseases on the example of diabetology

Professor Dirk Müller-Wieland | Clinical Trials Centre, Medical Clinic I, University Hospital of the RWTH Aachen

The German Pharmaceutical Market Reorganisation Act (AMNOG) of 2011 regulates both market access and financing of new pharmaceuticals by the self-governing bodies for our community of solidarity by means of early benefit assessment by the Institute for Quality and Efficiency in Healthcare (IQWiG) and final decision by the Federal Joint Committee (G-BA). In principle, this is basically welcome. According to the 5th German Social Codebook (Sozialgesetzbuch V, SGB V), criteria for benefit assessment include data on the patient's mortality, morbidity, symptoms and disease burden in comparison to an appropriate comparative treatment (ACT) as determined by the G-BA. Among other things, medical professional associations request to be involved procedurally in the determination of the research question, ACT, and operationalisation of endpoints or „determination of benefit", respectively. The problem with frequent and chronic diseases, e.g. type 2 diabetes, is to set up an optimal study design for the early identification of the criteria as determined by the SGB V and direct comparison to the ACT.

AMNOG and new diabetes medication

Just to avoid this misunderstanding, the objective of the evaluation of the additional benefit within the scope of AMNOG by the IQWiG after having been commissioned by the G-BA does not comprise an evaluation of effectiveness and safety, but rather the determination whether a new substance has an additional benefit within the approved application – as specified in the summary of product characteristics – as compared to a comparative substance that has been determined by the G-BA.

AMNOG does not consider sufficiently and procedurally the specialist knowledge of scientific professional associations, e.g. in the determination of an appropriate comparative treatment, determination and operationalised evaluation of an additional benefit, involvement of external experts and in case of deviations from guidelines.[1]

Between 2011 and 2018, 36 procedures with 110 subgroups were conducted in diabetology with no additional benefit being determined in almost 90 percent of the cases.[2] As a general rule, this was due to the legitimate evaluation on the basis of the SGB V (see above for the criteria) that a mere reduction of the surrogate parameter is not enough and data on the comparison to the ACT are missing. Further discussion points include different evaluations and assessments of hypoglycaemia and frequently missing data on the patient perspective, microvascular endpoints, and cardiovascular superiority. The situation has slightly changed since the US Food and Drug Administration (FDA) requests so called cardiovascular safety trials promptly after the approval in 2018 and has so far not revised its decision after ten years.[3,4]

Cardiovascular safety trials

There are three different designs of cardiovascular outco-

me trials (CVOT), e.g. comparison of treatment strategies, superiority, and safety trials (see overview[5,6]). In the comparison of treatment strategies, a potential difference between two substances with regard to cardiovascular events while ensuring a comparable reduction of the risk factor is assessed. At present, a study on oral diabetes medication is ongoing comparing glimepiride (sulfonylurea) with linagliptin (DPP-4 inhibitor). In superiority studies, the risk factor is reduced as compared to the placebo group; the best example is studies on statins.

The design of safety studies involves comparing the study substance with a placebo, but also reducing the risk parameter to be influenced, in this case HbA1c or blood glucose, respectively, in the placebo arm of the study, to a comparable level according to the study protocol. Thus, such a study design only examines whether a therapeutic molecule has its own effect or an effect beyond blood glucose reduction, respectively, on the cardiovascular risk. Cardiovascular safety studies confirmed the safety of DPP-4 inhibitors, superiority of individual GLP-1 receptor agonists like liraglutide, semaglutide, albiglutide and – according to a press release (results have not been published yet) – also for dulaglutide.

There was already consistent evidence in the safety studies for the renoprotective effect of the SGLT-2 inhibitors dapagliflozin, canagliflozin and empagliflozin, reduced hospitalisation for cardiac failure, as well as reduced MACE (major adverse cardiovascular event) and cardiovascular mortality in high-risk patients. The latter parameter was most evidenced in the EMPA-REG-OUTCOME study and thus the G-BA concluded that empagliflozin provides a considerable additional benefit in this patient group. Furthermore, this study data led to the inclusion of liraglutide and empagliflozin as ACT in the treatment of patients with diabetes and pre-existing cardiovascular disease.

Professor Dr. Dirk Müller-Wieland *is President (2017-2019) of the German Diabetes Association (DDG) and since 2011 speaker of the DDG-Commission Healthcare Policy. After his studies in human medicine in Hamburg he completed a two-year research stay at the Joslin Diabetes Center of Harvard Medical School. He completed his training as a specialist in internal medicine with a focus on endocrinology, diabetology and metabolism at the UKE in Hamburg and Cologne. Since 2016, he works at the Medical Clinic I, University Hospital of the RWTH Aachen.*

Transfer of these results to the „medical standard" in the treatment of type 2 diabetes patients from a national and international perspective

Pharmacotherapy of type 2 diabetes is a graded treatment that – according to the guidelines – should be patient-oriented and evidence-based. Annually updated practice recommendations of the DDG and common recommendations of the European Association for the Study of Diabetes (EASD) and American Diabetes Association (ADA) that were agreed upon and published simultaneously on 5 October 2018, were integrated into the recommendations of the „medical standard" of the ADA in 2019 and are summarised in the following section.[7-9] Metformin remains the treatment of first choice. The preferred recommendation

for other medications in addition to metformin depends on the patient-individual criteria and available evidence for the respective pharmaceuticals with regard to their antihyperglycaemic effect, risk of hypoglycaemia, influence on the body weight, individual adverse event profile, and especially its influence on cardiovascular and renal endpoints. This means that the results of new cardiovascular safety trials can be transferred and implemented into the clinical practice in a timely manner.

The recommendation on the preferred use of a specific substance reflects the patient populations included in the cardiovascular outcome trials. Consequently, sodium glucose transporter 2 inhibitors (SGLT2i) or glucagon-like peptide-1 (GLP-1) receptor agonists are preferably used in patients with pre-existing atherosclerotic cardiovascular disease (ASCVD).

The specific active ingredient of the respective substance class is selected that has demonstrated positive or cardioprotective outcomes in one CVOT and is approved in the respective healthcare system; for GLP-1 analogues, this is currently the case in Germany both for liraglutide and semaglutide as well as for the SGLT-2 inhibitors empagliflozin and dapagliflozin.

Strengths and weaknesses of cardiovascular safety trials

Figure 1 shows a summary of the strengths of cardiovascular outcome trials (also see[5,10]). This design does not involve assessing whether the glucose-lowering effect of a new antidiabetic agent has an effect on cardiovascular endpoints, but whether this new pharmaceutical is safe regardless of its potential glucose-lowering effect. Therefore, the protocol provides that the diabetes medication in the placebo group shall be escalated in accordance with local, mostly guideline-based circumstances. Hence, most studies also predefine a range of fasting glucose as target value on the basis of which treatment shall be modified.

As this is an add-on approach in most cases, i.e. a new medication is added to an existing medication, usually a large pool of data is obtained within the scope of these large studies comprising several thousands of patients about the safety and tolerability of the medication in patients who already have to take several pharmaceuticals due to their co-morbidities. For the early benefit assessment and an appropriate comparative treatment, the placebo arm should, as a rule, reflect the actual medical care situation. As the underlying glucose-independent mechanisms of

Cardiovascular outcome trials

Goal and strengths
- Do NOT test a reduction of the blood glucose level
- Increase the number of investigated patients promptly after the approval
- Safety, tolerance, pharmaceuticals (additional)

Clinical implications
- Placebo reflects actual medical care situation
- Patient-"group" oriented – implications for "LL"
- Morbidität der Patientenpopulationen und Effekte?

Weaknesses also for early benefit assessment
- Add-on approach
- No comparison of ACT or therapy strategies
- Short study duration (rapid event number and "non inferior")

Source: Professor Müller-Wieland

Figure 1: Strengths and weaknesses of cardiovascular safety studies for the early benefit assessment within the scope of AMNOG.

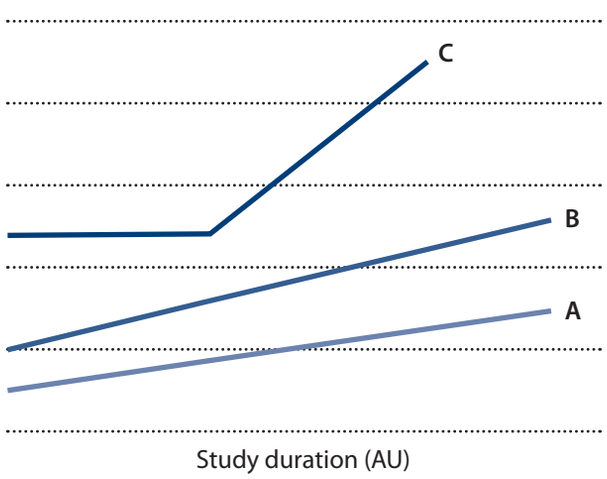
positive effects are not known, these studies are transferred to „preferred indications" in practice recommendations, however, referring to the patient population that was investigated in the respective study.

It remains an open question whether a particular clinical effect that was observed during the – mostly short – observational period is in fact directly related to the underlying morbidity or course of the disease, respectively; this fact limits a potential generalisation of the study results. This applies particularly to endpoints without a linear relationship between the drug-target-modulation and the clinical endpoint due to the disease stages (see figure 2).

Shortcoming of this study design may be that pharmaceutical companies want to rush these cardiovascular outcome trials, i.e. the recruited patient population has a high rate of clinical events and the design is event-driven. Thus, the study will be terminated after a statistically significant number of events has been reached. Therefore, study participants are usually cardiovascular high-risk patients, the study duration is short (two to three years), and the primary endpoint is the time until the first event in most cases.

For this purpose, a combined endpoint is usually selected, i.e. MACE usually comprising cardiovascular death, non-fatal myocardial infarction, and stroke. The short study duration and different probability or event rates of the individual components of the combined endpoint, respectively, often causes discrepancies, e.g. an effect is observed on the cardiovascular death, but not on the myocardial infarction, or an effect is observed on stroke without an effect on the other components. Thus, also effects on the myocardial infarction are observed „without effect" on cardiovascular death for in the interpretation that fact that these studies are limited in time is often neglected. For this reason, there are justified requests that these studies should have a minimum duration regardless of the event rate and that cumulative events in one individual should be assessed according to pre-specified terms.[10,11]

Relationship between event rates and duration of the disease



Source: Professor Müller-Wieland

Figure 2: There is a linear relationship between case A and case B, e.g. a relative risk reduction will be apparent in the two different populations, however with a different absolute risk reduction. In case C, the development of event rates depends on the underlying morbidity or risk (vertical axis), respectively, as well as the course of the disease that is not linear. Thus, a corresponding effect – even on the relative risk reduction – could not be observed in population A and B.

Cardiovascular endpoint studies and subgroup analyses versus ACT

Cardiovascular safety trials that are conducted on a global or internationalised scale, make a dedicated subgroup ana-

lysis versus an ACT that has been determined by the G-BA for the early benefit assessment mostly after beginning of the study impracticable. For this reason, the time when the study design was set up and the „medical standard" that was applicable at that time must be taken into consideration during the evaluation of these studies within the scope of early benefit assessment.

It would be desirable that approval authorities would consider further secondary endpoints in these studies – in accordance with the study protocol – for a potential approval extension. This would be more effective also from the patients' perspective and it wouldn't take such a long time until studies are available with the respective primary endpoint. The prerequisite is, of course, that the results are unequivocal and can be operationalised as if they corresponded to a primary endpoint.

Operationalisation of clinical results

It would thus be feasible that not only the appropriate comparative treatment is predefined at an early stage and in a transparent manner, but also that clinical endpoints for microvascular complications associated with diabetes, such as nephropathy, retinopathy and potentially also neuropathy, are operationalised and predefined in a transparent manner by the IQWiG (possibly in cooperation with the medical professional association). This would be fair to the pharmaceutical companies investing in these studies, as they would have a sound basis of national interests to be considered in the design of global studies and would also provide a higher degree of legitimacy and plausibility for methodological procedures during early benefit assessment.

This also applies for the identification of symptoms or adverse events in diabetes patients; the most frequent example of the discussion during early benefit assessment is the operationalisation of hypoglycaemic events. The problem with this symptom is that its clinical significance is not only dependent on the absolute low value achieved, but also on the duration, starting condition, speed of the reduction of the blood glucose level, as well as the underlying morbidity (e.g. with or without autonomous neuropathy, age, etc.).

Moreover, hypoglycaemia can be an indicator or mediator for complications and upon its determination a protocol-related bias between the treatment arms can occur. In this context, a new topic is the so-called „time-in-range" as determined by continuous measurement of the glucose levels in tissues; however, its clinical significance with respect to the course of the disease as well as criteria according to SGB V, i.e. mortality, morbidity, symptoms, and health-related disease burden are still completely unknown.

Adequate determination of the patient perspective

When dealing with this complex topic and diabetes – as an example for a chronic disease – both patient perspective and patient-reported-outcomes (PRO) are important parameters. It becomes clear that the tools normally used have not been sufficiently standardised to compare different studies and that the determination of the individual domains (e.g. physical functioning, anxiety, fatigue, sleep disorder, limitations imposed by pain, participation in social roles and activities, etc.) is not aligned with the specific underlying disease.

There is an urgent need for methodological and clinical development, as the sensitivity and specificity of these instruments are too low, i.e. the rate of false negative evaluations is unacceptably high and might be detrimental for an adequate consideration of patient-related treatment-relevant endpoints.

Outlook

According to the author, the patient's subjective treatment objective could be evaluated as a treatment-relevant endpoint. In chronic diseases, the patient's perspective on the treatment could be considered on an individual basis.

Moreover, if particular research questions shall be addressed during early benefit assessment, so-called real-world data and registry data should be taken into consideration wherever possible. However, prioritisation of study qualities for the procedure must be clearly predefined. In future, supplementary data from so-called cluster analyses or responder profile or subgroups established by means of new methods of deep learnings will create new perspectives allowing for a more patient-centred evaluation of an additional benefit for the large number of patients suffering from endemic diseases.

Objectives of diabetes treatment

Short-term
- Control HbA_{1c} without hypos and weight gain
- Increase of the number of investigated patients promptly after the approval
- Safety, tolerance, pharmaceuticals (additional)

Mid-term
- Avoidance of micro and macrovascular complications
- Reduction of morbidity and good quality of life

In future
- Remission – cure – prevention
- Stop the course (progression)
- Reduce neurodegenerative and malignant risk

Source: modified after[12]

Figure 3: Objectives of diabetes treatment and criteria for the development of new treatment interventions.

According to the author, „open questions" should be discussed and agreed upon in guidelines to provide a clear framework for action and research for the community of solidarity, the pharmaceutical company, and research policy. Fundamental issues of macro-social relevance that cannot be addressed by pharmaceutical companies due to conflicts of interest, should be determined by an independent body and implemented by special funding. Therefore, short and long-term diabetes therapy objectives must be discussed, specified, and operationalised (see figure 3).

Conclusions

According to AMNOG, an additional benefit is a patient-relevant therapeutic effect with respect to the patient's health condition, duration of the disease, survival, adverse events, and quality of life that is qualitatively or quantitatively higher as compared to that of the ACT. All stakeholders must accept the AMNOG criteria for an additional benefit. Involvement of scientific professional associations in the process of early benefit assessment should be defined procedurally.

Both the determination and evaluation of the respective clinical endpoints should be operationalised by the IQWiG or G-BA, respectively – possibly in cooperation with the respective scientific professional associations – in order to provide the necessary transparency at an early stage ensuring a reliable study design during the procedure. The patient's health-related quality of life determined by disease is an essential criterion for the patient perspective, is process-relevant during benefit assessment, but has been insufficiently addressed so far both methodologically and clinically.

References

[1] Gallwitz B, Kellerer M, Müller-Wieland D. „Diabetes mellitus in Deutschland – politische Handlungsfelder" (Diabetes mellitus in Germany – political fields of action). In: German Health Report Diabetes 2019. Ed. J. Kröger and D. Müller-Wieland. Verlag Kirchheim + Co GmbH, Mainz; p 276-282.

[2] Sackett DL (1979), Bias in analytic research. J Chronic Dis 32(1-2): 51-63. Wörmann B and representatives of various professional associations for the Association of the Scientific Medical Societies in Germany (AWMF), Ad hoc Commission Benefit Assessment and the German Society for Haematology and Medical Oncology (DGHO). Early benefit assessment of new pharmaceuticals in Germany 2011-2018, February 2019.

[3] US Department of Health and Human Services Food and Drug Administration. Guidance for Industry. Diabetes mellitus – Evaluating Cardiovascular Risk in New Antidiabetic Therapies to treat Type 2 Diabetes. 2008. www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm071627.pdf

[4] Phend C. FDA advisors: tweak, don´t dump CV outcomes trials.2018. www.medpagetoday.com/cardiology/prevention/75938.

[5] Cefau WT, Kaul S, Gerstein HC et al. Cardiovascular outcome trials in type 2 diabetes: Where do we go from here? Reflections from a Diabetes Care editors´ expert forum. Diabetes care 2018; 41: 14-31

[6] Home P. Cardiovascular outcome trials of glucose-lowering medications: an update. Diabetologia 2019; 62: 357-69

[7] Davies MJ, DÁlessio DA, Fradkin J, Kernan WN, Mathieu C, Mingrone G, Rossing P, Tsapa A, Wexler DJ, Buse JB. Management of hyperglycemia in type 2 diabetes, 2018. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes EASD. Diabetologia 2018; 61: 2461-2498 and Diabetes Care 2018; 41: 2669-2701.

[8] American Diabetes Association. Standards in Medical Care in Diabetes – 2019. Pharmacological Approaches to glycaemic treatment. Diabetes Care 2019; 42 (Suppl 1): S90-S102.

[9] Landgraf R, Kellerer M, Aberle J et al. „Therapie des Typ-2-Diabetes. Diabetologie und Stoffwechsel" (Treatment of type 2 diabetes. Diabetology and metabolism) 2018; 13 (Supplement): S83-290

[10] McGuire DK, Marx N, Johansen OE, Inzucchi SE, Rosenstock J, George JT. FDA duidance on antihyperglyacemic therapies for type 2 diabetes: One decade later. Diabetes Obes Metab 2019; 21: 1073-1078.

[11] Claggett B, Pocock S, Wie LJ et al. Comparison of time-to-first event and recurrent-event methods in randomized clinical trials. Circulation 2018; 138: 570-577.

[12] Müller-Wieland D. „Orale Antidiabetika ermöglichen Individualisierung der Behandlung" (Oral antidiabetics allow for treatment individualisation). MMW Fortschr Med 2015; 157 (Suppl1): 69-74

A glance into EUnetHTA's perspective on evidence requirements

Giovanni Tafuri, PhD, Senior Scientific Officer, EUnetHTA

The European Network of Health Technology Assessment aims to contribute to a sustainable model for the scientific and technical cooperation on HTA in Europe. Core activities are the production of joint assessment reports and the procedures of early dialogues with manufacturers, both in close collaboration with EMA. Although alignments of HTA and regulatory evidence requirements, as well as the definition of common HTA methodologies and joint assessments among EU Member States, play an important role in facilitating patient access to new treatments, the key hurdle to access remains price.

The European Network of Health Technology Assessment consists of 83 national or regional institutions dealing with HTA. The project, now in its third edition (Joint Action 3, 2016-2020), aims to contribute to a sustainable model for the scientific and technical cooperation on HTA in Europe in close collaboration with stakeholders and the European Commission. It is based on a voluntary cooperation in order to produce HTA joint work, but also to increase the uptake and implementation of such joint work at the national, regional, and local levels.

One of the core EUnetHTA activities is the production of joint HTA reports, the Relative Effectiveness Assessments (REAs), which can either be based on Pharmaceuticals or on Medical Devices. For pharmaceuticals, these reports are typically authored by a national HTA body with the support of another HTA body from a different Member State acting as co-author. The process of assessment also sees the involvement of a set of HTA bodies acting as reviewers of the report.

In particular, during the current Joint Action for Pharmaceuticals, the process has so far produced four reports, while five additional assessments are in progress and others are in the preparation stage (https://www.eunethta.eu/rapid-reas/). To facilitate the use of these reports at the national level, their publication occurs shortly after the time of the marketing authorisation issued by the European Commission, following CHMP approval.

Common features of REAs are the definition of a PICO (Population, Intervention, Comparators and Outcomes) as a starting point for the assessment of the technology, the use of GRADE, the risk of bias assessment through the quality rating tool of the Cochrane Collaboration, the presence of systematic reviews to collect all available evidence, indirect comparisons and network meta-analyses, as well as

the inclusion of patients and experts inputs. Of note, a recent document by EUnetHTA Partners has identified common methodologies to engage with patients and benefit from their inputs during the process of REA production.[1]

A comparison between the published REAs and the related European Public Assessment Reports (EPARs) published by EMA shows there are commonalities between HTA bodies and regulators regarding the uncertainties that surrounded the drug at the time of regulatory approval, but differing responses to those uncertainties (Figure 1).

The different remits of EMA and HTA bodies are also reflected in the evidence used for the assessment. Indeed, HTA bodies need to consider additional studies to conduct indirect comparisons and network metaanalyses in order to assess relative effectiveness versus therapeutic alternatives not used in the registration trials.

Another core EUnetHTA activity was the establishment of procedures of early dialogues with manufacturers during clinical development in collaboration with the EMA. The aim is to allow manufacturers to integrate specific HTA and regulatory needs into the development plan and, therefore, to fulfil the evidence requirements of both regulators and HTA bodies at the same time.

A retrospective analysis based on a cohort of procedures of early dialogues between 2010 and 2015 found that there was commonality in evidence requirements between regulators and HTA bodies.[2] Whilst there was somewhat less commonality for the advice on comparators, the investigators noted an overall high degree of alignment between the EMA and HTA bodies.

Another analysis by the same authors explored the actual impact of parallel scientific advice on clinical developments, assessing the uptake of regulatory and health technology assessment recommendations.[3] One of the key findings was that manufacturers tend to implement changes to the development programme based on both regulatory and HTA advice with regards to the choice of primary endpoint and comparator. However, the analysis also confirmed the challenging choice of the study comparator, for which manufacturers seem to be more inclined to satisfy the regulatory advice.

In general, the issue of different evidence requirements between regulators and HTA bodies has been widely debated within the scientific community. A single regulatory authorisation system in Europe, with single legislation and well-defined assessment criteria, is indeed compared with the environment for 28 different Member States, all of which operate under different legislations, HTA methodologies, criteria, and reimbursement systems.

Dr. Giovanni Tafuri *is the Scientific Director of the European Network for Health Technology Assessment (EUnetHTA) at the EUnetHTA-Secretariat in the Netherlands. Prior to his appointment, he worked at the HTA and in the approval department of the Italian Medicines Agency (AIFA) (between 2006 and 2018) and as designated national expert for scientific consultation at the EMA (from 2015 to 2017). He is a pharmacists with a Master of Science in International Health Care Management, Economics and Policy, a doctoral degree in Pharmaceutical Policy and Regulation and worked as scientific associate at several national and international institutions.*

In considering how these two very different systems might be aligned, Eichler and colleagues published a paper in 2010 outlining the current and potential future paradigms for interaction between regulators and health technology assessors.[4] Since that time, many other authors have examined the differences and consequences of those differences between the two groups of stakeholders.

The objectives of such analyses have been various. For example, a systematic evaluation of oncology approvals by the EMA in 2009-13 showed that most drugs entered the market without evidence of benefit on survival or quality of life, which is considered crucial for HTA and by pricing and reimbursement decision-makers.[5] In general, uncertainties related to the benefit-risk of products at the time of regulatory approval pose serious challenges to the definition of the added benefit of new technologies conducted by HTA bodies and of its overall value, which is essential for pricing and reimbursement decision-makers. Although obligatory post-approval studies are often imposed by regulatory authorities to fill the evidence gap existing at the time of regulatory approval, analyses of such studies showed that more than half were completed with a substantial delay, or not at all.[6] The current scenario is made even more complex by the variability in pricing and reimbursement decisions and time to reimbursement among EU countries as is widely documented.[7]

Regorafenib for patients with hepatocellular carcinoma (HCC)

EMA/CHMP EPAR[1]	EUnetHTA REA[2]
■ RESORCE study, OS increase (2.8 months) is considered a clinical benefit	■ RESORCE study, OS increase (2.8 months) is considered a moderate improvement
■ Insecurities: Sorafenib-intolerant patients; patients with ECOG PS > 1 and/or Child Pugh B → stipulated by changes in the summary of product characteristics (SmPC)	■ Insufficient evidence about the impacts on HRQoL ("regrettable" for final stage patients)
	■ Evidence gaps: Sorafenib-intolerant patients; patients with ECOG PS > 1 and/or Child Pugh B → further research data is required

Key:
1. EMA/CHMP EPAR EMEA/H/C/002573/II/0020.
2. EUnetHTA REA Project-D: PTJA02

Quelle: Dr. Tafuri

Figure 1: The different areas of responsibility of the EMA and HTA institutions are also reflected in the evidence used for the evaluation.

Although further alignments of HTA and regulatory evidence requirements, as well as the definition of common HTA methodologies and joint assessments, play an important role in facilitating patient access to new treatments, the key hurdle to access remains price.

It is now globally acknowledged that increasing drug prices pose serious concerns to the sustainability of healthcare systems. A common justification for high drug prices is the sizable research and development outlay necessary to bring a drug to market. Estimates of R&D spending range from $2.7 billion (2017 US dollars) to $650 million.[8] More transparency is therefore needed on price determinants from manufacturers. Many proposals have already been made, including the recent Italian resolution approved by the WHO which aims to improve public sharing of information on actual prices paid by governments and other buyers for health products, while promoting greater transparency on pharmaceutical patents, on the results of clinical trials, and other pricing determinants along the laboratory-to-patient value chain.[9]

High profits may be justified if novel products offer significant benefits to patients or if they represent significant pharmacological advances. However, in an analysis performed by Mailankody and Prasad on oncologic drugs approved by the US Food and Drug Administration in 2009 and 2013, a lack of relationship between the novelty or relative benefit of these drugs and their price was observed by the authors.[10] Indeed, little difference was found in price among drugs approved based on time-to-event endpoints and drugs approved on the basis of Response Rate. Results seemed to suggest that current pricing models are not rational, but simply reflect what the market will bear.

Villa et al also highlighted a mismatch between the value perceived by manufacturers and the value attributed by payers, analysing the negotiation process of 133 new compounds (44 orphan drugs and 89 new other molecular entities) conducted by the Italian Medicines Agency (AIFA) between 2013 and 2017.[11] Following the negotiation process, prices were lowered by 25.1% and 28.6% on average for orphan drugs and other molecules respectively. The price reduction was higher for innovative drugs (-32.2%). In addition, the authors found that the implementation of Managed Entry Agreements (in particular cost sharing agreements) were associated with higher price reductions during negotiations.

In this complex scenario, a new piece of EU legislation on HTA cooperation is currently under discussion.[12] One of its objectives is to increase cooperation on HTA evaluations which would be performed jointly by representatives of HTA national authorities. The underlying idea is to provide national EU authorities with reports synthesizing the available evidence using methodologies developed within the EUnetHTA Joint Actions, thus facilitating decision-making processes at a national level and reducing duplications, while still leaving pricing and reimbursement decisions to the individual Member States. The legislative process for such a regulation is currently ongoing and its outcome is expected within 2020

DISCLAIMER: The content of this article represents my own view only and is my sole responsibility; it cannot be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food Executive Agency or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use that may be made of the information it contains.

References

[1] Patient Input in Relative Effectiveness Assessments, EUnetHTA. Available at: https://www.eunethta.eu/wp-content/uploads/2019/06/Final_290519_Patient--Input-in-REAs.pdf

[2] Tafuri G, Pagnini M, Moseley J et al. Br J Clin Pharmacol.2016;82:965-973.

[3] Tafuri G, Lucas I, Estevão S et al. Br J Clin Pharmacol. 2018;84:1013-1019.

[4] Eichler HG, Bloechl-Daum B, Abadie E et al.Nat Rev Drug Discov. 2010, 4:277-91.

[5] Davis C, Naci H, Gurpinar E et al . BMJ. 2017;359:j4530

[6] Pease AM, 1 Krumholz MH, Downing NS et al. BMJ 2017;357:j168

[7] Variations between European Countries in Time to Reimbursement Decisions for New Cancer Treatments Approved by the EMA. ESMO 2018. Available at: https://www.ascopost.com/News/59385

[8] Vinay P, Mailankody S. JAMA Intern Med. 2017;177(11):1569-1575.

[9] 72nd World Health Assembly. Improving the transparency of markets for medicines, vaccines, and other health products. Available at: https://apps.who.int/gb/ebwha/pdf_files/WHA72/A72_ACONF2Rev1-en.pdf

[10] Mailankody S, Prasad V. JAMA Oncol. 2015; 4:539-40.

[11] Villa F, Tutone M, Altamura G, Health Policy 123 (2019) 595–600.

[12] Regulation of the European Parliament and of the Council on health technology assessment and amending Directive 2011/24/EU. Available at: https://ec.europa.eu/health/sites/health/files/technology_assessment/docs/com2018_51final_en.pdf

How to deal with real world data and registries? A political consideration

Michael Hennrich | Member of the German Bundestag

In light of an increasing number of orphan drugs their special status within the scope of the German AMNOG procedure needs to be further developed. The aim of health policy must be that these innovative pharmaceuticals continue to be available for patients as quickly as possible while ensuring their benefit. For this purpose, the Act for Greater Safety in the Pharmaceutical Supply System (GSAV) strengthens the role of registries and real-world data during benefit assessment.[1] In a first step, this abstract sums up the current knowledge regarding the existing legal situation and provides examples for registries. In a second step, the new provisions of the GSAV and resulting problems and challenges are illustrated.

New ways in benefit assessment: The perspective of politics of registries and real-world data

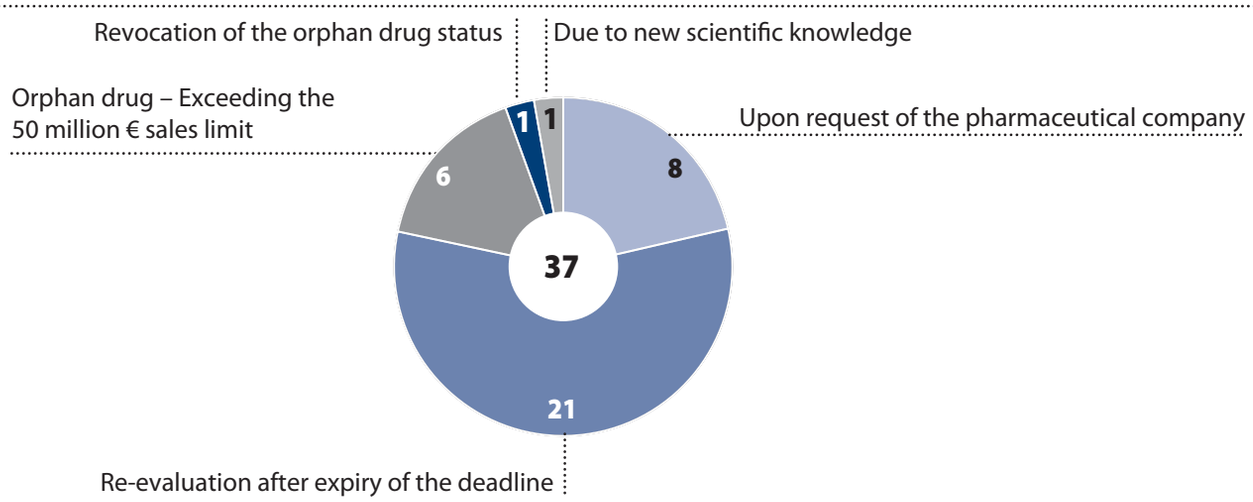
A constant factor of politics in our modern world is the challenge to acknowledge rapid technological progress, define its specific benefit, and make sure that it reaches people's everyday lives.

This maxim is particularly true in the pharmaceutical sector with its unique political environment: On the one hand, we see rapid progress here, and on the other hand, we deal with the precious asset of health. With the AMNOG procedure we have a modus vivendi that provides rapid access to the latest effective innovations for patients while keeping an eye on the total health care budget. At the same time, we are facing the challenges of an increasingly finer differentiation between innovative pharmaceuticals that are associated with improved chances for recovery for many patients. We as legislator still want that these innovations become available for patients as quickly as possible.

These innovations show that the existing provisions reach their limits and we need to further develop the exemption for orphan drugs. There's always an exception to the rule – and that's how it should stay. But against the background of an increasing differentiation between innovative pharmaceuticals, the existing requirements regarding the evidence base required both for the approval and the G-BA process need to be reviewed. The European Union adopted Regulation (EC) No 141/2000 as well as Regulation No 726/2004 and established a reasonable framework for the approval process. In order to gain a better understanding of the mode of action and benefit of orphan drugs and for documentation purposes, registries shall thus play a supplementary and increasingly important role during early benefit assessment.

With this instrument, we can combine two goals: To con-

Number of benefit re-assessment from 2011–2017 by reasons



Source: BPI-MARIS 2018

Figure 1: Orphan drugs play an important role during benefit re-assessment.

Michael Hennrich *studied law in Passau and Bonn. From 1991 to 1995, he worked as scientific associate for the member of the Bundestag Elmar Müller in Bonn. He has been a freelance lawyer since 1995. From 1998 to 2003 he was Regional Managing Director of the CDU Baden-Wuerttemberg Economic Council. Since 2002 he has been the representative for the Nuertingen constituency, since 2013 correspondent for the pharmaceutical supply and pharmacies. Since 2015, he works as Chairman of the Health Committee.*

tinue ensuring rapid availability of innovative orphan drugs for patients while providing long-term evidence for their additional benefit. Within the scope of the GSAV, we will substantiate these considerations at a legal level.

Status quo at a glance: Legal basis for benefit assessment

The starting point is Section 35a of the 5th German Social Codebook (Sozialgesetzbuch V, SGB V). It states that the G-BA must evaluate the benefit of all reimbursable pharmaceuticals with new active ingredients. Benchmark of this evaluation is the additional benefit as compared to the appropriate comparative treatment (ACT). For the implementation, Section 35a in Paragraph 1 Sentence 9 refers to an ordinance to be adopted by the Federal Ministry for Health (BMG) (without approval of the federal states). With regard to the requirements, the following points shall be stipulated:

„Principles for the determination of the appropriate comparative treatment and the additional benefit, and thus also in cases in which additional evidence is required as well as the conditions under which studies of a specific evidence level must be requested; the international standards of evidence-based medicine and health economy provide the basis for this."

The corresponding Ordinance on the Benefit Assessment of Pharmaceuticals (AM-NutzenV) defines the additional benefit as a therapeutic improvement that has to be proven by means of product information and clinical studies. Section 5 Paragraph 2 Sentence 3 also provides us with the instrument of choice for this purpose: „Clinical studies, in particular direct comparative studies with other pharmaceuticals of this reference price group with patient-relevant endpoints, primarily mortality, morbidity, and quality of life should be considered."

The following Paragraph 3 stipulates the exemption that has already been mentioned in the text of the regulation. In the event that studies with highest evidence are not feasible or cannot be reasonably requested, lower levels of evidence shall be used. Section 5 Paragraph 6 of the AM-NutzenV lists the levels of evidence. Moreover, in terms of legal consistency European legislation, we have a similar provision as provided by national law with Article 14 Paragraph 8 of Regulation No 726/2004.[2]

Against the background of the existing general regulation we can thus draw the following preliminary conclusion: According to Section 35a of SGB V, randomised, controlled studies (RCTs) are the gold standard of our approval sys-

Classes for the assessment of evidence

Class		Study requirements
I	Ia	Evidence on the basis of a systematic review of randomised controlled studies (possibly including meta-analysis)
	Ib	Evidence on the basis of at least one high-quality randomised controlled study
II	IIa	Evidence on the basis of at least one well-designed controlled study without randomisation
	IIb	Evidence on the basis of one well-designed quasi-experimental study
III		Evidence on the basis of well-designed non-experimental descriptive studies
IV		Evidence based on reports/opinions of experts, consensus meetings and/or clinical experiences of recognized authorities

Source: in line with Kleespies C; Kaise T; Sawicki PT; for the working group practical evidence-based medicine, St. Franziskus Hospital, Cologne; IeM – Institute for Evidence-Based Medicine (IQWiG), Cologne: Begriffe und Methode der evidenzbasierten Medizin – Ein Glossar (Terms and methodology of evidence-based medicine – a glossary).

Figure 2: The randomised, controlled study is the gold standard during early benefit assessment. At the same time, Section 35a includes exceptions for cases in which the highest evidence classes cannot be achieved.

tem as well as for early benefit assessment. At the same time, there are exceptions for cases in which the highest evidence classes cannot be achieved. And it should be noted that Section 5 Paragraph 5a AM-NutzenV also contains specific legal provisions for European approvals.

It allows the G-BA to determine an additional benefit for other patient groups or sub-indications, „if the transfer of evidence is admissible and well founded according to the current state of scientific findings also with regard to benefit assessment."[3]

At this stage, I would like to provide a first interim evaluation regarding existing provisions. From my point of view, only modifications of the drug regulation might be needed. Any further need for change might be difficult to realise on the level of the SGB V and would be in contradiction with the existing system. However, at the same time, I would express the request: Whoever sees a need for further modifications, should transparently specify and clearly point out what is required. Politics is sometimes baffled in light of far-reaching, but sometimes vaguely and not sufficiently elaborated demands communicated to us.

Where do we stand when it comes to registries?

Looking at existing registry examples provides clarity, and I would thus like to mention three areas.

1. Clinical cancer registries

Clinical cancer registries and the underlying provisions first come into mind when we talk about additional evidence from registries and real-world-data. In this context, the Federal Cancer Registry Data Act (BKRG) of 3 April 2013 and the new Section 65c in SGB V should be mentioned. The key phrases are:

1. „For the improvement of the quality of oncological care, the federal states will establish clinical cancer registries." (Paragraph 1, Sentence 1).

In sentence 2, the tasks and parameters that need to be fulfilled are specified – sub-divided into eight sub-items: In which area data should be collected, evaluation and feedback of the results to care providers, matters of cooperation, etc.

2. „Clinical cancer registration will be performed on the basis of the nationwide standardised data set […]" (Paragraph 1, Sentence 3). Moreover, the following crux should be mentioned:

3. „The necessary provisions for the establishment and operation of clinical cancer registries according to Sentence 2 including data protection provisions shall remain subject to national law."

A Prognos study from 2017 shows how heterogeneous the landscape of cancer registries is today. According to the study, there are different procedural stages, each federal state has developed its own concept, and often only base data sets are available.[4] The study demonstrates that the different stages can be attributed to various factors, such as the further development of epidemiological to clinical registries. The study states that:

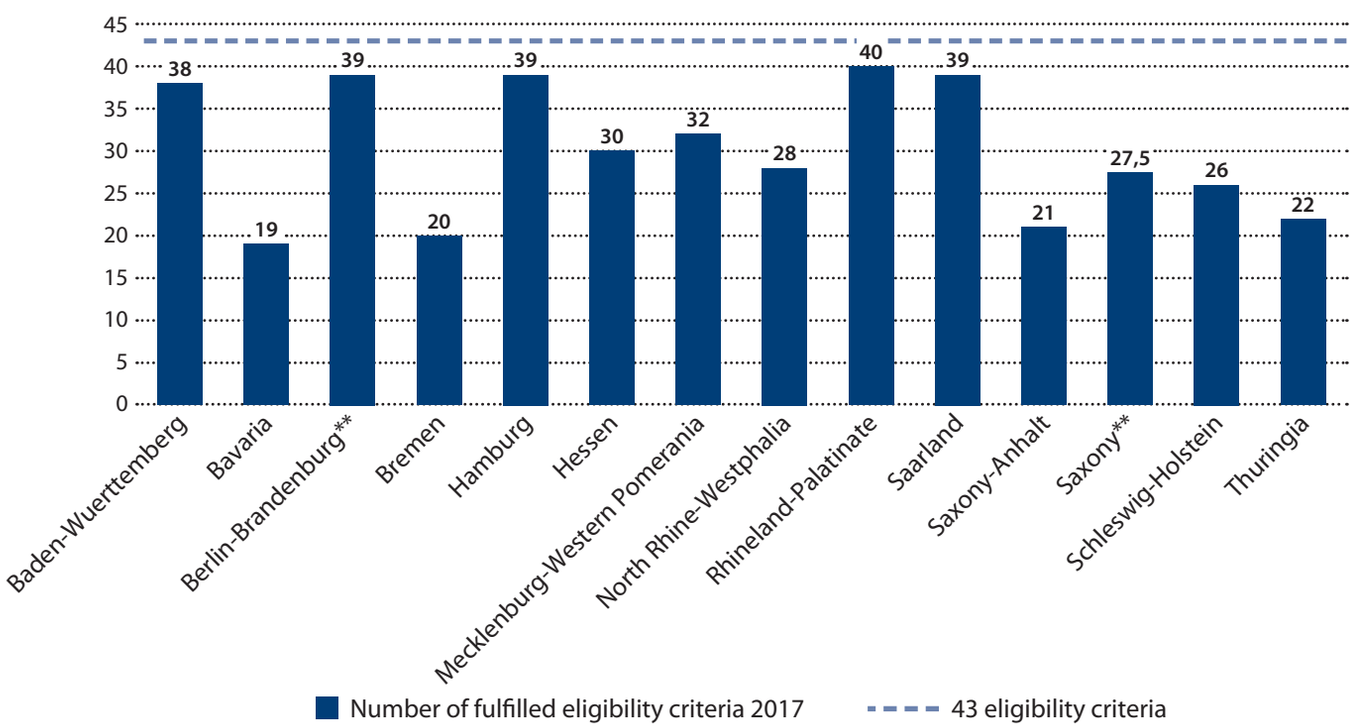
„By 31 December 2017, the fulfilment level of eligibility criteria of the individual federal states varies between the clinical cancer registries. The clinical cancer registries of Baden-Wuerttemberg, Berlin-Brandenburg, Hamburg, Rhineland-Palatinate, and Saarland for example had mostly fulfilled the eligibility criteria by the reference date, while the remaining clinical cancer registries lagged partly far behind in fulfilling the 43 eligibility criteria. According to the fulfilment reports, for some clinical cancer registries less than half of the eligibility criteria are considered fulfilled."[5]

The status of cancer registries shows that despite the fact that a lot has already been achieved, there is still a lot of work to do for the responsible stakeholders. Moreover,

there are several fundamental problems. From my point of view, the network of stakeholders and at data level needs to be improved. This means that we have to ensure compatibility of data and avoid double entries. In general, there are two possibilities: We can optimise existing system or we could establish a new acquisition system. However, consolidation of the results is mandatory.

2. Case study RABBIT – data collection in rheumatology: With its logo of a quick moving rabbit, RABBIT is also (almost) an acronym for the full title of the registry: „Rheumatoid arthritis: Observation of therapy with biologics." The project was established in 2001 by the German Rheumatism Research Centre Berlin (DRFZ) in coordination with the German Society of Rheumatology (Deutsche Gesell-

Number of fulfilled eligibility criteria of clinical cancer registries by federal state* by 31 December 2017



This information is based on: Fulfilment reports 2017; *without Lower Saxony; **joint CCR Berlin-Brandenburg; ***Average of decentral CCR Chemnitz (27 fulfilled criteria), Dresden (28 fulfilled criteria), Leipzig (28 fulfilled criteria), Zwickau (27 fulfilled criteria)
Source: Prognos AG, 2019

Figure 3: By the end of 2017, only five federal states had established cancer registries that fulfilled almost all eligibility criteria. Other federal states partly lagged far behind these specifications.

schaft für Rheumatologie), the Professional Association of Rheumatologists (Berufsverband Deutscher Rheumatologen), and the Competence Network Rheumatology (Kompetenznetz Rheuma). The registry assesses disease and therapy courses of more than 17,000 patients affected by rheumatoid arthritis. On the medical side, some 400 rheumatologists act as attending physicians in Germany.

RABBIT investigates the long-term safety of biologic disease modifying antirheumatic drugs (DMARDs). It is a mere observational study with a minimum period of 5 years. The registry is financially supported by a joint grant from all companies that provide biologics for therapy of rheumatoid arthritis. Both with its design and its realisation, RABBIT stands for the possibility to generate evidence by means of registries. It can therefore be used as an example in various aspects for the changes intended with the GSAV to increasingly generate evidence by other means.

3. Case study CRISP (Clinical Research platform into molecular testing, treatment and outcome of non-Small cell lung carcinoma Patients): Testing and treatment reality in lung cancer:

The Oncology in Internal Medicine Working Group (AIO) within the German Cancer Society is responsible for this large study. CRISP is a large, open, and also non-interventional, prospective registry study with 117 active participating centres. More than 8,000 patients with metastatic non-small cell lung cancer were recruited and followed-up to investigate the mode of action of these therapies and their effect on the patients' quality of life. Patients are followed up until their death or for a maximum of three years.

These three examples clearly show both chances and limitations of the further use of registry dates within the scope of the AMNOG processes. It is thus absolutely desirable that especially in case of extremely expensive therapies that often enter the market with only limited – often single-arm data sets – the introduction of these products on the market should be associated with a systematic data collection in daily clinical practice. On the other hand, cancer registries demonstrate e. g. how difficult it is to generate nationwide standardised data sets. But also the exemplary RABBIT registry has its starting point and focus in the evaluation of the risk profile of biologics in rheumatology. Statements on the comparative efficacy were also hardly made on the basis of the RABBIT registry.

The stringent methodology of prospective randomised data sets from clinical studies can hardly be reproduced under treatment conditions. For this reason, the collection of registry dates as specified by the GSAV to furnish comparative proof for an additional benefit requires further methodological discussions about the adequacy and use of such treatment data for comparative efficacy evaluations according to AMNOG.

New provisions in Section 35a SGB V – Content and goals

According to the new provisions in Section 35a Paragraph 3b, the G-BA can request post-market data collection for certain pharmaceuticals for the assessment of a benefit. These pharmaceuticals are specified in two sub-items, whereas Point 2 comprises pharmaceuticals for rare diseases.

In the paragraph, it is further stipulated:

„The Federal Joint Committee can limit the authorisation to prescribe the pharmaceutical at the expense of the statutory health insurance to those statutory health insurance physicians or approved hospitals participating in the post-market data collection." There is a clear objective, but I would like to point out that the regulation includes a „can" provision. For any further details – duration, type and scale, evaluation and formats – the text in the Section refers to the G-BA.

Also new: The consultation obligation in Section 35a Paragraph 7 will be extended to include post-market data collection. The G-BA must review the acquired data and conduction of data collection on a regular basis, at least annually. The review is performed in coordination with the Federal Institute for Drugs and Medical Devices (BfArM, Bundesinstitut für Arzneimittel und Medizinprodukte) and the Paul-Ehrlich-Institut. Target group of the provision are the pharmaceutical companies who have to bear the advisory costs incurred by the two institutions (Paragraph 7).

The justification includes further details. Thus, the act does not include a general limitation to institutions or centres. Nor does the legislator impose any specifications for the study design, such as randomisation. The type of data collection required depends on the purpose and proportionality. The European regulations to which Section 35a Paragraph 3 b refers should also be mentioned.

In this context, I would like to mention Article 14 Paragraph 7 and 8 of Regulation (EC) No 726/2004:[6]

„(7) Following consultation with the applicant, an authorisation may be granted that is subject to certain specific obligations and must be reviewed annually by the Agency. The list of these obligations shall be made publicly accessible.

„(8) In exceptional circumstances and following consultation with the applicant, the marketing authorisation may be granted subject to certain conditions, in particular relating to the safety of the medicinal product, notification to the competent authorities of any incident relating to its use, and action to be taken. The marketing authorisation may be granted for objective and verifiable reasons only and must be based on one of the grounds set out in Annex I to Directive 2001/83/EC. Continuation of the authorisation shall be linked to the annual reassessment of these conditions."

With respect to the approval of pharmaceuticals for the treatment of rare diseases, Regulation (EC) 141/2000 is also relevant.[7] It specifies the obligations and conditions for the approval of pharmaceuticals for the treatment of rare diseases by the European Medicines Agency (EMA). Regulation (EC) 507/2006 is even more specific with respect to the implementation. I would like to point out in particular Articles 3 to 5.[8] The key points include:

- Even in case of conditional approvals, the aim is a positive benefit risk balance.[9]
- Conditional approvals shall be limited to those cases „ in which only the clinical part of the application documents is less comprehensive than usual" (Paragraph 4).
- The aim is to close supply gaps.
- Approvals shall be subject to conditions, such as conduction of studies to furnish proof for a positive benefit risk balance.

At EU level, a comprehensive legal system has already been established to find the right balance between rapid availability of pharmaceuticals and later submission of medical evidence. One point of criticism of the GSAV is certainly that these existing regulations have somehow been neglected in the past. This reminds me very much of the discussion about orphan drugs when the then chairman of the G-BA did not know the regulation and requirements to grant an orphan drug status in the legislative process.

Criticism of the new provision

Patients in Germany have a good and rapid access to orphan drugs.[10] The challenge of the amendment is especially to further improve an already good status quo. The following points of criticism were mentioned:

1. The expansion of the calculation basis of €50 million that is defined as a sales limit for orphan drugs in Section 35a Paragraph 1, Sentence 12. So far, this sales limit relates

to statutory healthcare, but in future it shall also comprise inpatient treatment.

Moreover, manufacturers shall be required to provide information on the distribution of pharmaceuticals. Various arguments were made against this.

I. One of the criticisms made is that this self-commitment would compromise the established balance of interests.

II. Companies cannot accomplish that, as they cannot track the way of the pharmaceuticals and for which purpose they are used.

III. A sales limit would at least have to be adjusted to the rate of inflation.

2. Another point relates to the obligations for data collection and their subsequent acknowledgement by the G-BA. This demonstrates that the G-BA only rarely attested an additional benefit to an active ingredient in the past on the basis of single-arm studies. Registry studies can complement randomised controlled studies and create additional evidence, but whether they can replace them in future and whether they will be accepted as an alternative means of proof, must be viewed critically against the background of previous experiences.

In current practice, however, the determination of an additional benefit requires comparative evidence. In order to furnish proof by means of registry studies this means quality-secured data collection prior to the scheduled market entry as well as indication-specific registries.

3. For the evaluation of the benefit of pharmaceuticals for the treatment of rare diseases, in case of conditional approvals and under exceptional circumstances, the G-BA shall be authorised under exceptional circumstances to request post-market data collection by the pharmaceutical company according to Section 35 Paragraph 3b SGB V. If the pharmaceutical company does not comply with this obligation or if a quantifiable additional benefit cannot be determined, adequate markdowns of the reimbursement amount can be agreed upon (Section 130b SGB V). Thus, the question is what happens, if physicians or patients refuse consent to data collection. Does a patient have to seek treatment at a distant centre then in case of doubt?

4. One specialist question deals with the requirements on the evidence for the annual review as specified in Section 35a Paragraph 3b Sentence 6. However, it must be pointed out that sufficient time should be provided for a valid data collection and that availability of a sufficient number of patients is required especially in rare diseases.

Thus, the general question in connection with these aspects that refer to specific regulatory content is how future data collection can be better coordinated at European level and how the G-BA deals with these data. Post-market data collection is already part of the European approval process. Therefore the question is justified whether this might lead to a duplication of structures. In this regard, reference is also made to the current initiative of the European Network for Health Technology Assessment (EUnetHTA) for the establishment of methodological pillars for the establishment of registries.[11]

Pricing[12]

Once the deadline for the conduction of post-market data collection has expired and after a new decision has been taken about the benefit assessment, the reimbursement amount will be renegotiated (Section 130b Paragraph 3 Sentence 7 SGB V). If no quantification of an additional benefit can be proven, low annual treatment costs shall be determined that are below those of the previously agreed reimbursement amount.

From my point of view, further legal restriction is required. The draft law did not provide for a justification for cases in

which no additional benefit can be determined despite data collection. However, various scenarios would be possible in which the manufacturer would not be responsible for that. For example, if data are not documented appropriately by physicians or in case of technical problems in the digital management of registries. Moreover, it should be taken into account that according to the mentioned draft by the EUnetHTA, manufacturers shall not be involved in the governance of the registry which would certainly limit their influence.

One possible solution could be to limit the requirements and grant the G-BA more freedom of manoeuvre. Instead of an actual provision this would be the case with a directory provision. The wording of a new provision could be: „If […] on the basis of the collected data, a quantification of the additional benefit cannot be determined, a reimbursement amount shall be determined that results to reasonably reduced annual treatment costs as compared to the previously agreed reimbursement amount."

Pharmaceuticals for advanced therapies in the Medicinal Products Act (AMG, Section 4 Paragraph 9)

Starting point is the decision of the G-BA. This can be illustrated with a schematic overview: In this context, the Federal Social Court pointed out that valuation of the physician treatment share within the therapy is decisive.[13]

In case of an ATMP (advanced therapy medicinal product), the minimum requirements regarding the quality of structure, process and outcomes must be determined. These include in particular the qualification of care providers, structural requirements as well as other quality assurance requirements. The G-BA adopts the implementing provisions according to section 136 Paragraph 2 and 3 SGB V. It is both desirable and reasonable to involve the Paul-Ehrlich-Institut in these processes. During the final design of the law, this aspect will certainly play a role.

What is the role of registries and real-world data?

Registries are already used according to Section 35 SGB V in the context of prevalence, incidence, and course of treatment. They already play an important role in rare diseases and long-term courses. But how can they also create additional evidence? And how can they possibly complement RCTs?

Thinking about these questions, we should remember that registries can work in two directions: They can substantiate evidence with regard to an additional benefit, but they can also put positive findings into a different light. This becomes clear when we take a look into the Deutsches Ärzteblatt: „Orale Antikoagulation: Wenn Studienergebnisse und die reale Welt divergieren." (Oral anticoagulation: When study results and the real world diverge.)[14] This article is based on the data of three health insurances and shows that direct oral anticoagulants (DOACs) show inferior outcomes under routine conditions with respect to morbidity and mortality as compared to common vitamin K antagonists.

An obvious cause might be that the patients included in the RCT are kept uniform especially to ensure that study results remain comparable. In contrast, patients for whom RWD were collected were older and generally sicker and suffered from various pre-existing conditions. If real world data like in this case demonstrate that fewer patients die or suffer a stroke, we have to find an answer how to deal with these facts.

Within the scope of legislation and practise we must find detailed answers to these questions. However, there are certain framework conditions:

- The registry should be organised by independent institutions.
- We need a binding funding of registries with participation of the industry.

Pharmaceuticals or ATMP

Classification as pharmaceutical (outpatient or inpatient)	Classification as treatment method
Dossier requirement ■ Pharmaceutical nature ■ Easy administration	No dossier requirement ■ More complex administration ■ Administration associated with medical intervention
→ Early benefit assessment according to §35a SGB V	→ Assessment according to §135 or 137c SGB V, respectively

Source: Own presentation

Figure 4: For the determination of whether a specific pharmaceutical is an ATMP, the valuation of the medical treatment share within the therapy is decisive.

- The collected data must be complete and suitable for use in the system of benefit assessment.

It is foreseeable that further legal and systematic questions will arise from these changes: Does the AM-NutzenV have to be adapted? Do we need modifications of the G-BA's Rule of Procedures? Does the underlying Methods Paper of the IQWiG have to be adapted?

The good news is that medical progress will continue relentlessly creating new chances for recovery for many patients. But the challenge remains to find a way to make it available to patients quickly and effectively without overburdening health care systems. For politics and self-governing bodies this means: The end of one reform is the start of the next.

References

[1] This abstract is based on a presentation I held during the ongoing legislative process at the „INTERDISCIPLINARY PLATFORM ON BENEFIT ASSESSMENT" meeting in March 2019. The GSAV was adopted on 5th June by the Bundestag prior to the printing date.

[2] „In exceptional circumstances and following consultation with the applicant, the marketing authorisation may be granted subject to certain conditions, in particular relating to the safety of the medicinal product, notification to the competent authorities of any incident relating to its use, and action to be taken. The marketing authorisation may be granted only when the applicant can show that he is unable to provide comprehensive data on the efficacy and safety of the medicinal product under normal conditions of use, for objective, verifiable reasons and must be based on one of the grounds set out in Annex I to Directive 2001/83/EC. Continuation of the marketing authorisation shall be linked to the annual reassessment of these conditions.", Regulation (EC) No 726/2004 of the European Parliament and of the Council dated 31 March 2004 laying down Community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency.

[3] During the evaluation of pharmaceuticals with approval for paediatric use for the purposes of Article 2 Paragraph 4 of Regulation (EC) No 1901/2006 of the European Parliaments and of the Council dated 12 December 2006 on medicinal products for paediatric use and amending Regulation (EEC) No 1768/92, Directives 2001/20/EC and 2001/83/EC as well as Regulation (EC) No 726/2004 (Official Gazette L 378 dated 27.11.2006, p. 1), as last amended by Regulation (EC) No 1902/2006 (Official Gazette L 378 dated 27.12.2006, p. 20), the Federal Joint Committee evaluates whether for certain patient groups or sub-indications that are included in the approval, but are not or only insufficiently represented in the study population and for which approval was issued as a result of a transfer of evidence, an additional benefit can be determined. The G-BA can determine an additional benefit, „if the transfer of evidence is admissible and well founded according to the current state of scientific findings also with regard to benefit assessment."3 The G-BA specifies the detail in its Rule of Procedures.

[4] Status quo of clinical cancer registration. Results of the evaluation of eligibility criteria by 31.12.2017, expert assessment by Prognos-AG, p. 5 et. seqq. With the study, the National Association of Statutory Health Insurance Funds (GKV-Spitzenverband) complies with the legal requirements of § 65c Paragraph 10 SGB V to publish a report with the nationwide results of clinical cancer registration every five years starting 2018.

[5] S. XV.

[6] Regulation (EC) No 726/2004 of the European Parliament and of the Council dated 31 March 2004 laying down Community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency (relevant text for the EEC), Official Gazette L 136 dated 30 April 2004 p. 0001-0033.

[7] Regulation (EC) No 141/2000 of the European Parliament and of the Council dated 16 December 1999 on orphan medicinal products, Official Gazette L 018 dated 22 January 2000 p. 0001-0005.

[8] Article 1 No. 28a of Directive 2001/83/EC.

[9] Article 1 No. 28a of Directive 2001/83/EC contains the definition of the benefit risk ratio.

[10] Cf. e.g., Deticek A, Locatelli I & Kos M (2018). Patient access to medicines for rare diseases in European countries. Value Health, 21, 553-560. The results of the study show that Germany provides the best medical care both in terms of promptness of availability and expenses (calculated on the basis of the number of inhabitants).

[11] EUnetHTA. (2018). NetworkREQueST® tool and its Vision paper are now available for Public Consultation. Accessed on 22 May 2019 https://www.eunethta.eu/request-tool-and-its-vision-paper-are-nowavailabble--for-public-consultation/.

[12] During the negotiations on the law, the present suggestion of an extension of the transitional period was rejected.

[13] BSG, decision dated 19 October 2004 – B 1 KR 27/02 R –, BSGE 93, 236-252, SozR 4-2500 § 27 No 1.

[14] Eckert N (2019). „Orale Antikoagulation: Wenn Studienergebnisse und die reale Welt divergieren." (Oral anticoagulation: When study results and the real world diverge). Dtsch Arztebl 2019; 116(9): A-416 / B-344 / C-340. Accessed on 22 May 2019, https://www.aerzteblatt.de/archiv/205793/Orale-Antikoagulation-Wenn-Studienergebnisse-und-die-reale-Weltdivergieren.

Generation of additional evidence: Consensus-based rules are required

By Florian Staeck

If randomised clinical trials (RCT) are not available for the early benefit assessment of pharmaceuticals or if their validity shall be further evaluated, the question of suitable methods for the creation of contextual evidence arises and their role in the overall structure of early benefit assessment.

On the occasion of the 9th meeting of the „Interdisciplinary Platform on Benefit Assessment" on 15/16 March 2019 in Fulda, the participants discussed which instruments could be used to process data and made available for the determination of a potential additional benefit beyond RCTs. One goal was to learn from procedures other HTA authorities use to explore how they could be applied in Germany in addition to the AMNOG context. During the course of the meeting, it became apparent that this discussion has only just begun in Germany.

Participants discussed one available tool for the evaluation of confidence in effect estimates for study endpoints, i.e. GRADE methodology (Grading of Recommendations, Assessment, Development and Evaluation) which has already been conceptualised abroad for many years. The ultimate goal is to reflect the quality of evidence not only as a bias risk of individual studies, but to take a closer look at the precision and potential inconsistencies of the study results using a systematic approach. This approach employs unambiguous criteria for the evaluation of the quality of the available evidence, such as study design, bias risk, inconsistencies or strength of the effects. GRADE might help structuring the steps of evaluation thus creating the basis for a transparent decision-making.

The participants emphasised that this involves an enhanced operationalisation of evidence-based medicine and not a reduction of standards. Moreover, no new evaluation domains had been added during the past 10 to 15 years in the GRADE approach – which has gradually lead to an expert consensus as to which domains are crucial for the evaluation of the quality of evidence.

During the discussion whether GRADE could be used efficiently for German evaluation procedures, various positions were presented. On the one hand, participants emphasised that the GRADE approach was keenly observed in Germany and individual elements had already been adapted. On the other hand, they pointed out that questions like distributive justice or equal access to new pharmaceuticals or treatment methods were rarely or not at all addressed in the German evaluation context. Insofar, the GRADE approach could make a valuable methodological contribution.

Benefit risk analysis during approval

Not only during benefit assessment, but also during the approval of pharmaceuticals or vaccines, evaluators face certain insecurities in the assessment of evidence. A benefit risk analysis shall be conducted, regardless of the fact that the term „evidence" is not even used in the Medicinal Products Act (AMG). During the German approval process, a full data set about a new product is of particular importance. Authorities have an arsenal of regulatory measures to react appropriately in the event of insufficient evidence – e.g. by means of conditional approval subject to certain predefined conditions. Participants noted that the regulatory body had the explicit possibility of enforcing subsequent data submission – a problem where the Federal Joint Committee (G-BA) often failed in the past. They continued that although there was a clear preference for RCTs during approval, also single arm studies or historic comparisons would be used in exceptional cases to provide further evidence. As many products go through the approval procedure, experts from the competent authorities have a wide range of experience enabling them to make compari-

sons regarding the evaluation of evidence.

Different opinions were expressed as to whether the commitment of an assessor of the approval authority to a product beyond its life cycle might lead to conflicts of interest. This may in particular occur if a new product is accompanied by an assessor from an early stage of development. Other participants replied that decisions by approval authorities were always taken in the group. Besides the co-rapporteur, a peer review procedure was always performed as an additional security system so that conflicts of interest in the approval process are largely avoided.

Participants discussed the actual challenges of dealing with deviating evidence on the example of single arm studies. They explained that there was a broad scientific consensus that the application of non-adjusted indirect comparisons – e.g. by considering individual study arms of various studies – was not appropriate. But even by means of adjustment with a view to the study population it would never be possible to achieve the same effect as blinding in RCTs. However, various constellations were possible in which data from these types of studies couldn't be totally rejected in order to create contextual evidence, participants commented prudently. Nonetheless, the clinically relevant benefit in the context of early benefit assessment depends on the circumstances of the individual case – e.g. a high adverse event rate of the appropriate comparative treatment (ACT) or suitability of these data in relation to the ACT.

Based upon previous experience with early benefit assessments, this type of studies has been submitted quite often in connection with active ingredients against hepatitis C – and the G-BA disproportionately often derived an additional benefit from the additional data. However, from their perspective, such data sets were only rarely suitable for the indication oncology. Even under the best possible circumstances, only an indication for an additional benefit was derived from data sets of e.g. of single arm studies. Participants reported that the extent was rated as not quantifiable or too low.

Feasible instrument or waste of resources?

During the discussion, these results were reviewed and scepticism prevailed as to whether – except for cases with dramatic effects – investing money in these studies might be „programmed waste if resources" from the pharmaceutical company's perspective. On the other hand, equality of observation might be achieved with single arm studies, as endpoints were determined in the same way by different studies. Or by means of high-quality prospective data collection, evaluation of data could generally be improved. However, it depended on the circumstances of the individual case whether an overall evaluation of the additional benefit of a certain product can be improved.

The current lack of reliability of expectations of many stakeholders also shaped the discussion about the significance of contextual evidence that was examined using the examples of oncology and the treatment of type 2 diabetes patients. In October 2018, a paradigm shift has been initiated in diabetology with the new guidelines of the American Diabetes Association (ADA) and European Association for the Study of Diabetes (EASD) for the treatment of type 2 diabetes patients. Treatment no longer primarily depends on the adjustment of the glucose metabolism, e.g. based on the HbA1c level, but becomes more patient-centred.

More than ever, treatment shall focus more on the patient's cardiovascular situation and other clinical characteristics. Consequently, the perspective shifted from an „one size fits all" approach towards the consideration and treatment of micro and macrovascular complications. Thus, the

focus is placed on the reduction of the patient's morbidity and preservation of a high level of quality of life. Through this change of emphasis, the focus is more on a holistic approach than in the past.

This also has an impact on the design of clinical studies. Hence, the appeal was to place equal importance on all criteria of the 5th German Social Codebook (Sozialgesetzbuch V, SGB V), i.e. to not consider the patient's mortality, but also his morbidity and health-related quality of life. This would have to be operationalised and evaluated based on the endpoints. All too rarely, studies would address the question as to whether the patient's health-related problems could in fact be solved. Accordingly, the problem is that the increased effort for the consideration of e.g. retinopathy or neuropathy is only worthwhile in studies, if it is foreseeable how these can be handled during early benefit assessment.

As a consequence of this paradigm shift, the focus shifted from the individual disease to a risk intervention. For this purpose, instruments used in diabetology have to be more specific and sensitive. With a certain amount of scepticism, participants referred to the widely varying starting positions in diabetology and oncology. Although it was possible to collect data about the patient's morbidity and quality of life, participants argued that this would further increase complexity of studies.

Quality of life is increasingly often measured

Starting conditions in oncology are significantly different. Thus, the proportion of studies in which the quality of life was determined continuously increased from 60 in the past to close to 80 percent recently. Moreover, molecular diagnostics play a fundamental role in the creation of evidence in the indication oncology. They reported that companion diagnostics had a high significance in Germany. The promotion of the German Cancer Society in the development of regional diagnostics centres together with the respective referrer networks was very helpful.

In the majority of indications, it would be possible to use RCTs. Particular challenges would arise if individual cancer entities were biologically very heterogeneous making it sometimes difficult to clearly define the respective patient collective. Similar to diabetology, a „one size fits all" approach wouldn't be suitable here. That applied particularly, if treatment standards and biomarkers change during the course of the study. Against the background of these challenges, early benefit assessment of new pharmaceuticals would continue to be associated with high levels of data uncertainty.

On the example of the planned Act for Greater Safety in the Pharmaceutical Supply System (GSAV), participants discussed current topics of actual political design of data collection to create additional evidence. The draft of March 2019 stipulated that the G-BA should have the option to request post-market data collection or analyses for the purpose of benefit assessment for orphan drugs. According to the draft law, the G-BA should review the acquired data and conduction of data collection at least annually.

The proposed provision gave rise to a controversial discussion among the participants. On the one hand, they welcomed that the G-BA would have the possibility of requesting missing or insufficient evidence. On the other hand, they criticised that the legislator de facto requested non-interventional studies that weren't – at least up to now – of use for stakeholders during additional benefit assessment. So far, even after eight years of AMNOG it had not even been possible e.g. to derive a quantifiable additional benefit from registry data.

They also criticised that the proposed provision in the GSAV relating to potential deductions was an inconsistent

element within the AMNOG system. So far, this type of sanctions was only possible in distinct cases of non-compliance. Participants were particularly sceptical about the fact that the G-BA should already start assessing the results one year after the beginning of data collection. This period was much too short with respect to registry data.

Another argument was to see the legislator's signal that also other modes of evidence could be considered during the AMNOG procedure. Thus, post-marketing data collection by pharmaceutical companies should be seen as a chance instead of a threat. This was underlined by the remark that collecting data in the German treatment context during post-marketing data collection would constitute an intrinsic added value.

High standards for data collection

During the AMNOG procedure, the G-BA would only be obliged to consider subsequently collected data. By contrast, the manufacturer submitting the dossier would not have the right to derive a positive additional benefit assessment from the submitted data. Moreover, only small patient populations would be used in practice. Against this background, data could not be collected by means of inclusion models for patients which would still remain incomplete. The participants concluded that the new procedures of evidence creation shouldn't be implemented too fast in order to continue developing the AMNOG system as a learning system.

They also observed the first impulses emanating from the European Network for Health Technology Assessment's (EunetHTA) activities with interest. Since 2004, it was the aim of the EU partners to strengthen the cooperation on the evaluation of health technologies. They reported that in the past manufacturers showed only little interest to start a conversation with EunetHTA scientists. However, this was gradually changing. This development was also underlined by six further reports that are currently developed by the expert panel. The EunetHTA project was designed as a learning process offering the opportunity to further develop the initiative from an academic project to a kind of reference that might be used by national HTA authorities. The aim was to provide member states with a large pool of evidence data by means of a well-structured procedure of data collection at the EunetHTA.

ADVISORY BOARD OF THE INTERDISCIPLINARY PLATFORM ON BENEFIT ASSESSMENT

Dr Jürgen Bausch
Wolfgang van den Bergh
Dr Markus Frick
Dr Antje Haas
Dr Harald Herholz
Prof Dr Wolfgang Greiner
Dr Ulf Maywald
Dr Heinz Riederer
Prof Dr Jörg Ruof
Dr Sibylle Steiner
Dr Florian Staeck
Prof Dr Bernhard Wörmann

DISCUSSANTS

Dr Jürgen Bausch
Sabine Beckmann
Dr Antje Behring
Dr Johannes Bruns
Claus Burgardt
Prof Dr Klaus Cichutek
Prof Dr Eva-Susanne Dietrich
Dr Mathias Flume
Dr Markus Frick
Prof Dr Wolfgang Greiner
Dr Antje Haas
Michael Hennrich
Dr Harald Herholz
Philipp Huwe
Dr Thomas Kaiser
Dr Dietrich Knoerzer
Lothar Kuntz
Dr Wolfgang Lang-Heinrich
Martin Lack
Dr Eberhard Lüdtke
Dr Ulf Maywald
Dr Thomas Mittendorf
Prof Dr Dirk Müller-Wieland
Dr Heinz Riederer
Prof Dr Jörg Ruof
Prof Dr Dr Holger J. Schünemann
Dr Florian Staeck
Dr Giovanni Tafuri
Dr Katharina Thiele
Wolfgang van den Bergh
Dr Julia Wagle
Frank Wallbrecht
Prof Dr Jürgen Wasem
Jan Weißflog
Dr Elena Wolff-Holz
Prof Dr Bernhard Wörmann

IMPRINT

PUBLISHER
Editorial Advisory Board of the
Interdisciplinary Platform:
Dr Harald Herholz

Kassenärztliche Vereinigung Hessen
Europa-Allee 90
60486 Frankfurt
Germany

PUBLISHING COMPANY
Springer Medizin Verlag GmbH
Wolfgang van den Bergh,
Chief Editor
Am Forsthaus Gravenbruch 5
63263 Neu-Isenburg, Germany
German Commercial Register (HRB):
Berlin Charlottenburg District Court
HRB: 167094 B
VAT-ID: DE 230026696

Telephone: +49 6102 5060
E-mail address: info@springer.com

EDITORIAL WORK
Dr Florian Staeck
Wolfgang van den Bergh

AUTHORS
Prof Dr Klaus Chichutek
Michael Hennrich
Dr Thomas Kaiser
Prof Dr Dirk Müller-Wieland
Prof Dr Dr Holger J. Schünemann
Dr Florian Staeck
Dr Giovanni Tafuri
Prof Dr Bernhard Wörmann
Dr Elena Wolff-Holz

PICTURE CREDITS
Titel: sdecoret / stock.adobe.com

LAYOUT/GRAPHICS
Sandra Bahr
Oliver Hippmann

PRINT
F&W Druck- und Mediencenter GmbH
Holzhauser Feld 2, 83361 Kienberg,
Germany

Springer Medizin

© Springer Medizin Verlag GmbH,
Berlin, October 2019

In cooperation with and with the kind support of AbbVie Deutschland GmbH & Co. KG, DAK Gesundheit, MSD Sharp & Dohme GmbH, Novo Nordisk Pharma GmbH, Roche Pharma AG, Association of Research-Based Pharmaceutical Companies (vfa, Verband Forschender Arzneimittelhersteller e.V.), and Xcenda GmbH.

INTERDISCIPLINARY PLATFORM ON BENEFIT ASSESSMENT

Contextual evidence – Strategies for targeted therapy

Volume 9
October
2019