INTERDISCIPLINARY PLATFORM ON BENEFIT ASSESSMENT

Vol. 2
March 2016

Clinical Studies – which endpoints count?

Springer Medizin

Volume 2
March 2016

Contents

Goals of the platform

In 2011, the legislator initiated a paradigm shift in the field of pharmaceutical supply in Germany, with far-reaching consequences. The principle, based on the AMNOG, provides that: for new active substances brought on the German market, the pharmaceutical company must prove an additional patient-relevant benefit compared to the available standard of treatment – the appropriate comparative therapy (ACT) – if a higher reimbursement price is sought than for the ACT.

The additional benefit is evaluated and determined by the Federal Joint Committee (Gemeinsamer Bundesausschuss), generally on the basis of proposals from the IQWiG. The pricing is determined largely by the result of this additional benefit assessment. In Germany the price is for the first time negotiated between the National Association of Health insurance Funds and the pharmaceutical company.

The assessment of the additional benefit by the G-BA is the result of expert work based on a law (AMNOG) and on procedural and methodical regulations (e.g. IQWiG methods). 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 (Medizinischer Dienst der Krankenkassen, MDK) 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 drugs, likewise classified and interest-based, are submitted by the pharmaceutical companies to the G-BA, 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 drug approval,
- Working towards international standards of evidence-based medicine and of health economy being adhered to and applied,
- Determining whether and to what extent actual 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,
- Identifying 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.

The Interdisciplinary Platform would like to make a contribution to ensuring that new active substances are transparently and fairly assessed. The Advisory Council considers an interdisciplinary discussion regarding the results of the assessment and the applied benefit assessment methods to be essential. Furthermore, in the benefit assessment process it sees a good opportunity to inform the prescribing physicians of the expected additional benefits of new drugs for patients earlier than it was previously the case.

The interdisciplinary platform resulted from 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 Roche Pharma AG, DAK Gesundheit, Xcenda GmbH and Springer Medizin.

The Advisory Council of the Interdisciplinary Platform on Benefit Assessment

Early benefit assessment is suffering from an information deficit

By Dr. Pamela Aidelsburger and Dr. Jürgen Bausch

For new medications, both the market authorisation decision in the German summary of product characteristics (Fachinformation) and the package insert are readily available. In contrast, the decision of the Federal Joint Committee (Gemeinsamer Bundesausschuss, G-BA) on the additional benefit is not available as a summary form convenient for clinical practice. Health care practitioners are often unaware of the details of the G-BA decisions. Although five years have passed since the introduction of the AMNOG, the early benefit assessment is still surrounded by controversies regarding the value of surrogate parameters as valid estimators for patient relevant endpoints. Accordingly, one focus of this report is the acceptance and patient relevance of study endpoints.

The solving of patient problems by using drugs is the standard of care beyond surgical and psychotherapeutic medicine. You would have to be incredibly ignorant not to realise that general practioners, specialists in internal medicine and paediatricians would be rather helpless and unsuccessful without the host of effective drugs available. But not everything that has achieved market readiness and approval as a drug also has a proven clinically relevant additional benefit for patients.

However banal this statement may be, those involved in the provision of care have real difficulty in differentiating approval and additional benefit from one another, and in taking this into account when prescribing drugs. Especially the approval decision can be easily identified in the summary of product characteristics (Fachinformation) for the drug. However, the ruling of the G-BA regarding the benefit assessment is not always automatically to hand, as some sort of package insert, but rather must be found online and laboriously read through. There is no practical abridged version.

The effect of new drugs is studied within the context of the approval procedure, generally in randomised clinical studies. The methodical framework conditions for these approval studies, and in particular the endpoints to be investigated, are, as a rule, agreed with the approval authorities during the planning phase of the study. Here it is frequently the case that instead of endpoints, which are directly perceived by the patient (patient-relevant endpoints), surrogate parameters are ascertained as best estimated and accepted.

Surrogate parameters have several practical advantages. Thus, by measuring such parameters, the duration of a study can be shortened, if, for example, blood pressure is recorded as a risk factor rather than waiting for a myocar-

dial infarction in ten years time. Or if the improvement in lung function (FEV1 improvements in the case of Asthma and COPD as well as FVC in cases of pulmonary fibrosis) are deemed relevant to the decision. Likewise, the number of study participants can be reduced by considering surrogates. The pre-requisite for this is that there is a causal connection between the surrogate and the patient-relevant endpoint and that this has been proven to be valid. Unfortunately it has been shown in the past that presumed surrogate parameters are not always valid estimators for patient-relevant endpoints, although they are in some cases.

For these reasons, surrogate endpoints are accepted to varying degrees by the approval authorities and the HTA institutions, such as the IQWiG. Thus, it is normally the case that in randomised clinical studies, surrogates are investigated during the approval procedure and are recognised by the approval authorities, but are then rejected in Germany within the scope of the early benefit analysis. Furthermore, the validity of surrogates is in no way assessed in a standardised manner internationally.

Some of the controversially discussed points at the beginning of the early benefit assessment have lost their explosive effect over the course of five years. However, the acceptance and patient-relevance of study endpoints has not decreased in importance as an issue.

In particular, answers to the following questions must now be found:

- How can surrogate parameters be pragmatically and methodically validated? The requirements for the methodically flawless validation of a surrogate are very high. Strictly speaking, there must be a study in which both the surrogate and the patient-relevant endpoint have been recorded; and this applies to every active substance. For the early benefit assessment this means that – if the patient-relevant endpoint has already been measured – a validation is no longer necessary.
- At what time in the life cycle of a drug is the validation of a surrogate endpoint worthwhile? During the approval procedure or later during the benefit assessment?
- Who decides what is patient-relevant? Endpoints deemed patient-relevant are those that are perceived and experienced by the patient. But how are endpoints such as progression-free survival (PFS), which are measured using laboratory parameters or imaging findings, to be evaluated? By definition, these are surrogates, but only - a point that is denied - provided the patient's knowledge of progress can result in psychological effect and therapeutic consequences.

Alongside the discussion regarding surrogates, a second – very important – aspect must not be forgotten. How can different endpoints be weighted against one another? We know little to nothing with regard to how patients assess the benefit of a therapy in relation to possible side effects, or how different endpoints can be weighted against one another.

At its second meeting, the Interdisciplinary Platform on benefit assessment considered and discussed this topic from various perspectives. This report shows the desire to join together to identify and describe future paths. This also includes the knowledge that there is great uncertainty among doctors on a care provision level as to how the AMNOG functions and how the results of the early benefit assessment of drugs can be made as easily accessible as it is the case with the summary of product characteristics (Fachinformation) and the package insert.

Contact:

Springer Medizin Verlag GmbH

gp@springer.com

Clinically relevant study endpoints and surrogates during the early benefit assessment

Dr. Thomas Kaiser | Head of the Drug Assessment Department / IQWiG

The additional benefit of a new medicine is evaluated according to patient-relevant endpoints in the early benefit assessment (EBA). So far, no additional benefit was granted in approximately 60 per cent of assessments. In certain indications the required patient-relevant endpoints are rarely evaluated, leading to unsatisfying results in the EBA for the manufacturer. Additionally, such data are insufficient for physicians and patients. A change in perspective is needed, one that emphasises the identification of relevant questions first, rather than assuming an investigation topic is relevant. Surrogate endpoints can be used, if they allow conclusions on patient-relevant endpoints, but this is seldom the case. Availability of unpublished data is also necessary for validation studies of surrogate endpoints. The importance of the endpoint and the magnitude of the observed difference is assessed as part of the overall additional benefit evaluation.

In the view of the Institute for Quality and Efficiency in Healthcare (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, IQWiG) the early benefit assessment gives rise to three important questions on the topic of „relevant endpoints“:

- Were the „correct“ (patient-relevant) endpoints investigated?
- Are surrogate endpoints an alternative?
- How are the endpoints weighted in order to answer the following question: Does the new drug offer an overall additional benefit?

The scientific work of the IQWiG with a view to answering these questions is based primarily on the assessment of the data regarding the new active substance in comparison to the appropriate comparable therapy (ACT) , which must be provided by the company in a dossier.

During the approval, the following question is asked: Can the drug be used on humans on account of it doing more good than harm?, whereas in the benefit assessment the vital question is: Is there an additional patient-relevant therapeutic effect compared to the previous standard therapy?

In accordance with European law, the approval is almost always issued by the European Medicines Agency (EMA). The early benefit assessment is based on the German Drug Market Restructuring Act (Arzneimittelmarktneuordnungsgesetz, AMNOG), which is anchored in the 5th German Social Code book (Sozialgesetzbuch V, SGB V). The AMNOG is further supplemented by the Drug Benefit Assessment Ordinance (Arzneimittel-Nutzenbewertungsverordnung, AM-NutzenV) of the Federal Ministry for Health (Bundesministeriums für Gesundheit, BMG).

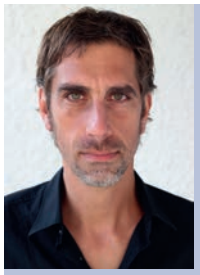
The benefit of a drug is defined as „the patient-relevant, therapeutic effect, in particular with regard to the improvement of the patient’s health status, the shortening of

the period of illness, the extension of the survival period, the alleviation of side effects or an improvement in quality of life. The additional benefit of a drug is a greater benefit than the appropriate comparable therapy" (Paragraph 2 Section 3 AM-NutzenV).

Were the correct, patient-relevant endpoints investigated?

The endpoints are selected on the basis of the statutory framework of the SGB V, supplemented by the Arzneimittel-Nutzenverordnung: mortality, morbidity, quality of life and side effects.

In the dossiers the manufacturers repeatedly postulate that an endpoint is relevant because it was investigated in the approval studies. However, on the basis of the mentioned statutory requirements, this is an erroneous conclusion: The mere fact that an endpoint has been investigated in studies no longer means that it is a relevant endpoint.



Dr. Thomas Kaiser *is a doctor and system developer. Following his time working as a programmer, he studied medicine in Cologne and worked in the field of internal medicine for a number of years. In 2002 he founded the Institute for Evidence-Based Medicine in Cologne. Since 2004, the year in which the IQWiG was founded, he has served there as the Head of the Drug Assessment Department, and since 2011 as the joint chairperson, alongside with Dr. Beate Wieseler.*

Selection of endpoints for the (early) benefit assessment

- Specifications of the SGB V / Social Code Book V
  - Mortality, morbidity, quality of life, side effects
- Investigated in studies ≠ relevant endpoint
- Symptoms and consequences of the disease → therapeutic goals → relevant endpoints
- Relevant endpoints → investigated in studies?

Figure 1: Process for the determination of patient-relevant endpoints.

Patient-relevant endpoints are derived from the disease and the associated symptoms and consequences, as well as the therapeutic objectives. Here the vital question is rather: „Were the relevant endpoints investigated in these studies?" (Figure 1).

What specific approach does the IQWiG take regarding the identification of relevant endpoints? First of all, literature regarding the disease is referred to, for it is essential that the disease is understood. The IQWiG also receives patient and expert surveys, in which information regarding patient-relative endpoints is sought out in a targeted manner. Furthermore, the manufacturer also has the opportunity to argue why a certain endpoint is relevant for patients in the documents he submits. Here it is important that the argument does not adopt the perspective that says „the endpoint was investigated, therefore it is also relevant", but rather that the endpoint fulfils patient-relevant therapeutic objectives within the therapeutic area. Finally, findings from previous projects are also important for the identification of relevant endpoints.

In the overview, therapeutic objectives can generally be deduced from the symptoms and consequences of a disease, from which patient-relevant endpoints can be de-

rived. If this data is missing, because it has not been investigated, then the conclusion can be that the additional benefit of the new drug has not been proven.

Assessment results

If you consider the results of the IQWiG assessments dated 22 September 2015, it is evident that no additional benefit was identified in any group for 57 percent of the drugs. Nevertheless, this also means that 43 percent revealed an additional benefit in at least one patient group. It is therefore fundamentally possible to investigate patient-relevant endpoints and to prove superiority over the appropriate comparable therapy (ACT).

Here the results differ greatly depending on the therapeutic area. In the field of oncology, around just one quarter of drugs had no additional benefit, while three quarters had an additional benefit (to varying extents).

Taking as an example Diabetes Mellitus Type 2, from the perspective of the IQWiG the endpoint problem is particularly clear: The main therapeutic objectives for this condition are the reduction of micro- and macrovascular complications, the alleviation of diabetes-related symptoms and the improvement of quality of life. None of the previous studies in the early benefit assessment were focussed on these therapeutic objectives (see Figure 2).

Approval studies for Diabetes Mellitus Type 2

The number of studies in early benefit assessments, which focus on the proof of a benefit with regard to the generally recognised therapeutic goals is:

0

Figure 2: The problem with regard to endpoint selection can be seen, for example, in the case of Diabetes Mellitus Type 2.

Because of the patients it is much to be hoped that the study currently published and not yet evaluated by the IQWiG/G-BA on Empagliflozin can for the first time prove a reduction in macrovascular events for this active substance (Zinman et al. 2015, EMPA-REG). It is to be assumed that, due to this study, Empagliflozin will soon be re-evaluated at the request of the manufacturer.

The proof of a reduction in macrovascular events was also the goal of a long-term study with Saxagliptin, published in 2013 (Scirica et al. 2013, SAVOR-TIMI 53). This study showed a reduction in cardiovascular events, but an increase in hospitalisation due to heart failure. The repeated assertion that, as a „safety study", the SAVOR-TIMI study was not intended to prove the superiority of Saxagliptin is false. This can be found in publicly accessible documents (publications, study registry).

Furthermore, the following should be noted: If the SAVOR-TIMI 53 had already begun when the Phase III studies on Saxagliptin were conducted (in 2006), the results would have already been available at the time of the approval (end of 2009) (Figure 3). The fact that it is also possible to begin endpoint studies very early in the case of diabetes mellitus type 2 is demonstrated in the current example of the EMPA-REG study, that had already commenced with Phase III (Figure 3). In addition to the fact that the available studies in the area of diabetes generally do not focus on patient-relative therapeutic objectives, it should be noted here that the results are frequently significantly influenced by the study design.

If, for example, you compare the gliptin „Linagliptin" directly with the sulfonylurea „glimepiride", then in principle this is an appropriate initial approach by which to verify an additional benefit. If, however, as it was the case in the corresponding direct comparative study, Linagliptin is administered in a fixed dose without a specific blood glucose

Endpoint studies in the case of Type 2 Diabetes - Example: EMPA-REG study

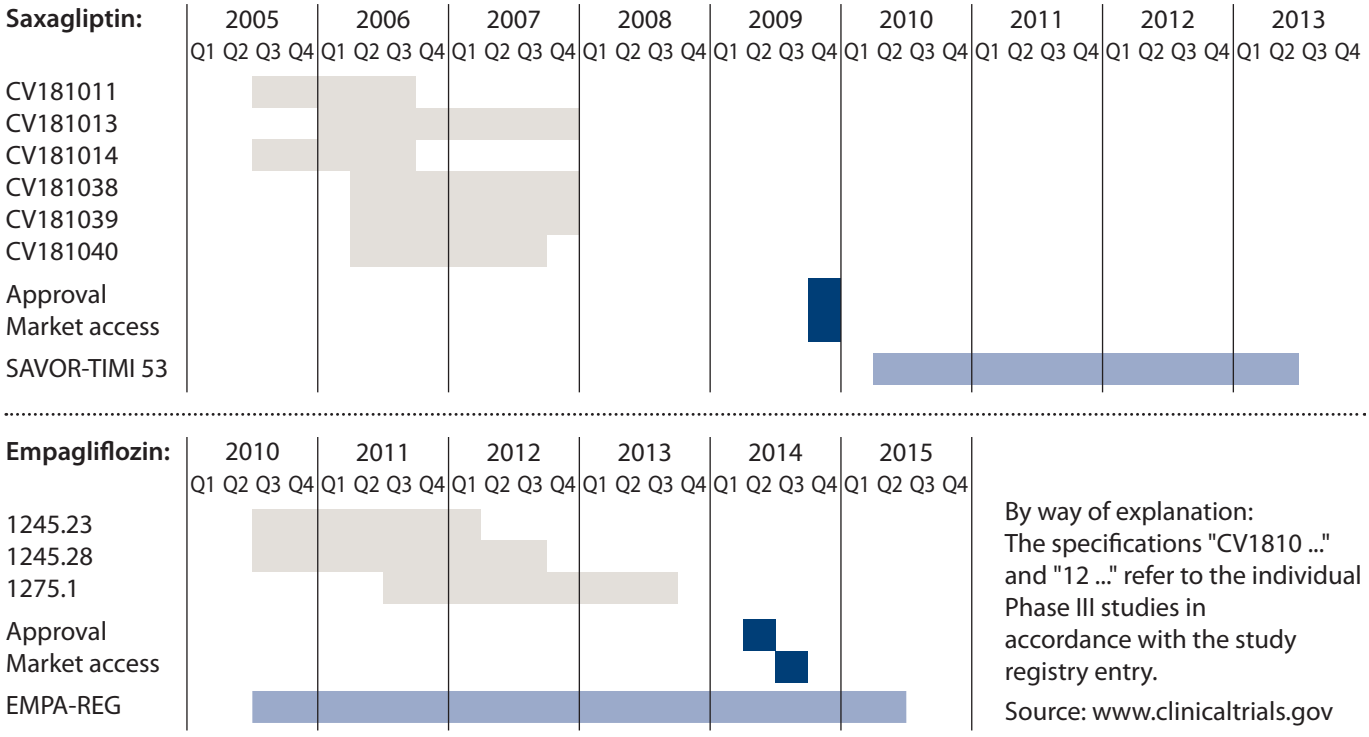


Figure 3: The EMPA-REG study clarifies that it is possible to commence endpoint studies at an early stage.

goal, but glimepiride is titrated with the goal of normoglycaemia (Gallwitz et al. 2012), this influences the interpretability of the study results considerably (IQWiG 2012). It is neither required in the approval, nor is it medically advisable or in accordance with medical guidelines to attempt to achieve very low blood glucose levels in all patients, regardless of the individual therapeutic objective and regardless, for example, of age and secondary diseases, by means of forced titration with glimepiride. As a result, the study design not only caused a higher rate of hypoglycaemia in the glimepiride arm in the weeks of the titration phase, but even resulted in more cerebral events under glimepiride during this time. Both are patient-relevant endpoints, however, due to the chosen study design no additional benefit of Linagliptin can be deduced.

Are surrogate endpoints an alternative?

The handling of surrogate endpoints is not simple. It is not just in Germany – and in the IQWiG – that the handling of surrogate endpoints is viewed as critical. Both in metaanalyses regarding the uncertain validity of surrogates in oncology and in the case of known, in some cases historic, examples (see Figure 4), it is evident how erroneously physicians can act if their decisions are based on the result of

surrogate endpoints. Thus, in the treatment of ventricular extrasystoles, in spite of a clear reduction in such rhythm disturbances, the antiarrhythmic flecainide resulted in an increased mortality in comparison to a placebo (Echt et al. 1991). Fluorides radiologically increase bone density, although in many studies the fracture rate was increased. The fact that hormone replacement therapy during menopause reduces cholesterol, yet the rate of heart attacks increases, was the finding of the Womens Health Initiative (Anderson et al. 2004). And in the case of diabetes mellitus type 2, a reduction of the surrogate HbA1c was observed with various anti-diabetics, without cardiovascular events having been reduced and in some cases with the rate of cardiovascular events even higher than in the comparison group (e.g. Scirica et al. 2013).

Thus, the pathophysiologic obvious conclusion that the normalisation of pathological values reduces the consequences of the disease, is not always correct. For this reason, caution is exercised when dealing with surrogate endpoints, including in the case of the early benefit assessment and in particular if relevant data is still lacking on a large scale. In the recent past, a surrogate endpoint caused a considerable stir in the case of early benefit assessment: Sustained Virological Response, SVR. This endpoint was also used by the IQWiG as a sufficiently valid surrogate for the reduction of the risk of hepatocellular carcinoma. The professional societies are in agreement that SVR is a valid surrogate endpoint for the reduction of the risk of cirrhosis and mortality, and last but not least also for patients with decompensated liver cirrhosis. However, with a view to the mortality risk in cases of decompensated liver cirrhosis itself, there is no proof of the validity of this surrogate. There are no corresponding investigations. Therefore it is essential to continually issue reminders regarding the validation of surrogate endpoints, but also to press ahead with it.

Here it is also essential to consider the evidence hierarchy, as published for oncology in Prasad et. al. 2015. In addition to the surrogate validation methodology being used, access to all data is required (including previously unpublished data). If you take into account that the more complete knowledge is applied, the more reliable the results are, then the stated considerations regarding evidence and hierarchy are all the more understandable.

Surrogate endpoints / Patient-relevant endpoints

- Anti-arrhythmic drugs – Ventricular extrasystoles ↓ – Mortality ↑
- Fluoride – Bone density ↑ – Fractures ↑
- Hormone replacement therapy – Cholesterol ↓ – Heart attacks ↑
- Several antidiabetic drugs - HbA_{1c} ↓ – Cardiovascular events ↑

Figure 4: Examples show how erroneously physicians can act if their decisions are based on the result of surrogate endpoints.

Information content of the sources

The experiences of the IQWiG show that the unpublished study reports have a much higher information content that the study registry reports or publications in scientific journals (see Figure 5). Therefore, including for the validation of surrogate endpoints, it should be encouraged that those who have access to the data (in the case of drugs this is generally the pharmaceutical companies), provide this data without exception. As in the case of the study evaluation itself, there is otherwise no way to ensure that the results are not considerably distorted as a result of targeted publication prioritising positive results.

Extent of the additional benefit/weighting

The weighting of the endpoints is by no means arbitrary, but is rather subject to statutory regulations specified by the AM-NutzenV. Here quantitative and qualitative aspects must be considered. The reasoning of the AM-NutzenV is: The more serious the event and the greater the difference observed, the greater the extent of the additional benefit.

This can be substantial, considerable or minor. However, classification as „not quantifiable" does not indicate benefit lesser than minor, but merely indicates that it is not known in which of the three above categories the benefit should be classified, for methodical or any other reasons. Other categories are „no additional benefit" and as the poorest category „lesser benefit than the ACT".

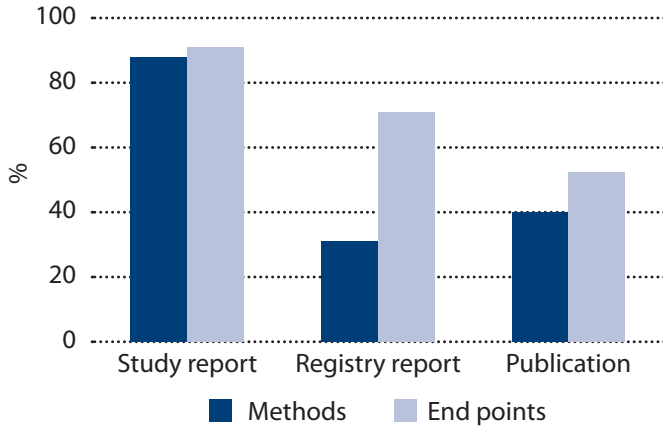
Without going into the methodology of the IQWiG in detail, quality of life should be taken into consideration in this context. Although this is a patient-relevant endpoint according to AM-NutzenV, quality of life is not addressed in the provisions for determining extent in AM-NutzenV. On the other hand, in its methodology the IQWiG equates changes in quality of life to a change in severe symptoms (IQWiG 2015).

The question as to when a side effect or symptom is actually serious and when it is non-serious is sometime simple and sometimes difficult to answer (see Figure 6). In the case of undesired events there are standard classifications by which to determine the level of severity (for example the generally recognised and applied definition of serious undesired events, SUE). In oncology in particular, the classification according to CTCAE is also used (Common Terminology Criteria for Adverse Events). In the event of therapy discontinuations, the IQWiG always asks: Which events are behind this? Was it serious undesired events that lead to the discontinuation or was it non-serious events?

IQWiG's goal is always to make an allocation that is backed up by data. In case of doubt, the allocation must be expressly justified in the evaluation.

Within the written reporting procedure and the obligatory hearings in the G-BA, corrections by manufacturers and other parties submitting comments are possible, especially with regard to the weighting and classification into assessment categories. And, as a look at the now 140 assessment procedures shows, they are also reality.

Information content of various document types



Source: Wieseler B, Kerekes MF, Vervoelgyi V, McGauran N, Kaiser T. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. BMJ 2012;344:d8141

Figure 5: Study reports generally have a higher information content than other document types.

How is ignorance weighted?

At the time of the benefit assessment, we at least know something about the endpoints that have been investigated, but we do not know always everything. However, there is no way that we can know anything about relevant endpoints that have not (yet) been investigated. How is this ignorance considered when weighing up the addition-

Serious or non-serious event? Goal: data-supported allocation

- Undesired events (UE)
  - Classification of the level of severity (in articular SUE, CTCAE-level)
  - Therapy discontinuations: caused by serious UE?
- Symptoms / complications
  - In studies, partial classification of the end point itself as minor, moderate or severe
  - Using the descriptions in the survey (e.g. low impact on daily life vs. severe impact caused by symptoms)

In case of doubt, the allocation must be expressly justified in the evaluation.

Figure 6: In the case of UEs there are standard classifications by which to determine the level of severity.

al benefit? Here are two examples:

- If there is no reliable data regarding the presumed principle effect (beneficial) of a drug, it makes little sense to deduce an additional benefit solely due to rarely occurring side effects.
- Essentially, all parties are in agreement that quality of life is frequently of high priority in oncology. Nonetheless, in the past the lack of such data has only resulted in the devaluation of the additional benefit in exceptional cases.

If important data is lacking, there is the legal possibility of imposing a deadline with the goal of answering questions still pending, so as to then be able to draw a provisional and then a definitive conclusion (Figure 7). Until now this option has been availed of in around one in five assessments. We will now look briefly at the extent to which the requirements formulated by the G-BA have actually been implemented by the pharmaceutical companies. There are a host of deadlines approaching in the coming

How is ignorance weighted?

Drug Benefit Assessment Ordinance § 5

If valid data on patient-relevant endpoints is not yet available at the time of the assessment, the assessment shall be performed on the basis of the available evidence, taking into account the study quality, stating the probability of the proof of an additional benefit, and **a deadline can be defined by which time valid data on patient-relevant endpoints must be submitted.**

Figure 7: The legal option to impose a deadline has been availed of in around one in five evaluations.

months, including two relating to orphan drugs. In 2016 a number of anti-diabetics are due to be reassessed. How the G-BA makes decisions when the data it requires is still not available, is one of many unanswered questions.

For further discussion

The goal of all those involved should be to further develop early benefit assessment, as such that it represents the best possible basis for decision-making, not only for price negotiations, but first and foremost for doctors and patients. In relation to the topic of this meeting, the following topic areas in particular should be addressed:

- How do we progress from: „What is done is right" to: „The right thing is done"?
- How can we use the existing, but not always publicly available information from clinical studies to promote surrogate validation and, if the case may be, to determine that a surrogate is invalid?
- How should lacking knowledge regarding relevant endpoints be considered in future when weighing up additional benefits?

Summary

From the perspective of the IQWiG it is evident in many examples of the early benefit assessment process in accordance with the provisions of the German Act on the Restructuring of the Medicines Market (Arzneimittelmarktneuordnungsgesetz, AMNOG) that the correct and necessary patient-relevant endpoints have not been investigated at all. This inevitably leads to unsatisfactory assessment results from the manufacturer's perspective, but also in an insufficient data situation for doctors and patients. A different mindset is required here: What makes something relevant is not the fact that it is being studied, but rather it is the relevant matters that should be studied.

Surrogate endpoints can be used in individual cases if data on patient-relevant endpoints is lacking. However, the pre-requisite is that these surrogate endpoints enable sufficiently reliable assertions regarding effects on patient-relevant endpoints. The opposite is frequently the case, or there is a lack of scientific studies that prove this link (validation studies). For conducting surrogate validation studies, it is essential that all, even unpublished, data be made available.

In the overall weighting regarding additional benefits, the relevance of the endpoint (for example the severity of the event) and the extent of the observed difference must be considered. The extent to which lacking knowledge flows into the overall weighting regarding additional benefits should be the subject of further discussion.

References

Anderson GL, Limacher M, Assaf AR, Bassford T, Beresford SA, Black H et al. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. JAMA 2004; 291: 1701-1712.

Federal Ministry of Health. Ordinance on the benefit assessment of drugs in accordance with § 35a Paragraph 1 SGB V for reimbursement agreements pursuant to § 130b SGB V (Drug Benefit Assessment Ordinance - AM-NutzenV) [online]. 27.03.2014 [Access: 23 November 2015]. URL: http://www.gesetze-im-internet.de/bundesrecht/am-nutzenv/gesamt.pdf.

Act on the restructuring of the medicines market in the statutory medical insurance sector (Arzneimittelmarktneuordnungsgesetz - AMNOG). Federal Law Gazette (Bundesgesetzblatt) Part 1 2010: 2262-2277.

Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. N Engl J Med 1991; 324: 781-788.

Gallwitz B, Rosenstock J, Rauch T, Bhattacharya S, Patel S, by Eynatten M et al. 2-years efficacy and safety of linagliptin compared with glimepiride in patients with type 2 diabetes inadequately controlled on metformin: a randomised, double-blind, non-inferiority trial. Lancet 2012, 380: 475-483.

Institute for Quality and Efficiency in Healthcare. Linagliptin: renewed benefit assessment according to § 35a Paragraph 5b SGB V; Dossier assessment; Order A12-11 [online]. 29.11.2012 [Access: 23 November 2015]. (IQWiG reports; Vol. 144). URL: https://www.iqwig.de/download/A12-11_Linagliptin_Erneute_Nutzenbewertung_35a_Absatz_5b_SGB_V.pdf.

Institute for Quality and Efficiency in Healthcare. General methods: Version 4.2. Cologne: IQWiG; 2015. URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4-2.pdf.

Prasad V, Kim C, Burotto M, Vandross A. The Strength of Association Between Surrogate Endpoints and Survival in Oncology: A Systematic Review of Trial-Level Meta-analyses. JAMA Intern Med 2015; 175: 1389-1398.

Scirica BM, Bhatt DL, Braunwald E, Steg PG, Davidson J, Hirshberg B et al. Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus. N Engl J Med 2013; 369: 1317-1326.

Wieseler B, Kerekes MF, Vervoelgyi V, McGauran N, Kaiser T. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. BMJ 2012; 344: d8141.

Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S et al. Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. N Engl J Med 2015; 373: 2117-2128.

This article is based on a transcript of the original lecture by T. Kaiser. The transcript was drafted by Springer Medizin Verlag GmbH and converted into article form by Dr. Jürgen Bausch. The manuscript draft was then thoroughly revised by T. Kaiser. The figures were selected by T. Kaiser and match the respective slides used in the lecture.

Orphan drugs: Justified differences regarding study designs and endpoints?

By PD Dr. Michael Kulig | Federal Joint Committee

Orphan drugs are subject to reduced requirements in the early benefit assessment by the German health technology authorities. By law, market authorisation is considered proof of their additional benefit. Although EU regulations entitle patients with rare diseases to medicines of the same quality, safety and efficacy as other patients, there is a high proportion of non-comparative and non-blinded studies for orphan drugs. Specially adapted study designs for orphan drugs do not exist, and the same assessment criteria apply to orphan and non-orphan drugs. Similarly, the evaluation of data in orphan drug dossiers submitted to the G-BA faces the same underlying limitations as non-orphan drugs, particularly regarding the collection of data on patient relevant endpoints.

In late 1999, the European Union specified special regulations for drugs for rare conditions (orphan drugs) in a statutory ordinance [1]. Accordingly, a manufacturer can receive the status „Orphan Designation" if there was previously no satisfactory treatment option and a condition with a prevalence of no more than 5 in 10,000 people presents itself. By definition this still means a population of up to around 230,000 patients per rare condition in Europe. These regulations grant certain advantages with regard to drug approval and market access.

According to German Social Code Book V, due to the legal binding of early benefit assessment to approval, lighter requirements apply in comparison to drugs that do not have orphan drug status (non-orphan drugs). The additional benefit is already deemed to have been proven as a result of the approval, meaning that evidence of the medical benefit and the medical additional benefit in relation to the appropriate comparable therapy does not have to be submitted (Table 1).

Because the pharmaceutical company applies to the European Medicines Agency (EMA) for the orphan designation, not every drug for the treatment of rare conditions necessarily falls under the special regulation for orphan drugs. This can result in a drug with and another drug without orphan drug status being approved for comparable indications. In the case of drugs being marketed in Germany, benefit assessments and rulings with the G-BA may then take place using varying procedural regulations, as is evident, for example, with the orphan drug ibrutinib and the non-orphan drug idelalisib for the indication chronic lymphatic leukaemia (CLL).

Because the additional benefit in the case of drugs for rare diseases are deemed to have been proven, no probability for the additional benefits and no appropriate com-

Orphan drug regulations

Criteria of the European approval authority, the European Medicines Agency (EMA), for an orphan drug designation:

- Drug for the diagnosis, prevention or treatment of a condition, which is life threatening or results in chronic invalidity
- At the time of application there are no more than five in ten thousand people affected with the [European] community
- No satisfactory methods for the diagnosis, prevention or treatment of the respective disease have yet been approved or the respective drug — provided such a method exists — will be of considerable benefit to those affected by the condition
- Incentives for orphan designation: companies receive market exclusivity for ten years, fees are waived or reduced, the processing of the approval application is expedited

AMNOG regulations since 2011:

- Medical additional benefits are deemed to have been proven upon approval ("fabricated" additional benefits for approved orphan drug without an assessment – in accordance with the procedural code of the G-BA)
- Proof of the medical additional benefit *in relation to the appropriate comparable therapy* is not required to be submitted
- Restrictions do not apply if the turnover exceeds 50 million Euros in the last 12 months; then an unrestricted benefit assessment takes place in accordance with the procedural code of the G -BA

Table 1: Orphan drug status goes hand in hand with specific characteristics in the early benefit assessment.

Privat-Dozent Dr. Michael Kulig *is a doctor and Master of Public Health, and since 2009 has been Head of the Medical Consultancy Department at the G-BA. In addition to serving as an expert for international journals and in ethics commissions, he previously worked for the IQWiG, at Sanofi Pasteur MSD in France and at the Charité Hospital in Berlin.*

parable therapy are determined (Table 1).

According to the Ordinance of the European Parliament on medicines for rare conditions, it is stipulated that „patients with rare conditions must have the same right to good treatment as other patients" and the „patients with such conditions have the same claim regarding the quality, safety and effectiveness of drugs as other patients" (extract from the second and seventh recital of the European Ordinance [EC] 141/2000). These statements pose the questions as to whether or not these specific characteristics justify a separate assessment of orphan drugs or whether medicines for rare conditions should not therefore be subject to exactly the same, normal assessment procedure as non-orphan drugs.

Example for an equal indication for an orphan drug as well as a non-orphan drug

	Ibrutinib (Orphan)	**Idelalisib (Non-orphan)**
Approved area of application: Chronic lymphatic leukaemia (CLL)	A) Patients that have received at least one previous treatment B) As first line treatment in the case of 17p-deletion or TP53-mutation in patients that are not suitable for chemoimmunotherapy	in combination with Rituximab A) Patients that have received at least one previous treatment B) As first line treatment in the case of 17p-deletion or TP53-mutation in patients that are not suitable for chemoimmunotherapy
Patient numbers	N = 184	N = 220
Study design	RCT, active controls (**Ofatumumab**), open	RCT, active controls (**Rituximab**), double-blind
Primary end point **quality of life**	PFS Ascertained	PFS Not ascertained
Extent of the additional benefit **(probability of the additional benefit)**	A) Patients that have received at least one *previous treatment:* **Not quantifiable** *B) As first line treatment in the case of 17p-deletion or TP53-mutation in patients who are not suitable for chemoimmunotherapy:* **Not quantifiable**	A) Patients with recurrent CLL, for whom chemotherapy is not suitable: **Not quantifiable (reference point)** [ACT: Best-Supportive-Care] Patients with recurrent CLL, for whom chemotherapy is indicated: **Not documented** [ACT: chemotherapy in combination with Rituximab] Patients with refractory CLL, for whom chemotherapy or treatment with Ofatumumab is indicated: **Not documented** [ACT: patient-specific, optimised treatment] Patients with refractory CLL, for whom chemotherapy or treatment with Ofatumumab is not indicated: **Not documented** [zVT: Best-Supportive-Care] B) First line treatment of chronic lymphatic leukaemia (CLL) with 17p-deletion or a TP53-mutation not suitable for chemotherapy: **Not quantifiable (reference point)** [ACT: Best-Supportive-Care]
Specific characteristics		Limited in time until 1 April 2016

Table 2: Differing procedural rules and benefit assessments for an orphan drug and an non-orphan drug.

In a study of the underlying evidence from 44 drugs for rare conditions newly approved under the EU Orphan Drug Regulations prior to 2007, the authors essentially found that only in 57 percent of cases were the approval studies randomised, controlled studies (RCT), in 23 percent non-controlled Phase II study, and in one case a retrospective study, while in three cases the approval was granted on the basis of a literature review [2]. In addition, the authors criticise that the study duration was in some cases too short when compared to the natural course of the disease, the primary endpoints were mainly surrogate endpoints and in case of oncological indications relevant data regarding overall survival or quality of life was lacking.

Has this insufficient data situation, which resulted in the approval of orphan drugs, changed in recent years? In a more recent analysis regarding the approval of 188 new substances with 448 studies between 2005 and 2012, based on approval by the US Food and Drug Administration (FDA), there were relevant differences in the study design between orphan and non-orphan drugs [3]. The 56 pivotal studies on orphan drugs were randomised in just 53 percent of cases and blind in 37 percent, in contrast to 94 percent and 86 percent respectively in the 392 studies on non-orphan drugs (Figure 1).

In almost half of the orphan drug cases, approval was granted on the basis of a non-comparative study, while just eight percent of the pivotal studies on non-orphan drugs had no comparators. So too in the case of the surrogate endpoints and the resulting questionable patient-relevance of the endpoints was the proportion higher for orphan drugs (73 percent versus 45 percent). As expected, the medi-an number of patients involved in the studies on orphan drugs (N=150) was lower than for non-orphan drugs (N=480). A study that describes the characteristics of approval studies for cancer drugs between 2004 and 2010

Orphan versus non-orphan drugs

FDA approval procedure (2005 - 2012)

- **Design differences**
  In 448 pivotal studies for 188 new substances
- **In the case of orphan drugs**
  - Lower proportion of randomised studies
  - Lower proportion of blinding

Orphan status	randomised	double-blind
Yes (n = 56)	30 (53.6) [40.0–67.0]	21 (37.5) [24.4–50.6]
No (n = 392)	370 (94.4) [92.1–96.7]	335 (85.5) [82.0–89.0]
p-value	<0.001	<0.001

[1]Downing NS et al: Clinical Trial Evidence Supporting FDA Approval of Novel Therapeutic Agents, 2005–2012. JAMA. 2014;311(4):368–377.

Figure 1: Analysis shows differences in the study design depending on the orphan designation.

at the FDA produced similar results [4]. In spite of the lower patient numbers, studies for orphan drugs with meaningful results do appear possible, even with these low case numbers (Figure 2).

Significant limitations, such as a lack of randomisation, blinding or a control group, which were identified in the previous investigation of approval studies prior to 2007 of 44 orphan drugs in the EU were also evident in the subsequent analyses of the pivotal studies, which resulted in FDA approval up until 2012. On the basis of which evidence was the G-BA able to support the early benefit assessments of orphan drugs performed since AMNOG? Do the mentioned limitations and other factors make the as-

Orphan versus non-orphan drugs

FDA approval procedure (2005 – 2012; 188 Novel agents[1])

- **In the case of orphan drugs**
  - Lower proportion of comparable studies
  - Lower proportion of clinical (patient-relevant) end points

No. (%) [95%-CI]

	Comparator			Evidence base		
	Active	**Placebo**	**None**	**End points**	**Clinical end points**	**Clinical scales**
Orphan-Status						
Yes (n = 56)	12 (21.4) [10.3–32.5]	16 (28.6) [16.4–40.8]	28 (50.0) [36.5–63.5]	41 (73.2) [61.2–85.2]	10 (17.9) [7.5–28.2]	5 (8.9) [1.2–16.6]
No (n = 392)	131 (33.4) [28.7–38.1]	231 (58.9) [54.0–63.8]	30 (7.7) [5.0–10.3]	178 (45.4) [40.5–50.4]	120 (30.6) [26.0–35.2]	94 (24.0) [19.7–28.2]
p-value		<0.001			<0.001	

[1]Downing NS et al: Clinical Trial Evidence Supporting FDA Approval of Novel Therapeutic Agents, 2005–2012. JAMA. 2014;311(4):368–377.

Figure 2: Analysis of the clinical evidence for the FDA approval of new medicines.

sessment of data regarding the additional benefit more difficult? Of the 30 dossiers of orphan drugs assessed since 2011, seven (23 percent) were based on non-comparative studies and 23 (77 percent) on RCTs. Of these RCTs 17 were blind and had a „formal" low bias risk, which corresponds to a proportion of 57 percent of all dossiers. However the certainty of the results of patient-relevant endpoints in many of these RCTs was so limited, as a result of limitations such as restricted validity of the survey tools and the clinical relevance thresholds in patient-reported endpoints, low response rates, disparate study arms or the insufficiently long recording of endpoints (for example beyond the progression of the disease), that no reliable results could be deduced for an assessment of the additional benefit. In cases with an insufficient scientific data basis or a high level of uncertainty regarding the results, it is often impossible to quantify the additional benefit.

Such limitations make the reliable estimation of the additional benefit of a newly approved drug difficult in the early benefit assessment. These difficulties are not limited to Germany or the G-BA and its procedures.

For example, in November 2013 the English National Institute for Health and Care Excellence (NICE) did not issue a recommendation for the orphan drug bosutinib (for chronic myeloid leukaemia). The G-BA had previously been unable to quantify the additional benefit due to an insufficient data basis regarding the patient-relevant endpoints and the lack of a control group and blinding (Figure 3).

Low or non-quantifiable additional benefits or time limitations

Due to unreliable evidence:

- **Poor data conditions**
  - Data from the literature/case reports or very low case numbers
    - E.g. Alipogentiparvovec in the case of lipoprotein lipase deficiency (LPLD) (N=8 or N=17 [May 2015])
    - E.g. Bosutinib
      (in case of CML) → G-BA: not quantifiable [October 2013]
      → NICE: no recommendation for Bosutinib [Nov 2013]
- **Patient-relevant end points / Quality of life (QoL)**
  - Not always ascertained
  - Differing estimations regarding validity and patient-relevance
  - Lacking data → study organisation, logistics[1]
- **Lacking long-term data**

[1]EMA – Guideline on Clinical Trials in Small Populations (2006)

Figure 3: Rulings of the G-BA and NICE in the case of an orphan drug compared to CML.

In view of the described studies and the identified limitations, it appears rather surprising that the feasibility of RCTs and comparative studies for the assessment of the medical benefit of orphan drugs is repeatedly called into question. As long as a therapeutic effect is not sufficiently evaluated and guaranteed, a study without a comparison group is difficult to justify.

In order to facilitate the conducting of a study with rare diseases, so-called adaptive designs are generally mentioned. However, on account of this, an adaptation and which adjustments should improve the feasibility of studies to test the effectiveness of orphan drugs, are in most cases not accurately specified. As a result, such adjustments lead to more complex designs. This often negatively affects the feasibility.

This can – in order to make decisions regarding the adjustment – increase time and resource expenditure and, by means of multiplicity adjustment also results in an unreliable evidence base and unreliable data interpretation, compared to with non-adaptive designs. These specific characteristics of the adaptive design apply equally to orphan and non-orphan drugs. Thus, the EMA also explained in its Reflection Paper on adaptive design that „adaptive designs should not be seen as a means to alleviate the burden of rigorous planning of clinical trials. Instead, adaptive designs would be best utilised as a tool for planning clinical trials in areas where it is necessary to cope with difficult experimental situations".

With regard to the case numbers, which could render study planning and implementation difficult in the case of very rare conditions, no advantages are to be expected here as such adjustments based on an interim analysis do not, per se, aim to reduce the number of cases, but rather to monitor type I error. In the FDA Draft Guidance for adaptive designs, it is stated that due to such interim analyses, decisions should not be made in favour of an increase in case numbers, but rather in favour of a reduction in case numbers [5].

In the field of oncology, for example, a two-stage adapted design was proposed in order to test a certain targeted active substance with various types of cancer or patient populations [2]. This design, which the authors intended for Phase II studies, following the inclusion and analysis of a pre-defined number of treated patients, enables decisions to be made with regard to whether further patients should be included for the respective indication or whether the active substance is considered ineffective

for the respective indication and the study recruitment for this population should be stopped [7]. However, this study design can only be viewed as a „signal-generating study", especially taking into account that these studies are generally planned to be one-armed.

Likewise, in its report in which it evaluates the studies pursuant to the orphan drugs with EU approval from 2001 until Jan 2004, the IQWiG comes to the conclusion that „no scientific justification can be deduced for a different approach in the assessment of medical interventions for rare and non-rare diseases". The IQWiG also concludes that „conversely […] no specific designs and statistical methods [exist] that could also be relevant for (more) frequent diseases" [8].

Conclusion: In order to improve the often insufficient data basis and to increase the certainty of results in studies on orphan drugs, improvements and efforts are required both with regard to the validity of patient-reported endpoints and the underlying surveying tools, and in particular with regard to study logistics, in order to achieve sufficiently high response rates and long observation times. As such, these limiting factors do not differ fundamentally in the case of non-orphan drugs.

References

((1) European Parliament, Council of the European Union Ordinance (EC) No. 141/2000 of the European Parliament and Council of 16 December 1999 on drugs for rare conditions. OJEC 2000; http://www.eur-lex.europa.eu/LexUriServ/LexUriServ.do? uri=OJ:L:2000:018:0001:0005:de:PDF

(2) Joppi R, Bertele V, Garattini S: Orphan drug development is not taking off. Br J Clin Pharmacol 2009:67; 494-502.

(3) Downing NS et al: Clinical Trial Evidence Supporting FDA Approval of Novel Therapeutic Agents, 2005-2012. JAMA. 2014;311:368-377.

(4) Kesselheim AS, Myers JA, Avorn J.: Characteristics of clinical trials to support approval of orphan vs. non-orphan drugs for cancer. JAMA. 2011;305(22):2320-2326.

(5) US Food and Drug Administration (FDA) – Draft Guidance – Guidance for Industry Adaptive Design Clinical Trials for Drugs and Biologics (February 2010)

(6) Lin Y et Shih WJ: Adaptive Two-Stage Designs for Single-Arm Phase IIA Cancer Clinical Trials. Biometrics 2004,60,482-490.

(7) Hyman DM, Puzanov I, Subbiah V, et al. Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations. N Engl J Med. 2015;373:726-36.

(8) IQWiG Report – No. 241 of 05.09.2014: „Assessment and evaluation of studies regarding rare conditions".

Acceptance and weighting of endpoints in drug therapy

Prof. Dr. Bernhard Wörmann | DGHO Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie (German Society for Haematology and Medical Oncology)

More than 50 % of newly approved drugs were assigned „no additional benefit" in the German early benefit assessment according to the AMNOG process. The discrepancies are due to the different assignments of approval and health technology assessment, to differences in the definition of patient-relevant endpoints and to a lack of relevant data for HTA assessment in earlier trials. We need a common data platform for newly designed pivotal trials.

Introduction

The effectiveness and benefit of a new drug are recorded in relation to patient-relevant endpoints. In daily practice this takes place individually with the specific patients. The definition of standardised endpoints is required for studies. They determine the design of the study, the measurement parameters to be recorded, the required number of study participants, termination criteria and finally, the significance of a new drug to the future care of patients. The considerable heterogeneity of the clinical pictures and therapeutic indications results in a wide range of possible and reasonable endpoints. These differ in the various indications and specialist areas. Patient-relevant endpoints from oncology are shown in Figure 1.

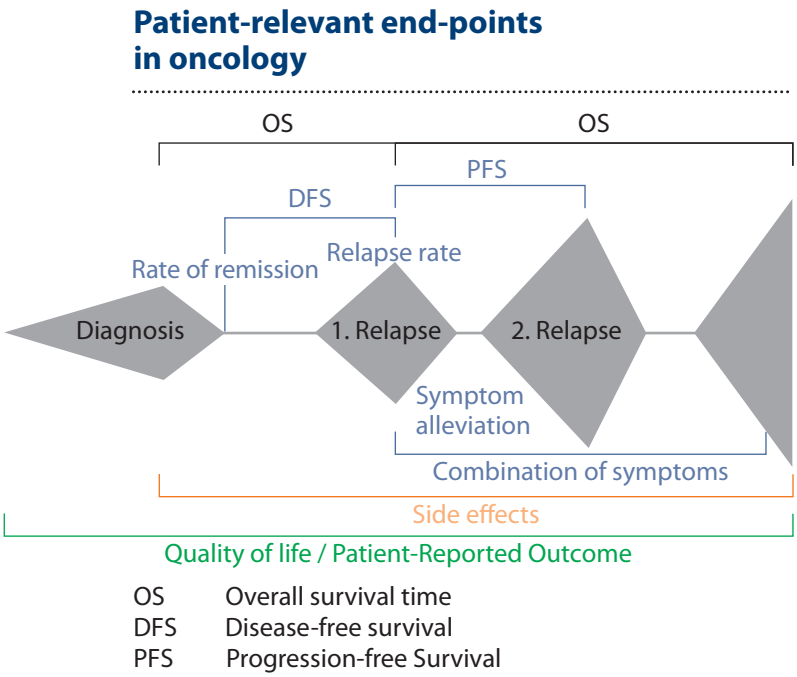


Figure 1: The heterogeneity of the clinical pictures and therapeutic indications require a host of possible endpoints.

In terms of content, the endpoints can be classified into these four categories:

- Mortality
- Morbidity
- Quality of life
- Side effects

Experience in recent years in Germany has shown that the analysis of the same study(/ies) from the perspective of approval, benefit assessment and the drafting of guidelines can result in divergent outcomes. The summarised publication of these decisions in various media causes irritation and uncertainty in the prescribing doctor and the affected patient, rather than providing qualified support when making a decision. Below, approaches and results of approval and benefit assessment procedures are first presented briefly.

Prof. Dr. Bernhard Wörmann *is a doctor specialising in internal medicine, haematology and internal oncology, and also works in palliative care.*
Since 2010 he has been Medical Director of DGHO Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie (German Society for Haematology and Medical Oncology). Since 2011 he has been employed as a doctor in the Outpatient Health Centre on the Virchow Campus of the Charité hospital in Berlin.

End points of the clinical trials for approval (EMA) - Benefit assessments 2011 to 2014

Category	End point		Specification of the end point
Mortality	Overall survival		
Morbidity	Disease-free survival		
	Progression-free survival		
	Response	Clinical symptoms	
		Imaging	
		Lung function	
		Blood pressure	
		Lab findings	HbA_{ic}[1]
			Viral load[2]
			LDL-C[3]
			Blood count
			Cytogenetics
			Others
	Relapse		
	Avoidance of diseases, complications, following surgical intervention	Attack rate (epilepsy)	
		Thromboembolic complications	
		Operations, Rejection	
Quality of life	Quality of life / Patient-Reported Outcome		
Side effects			
Diagnostics	Quality		Consistency with medical finding after imaging

[1] Glycated haemoglobin, [2] HIV, Hepatitis,
[3] Low Density Lipoprotein Cholesterin

Table 1: Endpoints of approval studies

Endpoints for approval

The European Medicines Agency (EMA) and the Federal Institute for Drugs and Medical Products (Bundesinstitut für Arzneimittel und Medizinprodukte, BfArM) are responsible for the approval of all new drugs in accordance with Directive 2001/83/EC, Art. 21, Paragraph 4. The decisive criteria for approval are quality, effectiveness and safety. Table 1 lists the endpoints from approval studies on new drugs, which were then analysed from 2011 until 2014 within the scope of early benefit assessment in Germany.

In the area of the measurement of morbidity in particular the EMA accepts a wide range of different endpoints. This also includes surrogates such as the virus load or HbA1c laboratory parameters. They are used in the approval studies as a replacement for another endpoint that is deemed to be of higher quality, if this cannot be identified as such during the course of the study.

Benefit assessment endpoints

The Drug Benefit Assessment Ordinance, Paragraph 2 Section 3 (AM-NutzenV of 28. 12. 2012, BGBl. I P. 2324) defines benefit as:

„The benefit of a drug in the sense of this ordinance is the patient-relevant effect, in particular with regard to the improvement of the patient's health status, the shortening of the period of illness, the extension of the survival period, the alleviation of side effects or an improvement in quality of life."

The results of an analysis conducted by scientific medical associations of all procedures completed by the end of 2014 are shown in Figure 2. For numerous drugs the benefit assessment is performed according to sub-groups, which are established by the G-BA prior to the assessment procedure commencing. Because the conclusive determination of the additional benefit is realised by the G-BA in these subgroups, the analysis follows this procedure.

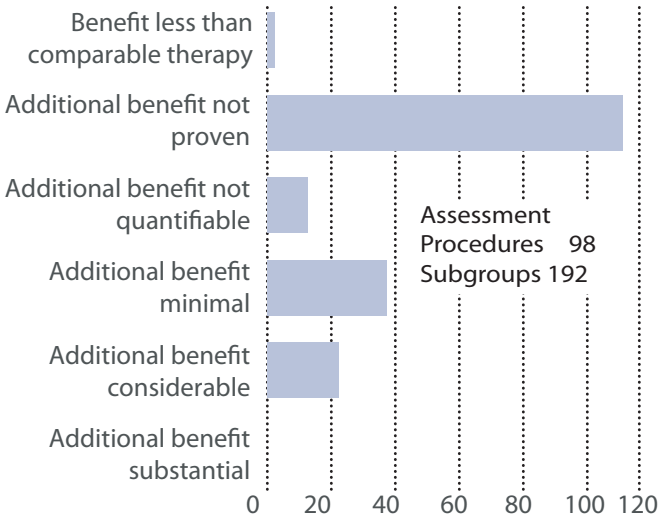


Figure 2: Results of an analysis conducted scientific specialist medical associations of all procedures completed by the end of 2014.

The figures show that more than half of the drugs approved by the EMA doe not fulfil the criteria of the G-BA for the determination of an additional benefit. The results of the benefit assessment differ greatly in the respective specialist areas. Whereas in oncology and infectiology the majority of subgroups were deemed to have an additional benefit, this is rarely the case for anti-diabetics and drugs for neurological conditions. The results from oncology are shown separately in Figure 3.

A significant justification for the relatively better assessments of oncological drugs is the weighting of the endpoints. The Institute for Quality and Efficiency in Health Care in Cologne (IQWiG) has been commissioned by the G-BA to draft a report for all regular procedures, i.e. for

Results of the early benefit assessment of new oncological drugs 2011-8/2015

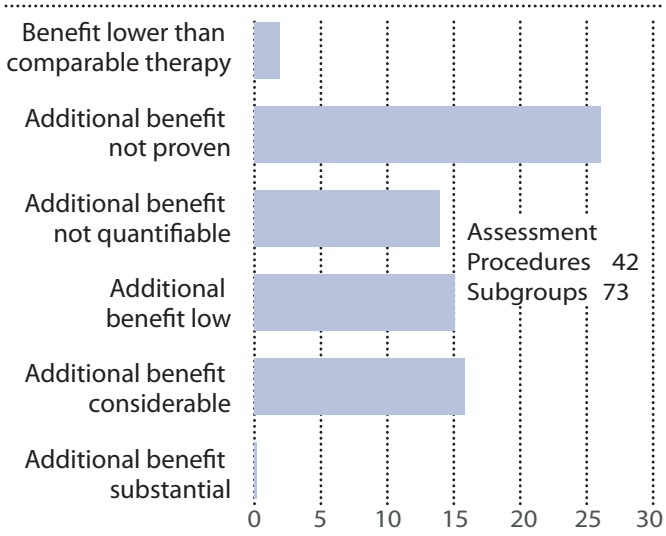


Figure 3: Results of the early benefit assessment for oncology products.

drugs without orphan drug status. Within the scope of the benefit assessment for Ticagrelor it developed methods for the operationalisation of the determination of the extent of the additional benefit, assumptions and threshold values, and then published them in its method papers [2], see Table 2.

In this methodology the overall survival is assigned a greater value than serious symptoms. In medicine there are many indications in which this assumption is in line with the priorities of the patient. However, in palliative situations at the end of a patient's life, and not only in the case of cancer patients, the alleviation of symptoms and an improvement in quality of life can have a greater significance than a short-term extension of the survival time. In situations such as this, this methodology is not appropriate. The relative preference towards drugs for indications in which a larger group of patients die from the disease during the course of the approval study is also proble-

Operationalisation of the determination of the extent of the additional benefit offered by the IQWiG

Extent of the additional benefit	**Overall mortality**		**Serious or severe complications and side-effects, health-related quality of life**		**Non-serious (or non-severe) symptoms or complications and side-effects**	
	Effect on relative risk	Confidence interval threshold value	Effect on relative risk	Confidence interval threshold value	Effect on relative risk	Confidence interval threshold value
Substantial	0,53 - 0,58	0,85	0,24 - 0,38		N/A	not specified
Considerable	0,84 - 0,85	0,95	0,69 - 0,71	0,75	0,34 - 0,48	0,80
Low	N/A	Every statistically significant extension of the survival time	N/A	0,90 1,00	0,69 - 0,71	0,90

Table 2: Operationalisation methods that the IQWiG has published in its method papers.

Patient-relevant endpoints in oncology and their possible assessment

	Relevance				
Parameter	none	low	moderate	high	very high
Mortality					
Cancer-specific mortality					
Disease-free survival					
Progression-free survival (PFS)					
PFS + symptoms					
PFS + change of therapy					
Remission rate + symptoms					
Relapse rate					
Symptoms					
Time until the occurrence of encumbering symptoms					
Avoidance of disease/symptoms					
Patient-reported outcome / quality of life					
Side effects					

Figure 4: Trial of the drafting of a common matrix for all data to be recorded for oncological drugs.

matic. This does confirm the public need for new drugs for these indications, however it is of disadvantage to the development of drugs for early stages of an illness, which is equally worthwhile from a health policy perspective.

Standardised database for approval and benefit assessment

For responsible parties and for participants in approval studies, it is highly unsatisfactory if the established endpoints and the measurement criteria are subsequently classified as non-patient-relative in a benefit assessment. The drafting of a shared matrix for all recorded data appears to me to be possible. An approach for oncological drugs is summarised in Figure 4.

Depending on the indication, the endpoints can be assessed differently. There is also an opportunity for an approach for the scientific recording of the priority of endpoints from the patients´ perspective.

References

1. Wörmann B et al.: Early benefit assessment of new drugs in Germany, 2014. https://www.dgho.de/informationen/gesundheitspolitische-schriftenreihe/band-6
2. IQWiG: General Methods 4.0, 2011. https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf

Study planning within the approval and benefit assessment procedure

Prof. Dr. med. Dietmar P. Berger, Genentech inc., San Francisco | Dr. Miriam Walter, nspm ltd., Meggen

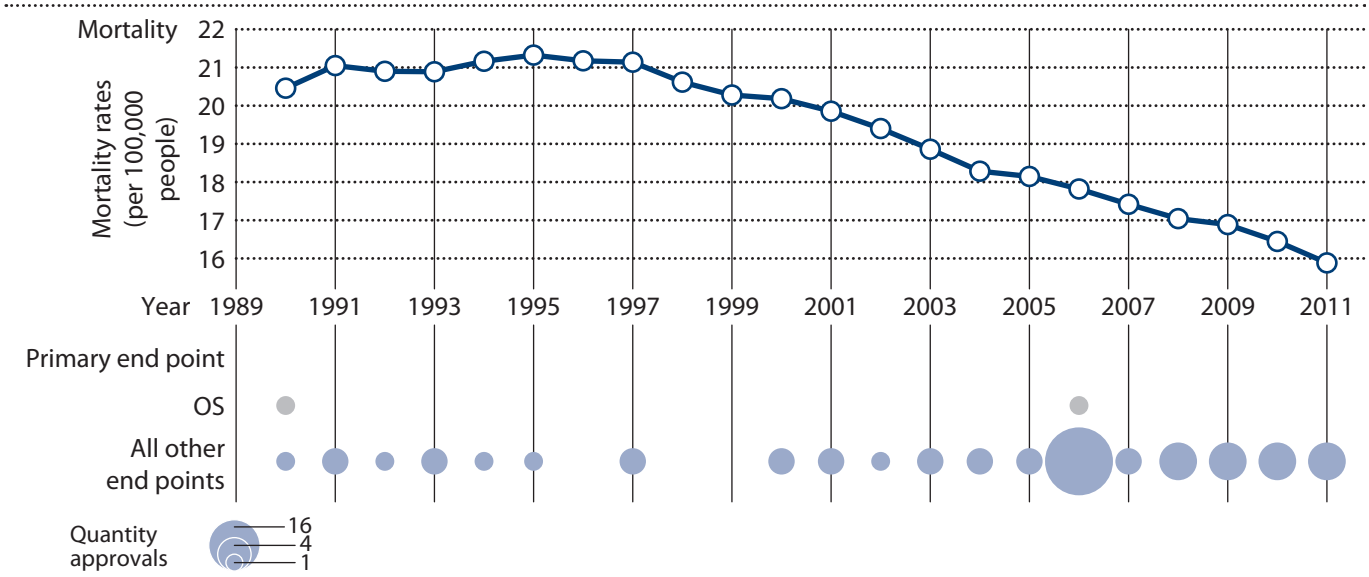
Last years' medical progress in the field of oncology, mainly due to improved therapeutic strategies and targeted molecular genetic profiling, leads to an impeded assessment of the endpoint category mortality. Licensing authorities recognise this hurdle and are gradually accepting patient-relevant morbidity endpoints when mortality data is limited or unavailable. The Federal Joint Committee also focusses evidence standards on mortality but, despite flexible licensing requirements, does not accept clinically well-established morbidity endpoints, e.g. progression-free survival. Although quality of life takes a high priority for patients, it plays a minor role in the licensing and benefit assessment processes. Harmonisation of evidence standards between licensing and benefit assessment is essential.

Five years ago, in January 2011, the Act on the restructuring of the medicines market (AMNOG) came into force in Germany. The new law requires a benefit assessment of drugs in comparison to the appropriate comparable therapy, in which both the Health Technology Assessment institutions (HTA institutions), namely the Federal Joint Committee (Gemeinsamer Bundesausschuss, G-BA) and the Institute for Quality and Efficiency in Health Care (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, IQWiG) are involved. The G-BA, commissioned by the legislator, makes a decision regarding the additional benefit of a medicine and generally commissions the IQWiG with the scientific evaluation of the available evidence and the drafting of a report in advance. To date around 190 assessment procedures have commenced, with around 150 completed (as of October 2015) (G-BA 2015).

The assessment of the additional benefit by the G-BA is based on patient-relevant endpoints of the categories which are mortality, morbidity and quality of life (QoL) (G-BA 2014). In clinical studies, various endpoints are recorded for the respective categories depending on the indication and study design, such as overall survival (OS) for mortality, progression-free survival (PFS) for morbidity and patient reported outcomes (PROs), such as sensitivity to pain, for QoL. A central critical point is the estimation of patient-relevance of the study endpoints used, since the G-BA only considers data from endpoints deemed to be patient-relevant.

In order to keep up with medical progress, innovative study designs and endpoints are required, which are not always in line with the requirements of the authorities. These developments are outlined below, with a focus on oncology and the respective perspectives of the approval authorities and the HTA institutions regarding endpoints

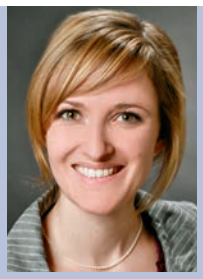
Decreasing mortality in haematology



Sources: 1) National Cancer Institute. Surveillance, Epidemiology, and End Results Program (SEER); URL: http://seer.cancer.gov; 2) Food and Drug Administration (FDA); URL: www.fda.gov; 3) Centerwatch. URL: http://www.centerwatch.com.

Figure 1: Reduction of the mortality rate in haematology between 1989 and 2011.

Prof. Dr. Dietmar P. Berger *is Senior Vice-President and Head of Product Development for Clinical Haematology and Oncology at Roche Pharma GmbH. In this position he manages the global strategic alignment of the company's oncological development portfolio. Among other positions, Berger was Head of the Clinical Research Centre at the Freiburg University Hospital and received the German Cancer Prize for work in the field of angiogenesis.*



Dr. Miriam Walter *is employed at nspm ltd. in Meggen, Switzerland. In addition to several years of professional experience in the field of pharmacies, she has diverse experience in the area of regulatory and medical writing, the drafting of medical information documents and the development of training materials. Another task area is the analysis of early benefit assessment in Germany.*

are addressed, with a view to ultimately submitting a proposal for the merging of these different views.

Therapeutic progress leads to the specification of patient populations

Through targeted prevention strategies, new therapeutic approaches and improved support therapies, pharmaceutical development has managed to achieve remarkable results. For example, the mortality rate in the area of haematology of over 20 per 100,000 patients in 1989 was reduced to 16 in 2011. Here the approval of new drugs introduced during this period was only based on the OS in two cases (see Figure 1) (Centerwatch, Food and Drug Administration [FDA], National Cancer Institute). So too in oncology, therapeutic progress could be achieved: between 1975 and 1977 the five-year survival rate across all surveyed cancer types was 49 percent, while between 2003 and 2009 it was 68 percent (American Association for Cancer Research [AACR] 2014). From 2002 to 2011 the overall mortality rate for oncological conditions in the USA fell by 1.8 percent in men, 1.4 percent in women and 2.1 percent for those aged under 20 (American Association for Cancer Research [AACR] 2015).

Clinical development in the field of oncology in recent years has been based on, among other things, the results of fundamental molecular genetic research. Through the characterisation of the mutation sub-types of a tumour, treatments that are more specific, and therefore more effective for patients, could be developed. For example, in 2004 two oncogenes, Epidermal Growth Factor Receptor (EGFR) and Kirsten rat sarcoma viral oncogene homolog (KRAS), were identified for bronchial adenocarcinoma. Nowadays it is assumed that there are ten oncogenes that represent potential therapeutic targets (Johnson 2013, Pao and Girard 2011).

In addition, biomarkers produced by cancer cells (for example Programmed Death-Ligand 1 [PDL1]) could be identified, which on the one hand occasionally overlap with the known oncogenes, and on the other hand identify patients for whom immunotherapeutic treatment may be particularly beneficial (Rolfo 2014, Sgambato 2015). It is to be expected that in the future molecular genetic profiling will be part of the standard diagnostic repertoire, with a view to supporting therapeutic decisions.

According to the new status of research, innovative approaches with regard to endpoints and study design are being pursued, especially in oncology. New clinical endpoints include, for example, minimal residual disease (MRD) in case of lymphoma or pathologic complete response (PCR) in the neoadjuvant treatment of breast cancer. So-called umbrella studies represent an innovative study design. In the umbrella study design, molecular genetic diagnosis is recorded within a histopathological indication, and the therapy is selected on this basis (American Association for Cancer Research [AACR] 2014).

The so-called basket study design, on the other hand, recruits patients on the basis of their molecular genetic characterisation and not, as it is classically the case, based on histopathological features (American Association for Cancer Research [AACR] 2014). For example, various melanoma cancer types feature the BRAFV600 mutation.

A basket study with the BRAFV600 inhibitor Zelboraf involved patients with metastasised solid tumours or multiple melanomas, exhibiting a BRAFV600 mutation. The various tumours were divided into eight cohorts, including non small-cell lung cancer, ovarian cancer and colorectal carcinoma. In the study, the response rate of the cohort of non small-cell lung carcinoma lay at 42 percent (95 percent confidence interval: 20, 67), which positively distinguishes itself from the response rate of seven percent, as it was re-

ported for the standard two-line therapy Docetaxel in molecular non-selected patients.

Patients with Erdheim Chester Disease or Langerhans cell histiocytosis, two related rare conditions without an approved therapy for adults, exhibited a response rate of 43 percent (95 percent confidence interval: 18, 71), which indicates a clinical significance of the effect of BRAF inhibitors in these conditions (Hyman 2015).

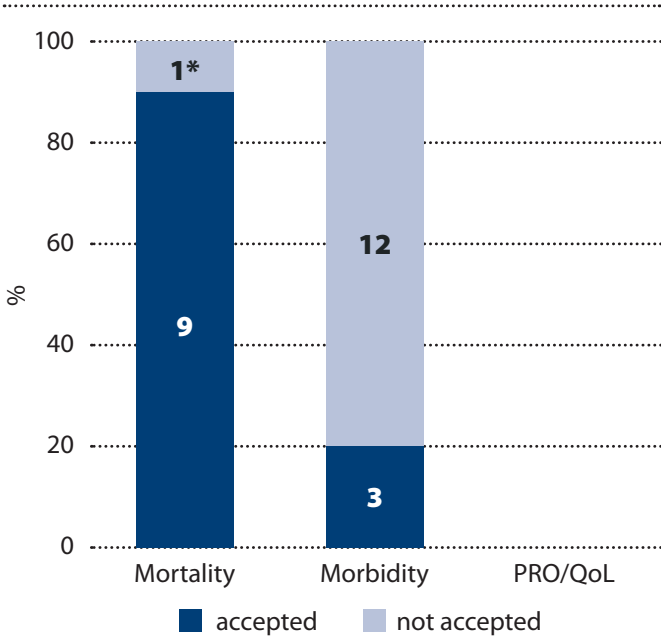
As a result of the major progress in oncological therapy, patient survival chances are improving considerably. In clinical studies this results in a decrease in mortality events. The focus on molecular genetic characteristics of disease entities also reduces the number of patients in study populations. Both developments make it difficult to prove a significant mortality advantage in new oncological drugs, especially for innovative, highly effective therapies.

Low acceptance of morbidity endpoints of oncological studies by the German HTA institutions

An analysis of the acceptance of primary study endpoints showed that primary mortality endpoints such as OS are generally accepted by the G-BA. The case of primary morbidity endpoints was a different matter, as shown in Figure 2. These are accepted to a lesser extent by the G-BA. Endpoints based on symptoms (such as pain), are viewed as patient-relevant by the G-BA. Other well-established clinical (primary or secondary) morbidity endpoints, such as PFS or response rate are, on the other hand, frequently overlooked in the benefit assessment. At the time of the analysis there was no data on primary PRO/QoL endpoints (Ruof 2014). Generally speaking, tools for the measurement of PROs, such as EORTC QLQ-C30, are accepted by the G-BA, however, they contribute less to the determination of the additional benefit.

This procedure of G-BA is clarified by taking Obinutuzumab as an example. In addition to the mortality endpoint OS, the summary of product characteristics (Fachinformation) for Obinutuzumab also provides data on six different morbidity endpoints, which are presented individually in Table 1. None of these morbidity endpoints were accepted by the G-BA in its benefit assessment for Obinutuzumab (G-BA 2015, Roche Pharma AG 2015). As already shown, therapeutic advantages of a drug cannot, however, be determined solely by means of mortality advantages in all cases. Therefore, a high level of non-acceptance of morbid-

Acceptance of primary study endpoints by the G-BA



* File incomplete
G-BA: Gemeinsamer Bundesausschuss (Federal Joint Committee); PRO: patient-reported outcome; QoL: Quality of life.
Source: Ruof J et al. (2014). Health Policy 118:242–54.

Figure 2: Comparison of the acceptance of endpoints for mortality, morbidity and patient-reported outcome.

ity endpoints is detrimental to innovative therapies in particular, and makes their development difficult.

Approval versus benefit assessment: a different approach to evidence-based medicine

When weighing up the risk-benefit profile of new drugs which leads to approval decisions, the approval authorities, and in particular the European Medicines Agency (EMA) and the Food and Drug Administration (FDA), refer primarily to mortality and morbidity endpoints, and not so much to PRO data. In particular, the FDA views OS as the gold standard, as this endpoint is free from bias caused by estimations made by the investigator. However, the authorities recognise that high patient numbers and a long follow-up are required to record OS. In addition, OS can be influenced by follow-up therapies. Accordingly, just two of 66 haematological approvals issued by the FDA between 1989 and 2011 are based on OS. PFS is an endpoint that is not biased by follow-up therapies, however, there is a risk of bias caused by the investigator.

An alternative is the response rate, which is viewed by the FDA as a direct measure of anti-tumour activity, although this does not take the progression-free time into consideration. PROs have previously not been the basis of an FDA approval and are viewed as methodologically immature (Centerwatch, Pazdur 2008). The EMA views OS as

G-BA endpoint acceptance, taking Obinutuzumab as an example

	Result: German summary of product characteristics (Fachinformation)[1]	Acceptance of the endpoints by the G-BA[2]
PFS (INV)	HR 0.39 (95% CI 0.31; 0.49)	×
PFS (IRC)	HR 0.42 (95% CI 0.33; 0.54)	×
Response rate	65%/78%*; p = 0.0001	×
MRD	3%/26%*; p = 0.0001	×
Event-free survival	HR 0.43 (95% CI 0.34; 0.54)	×
Time before another anti-leukaemic treatment	HR 0.59 (95% CI 0.42; 0.82)	×
OS	HR 0.66 (95% CI 0.41; 1.06)	Yes, but vs. Rtx+Clb not significant

* Rtx+Clb/Obinutuzumab; Clb: Chlorambuzil; G-BA: Gemeinsamer Bundesausschuss (Joint Federal Committee); HR: hazard ratio; INV: assessed by investigator; IRC: independent review committee; CI: Confidence interval; MRD: minimal residual disease); OS: overall survival; Rtx: Rituximab.

Sources: 1) Roche Pharma AG (2015). Gazyvaro® German summary of product characteristics (Fachinformation); 2) G-BA (2015). Obinutuzumab benefit evaluation procedure URL: https://www.g-ba.de/informationen/nutzenbewertung/131.

Table 1: Data on morbidity endpoints in the summary of product characteristics (Fachinformation) and acceptance in the G-BA ruling.

the endpoint with the highest informative value, but is aware of the fact that OS is not applicable in all situations. According to the EMA, PFS is viewed as ‚beneficial' for patients, however, a high measurement frequency can place a burden on the patient. According to the EMA the cure rate is a key therapeutic objective during the (neo-)adjuvant disease stage, however, it can only be recorded using surrogate endpoints (European Medicines Agency 2013).

Thus, mortality endpoints such as OS are accepted as the preferable endpoint by the approval authorities. Morbidity endpoints are generally accepted if they are established and clinically relevant. Endpoints used in oncology and accepted by the approval authorities are, for example, the cure rate, time until progression, PFS, relapse-free survival, response rate (clinical response rate, CRR) or symptoms. Endpoints for PROs are seldom represented in the approval process (see Figure 3). The approval authorities focus on the evaluation of the body of the available data, there is room for extrapolation and a stepwise approval, so-called adaptive licensing, is possible. In addition, approval authorities usually recognise primary endpoints as patient-relevant and their evidence requirements are based on the disease.

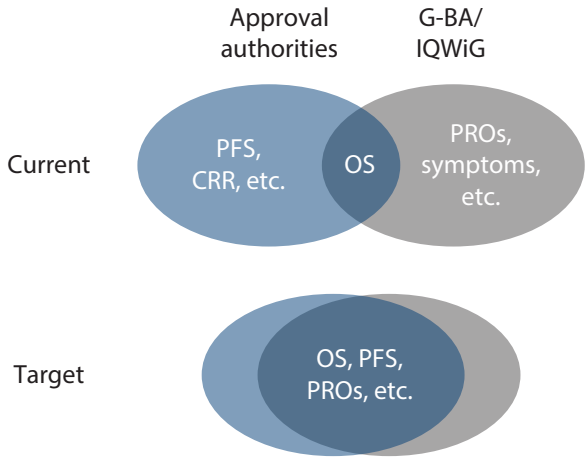
As already mentioned, mortality and, if they also have an impact on the benefit assessment, QoL endpoints are accepted by the G-BA. In contrast, even established morbidity endpoints such as PFS are generally not accepted. The focus of the G-BA for the assessment of the additional benefit is placed on the evidence provided by randomised clinical studies in comparison to the appropriate comparable therapy. An extrapolation of the data is rarely accepted and adaptive approaches are barely provided for within the scope of the AMNOG. The acceptance of primary endpoints as patient-relevant is, with the exception of OS, questionable within the context of the benefit assessment procedure. Furthermore, German HTA institutions tend to apply rigid evidence standards.

Proposed solution for the standardisation of approval and benefit assessment

In order to reduce the discrepancy between the requirements of the approval authorities and the German HTA institutions, various approaches are feasible. An acceptance of surrogate endpoints, such as MRD, facilitates the compilation of evidence, in particular in the case of less aggressive disease entities, as well as with specific, and therefore small, patient populations.

In addition, the HTA institutions should take into ac-

Acceptance of endpoints by approval authorities and German HTA institutions



CRR: clinical response rate; OS: overall survival; PFS: progression-free survival; PRO: patient reported outcomes.
Source: own diagram

Figure 3: Morbidity endpoints such as PFS are at present largely not accepted by the G-BA.

count a flexible adjustment of the acceptance criteria for endpoints, using the clinical picture and status. Patients with insufficient response to previous therapies or aggressive disease entities, such as small-cell bronchial carcinoma, acute myeloid leukaemia and HER2+ breast cancer in third-line therapy generally have a much reduced life expectancy, meaning that the overall survival definitely represents a reasonable and important endpoint. In the case of therapies in previous lines, in case of chronic diseases or if many treatment options are available, mortality is not only less relevant, but also much more difficult to measure.

Examples of such diseases in oncology are (neo-)adjuvance in cases of breast-, colon- or bronchial-cancer, chronic lymphocytic leukaemia or basal cell carcinoma. With such indications, morbidity endpoints should be given high priority, even if there is only indirect patient-relevance, for example if the endpoints are not based on the ascertainment of symptoms. Regardless of the significance of mortality and morbidity endpoints relative to one another, PROs should be measured as additional endpoints and should be considered in the benefit assessment.

Conclusion

The current framework conditions result in differences in the study requirements of approval authorities and HTA institutions, especially with regard to morbidity endpoints. In particular, smaller, more specific patient populations and a much improved life expectancy in many oncological fields, result in an intensification of an area of tension. While approval authorities consider the (temporal) feasibility of endpoints under certain circumstances, German HTA institutions require an extensive dataset with a strong focus on OS and a reduced acceptance of clinically accepted morbidity endpoints. This raises the question as to whether the evidence requirements of the G-BA, which tend to be rigid and far-reaching, oppose the procedures of the approval authorities, which are becoming more flexible on account of research progress.

The goal should be to minimise the area of tension in terms of acceptance of endpoints when it comes to drug development, approval and benefit assessment, and thus to provide sufficient room for innovative therapies. The determination of patient-relevant endpoints at an early stage, in collaboration between approval authorities, HTA institutions and pharmaceutical companies, taking into account the disease intensity and the stage of the illness (see Figure 3) enables the standardisation of various evidence requirements and paves the way for innovative therapeutic approaches.

References

American Association for Cancer Research [AACR] (2014). Cancer Progress Report; URL: http://cancerprogressreport.org/2014/Documents/AACR_CPR_2014.pdf [last accessed on 8 October 2015].

American Association for Cancer Research [AACR] (2015). Cancer Progress Report; URL: http://cancerprogressreport.org/2015/Documents/AACR_CPR2015.pdf [last accessed on 8 October 2015].

Centerwatch. URL: http://www.centerwatch.com [last accessed on 8 October 2015].

European Medicines Agency (2013). Guideline on the evaluation of anticancer medicinal products in man; URL: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/01/WC500137128.pdf [last accessed on 15 October 2015].

Food and Drug Administration [FDA]. URL: www.fda.gov [last accessed on 8 October 2015].

G-BA (2014). Rules of procedure; URL: https://www.g-ba.de/downloads/62-492-1002/VerfO_2014-12-18_iK-2015-04-16.pdf [last accessed on 06 October 2015].

G-BA (2015). Overview of active substances; URL: http://www.g-ba.de/informationen/nutzenbewertung/ [last accessed on 6 October 2015].

Hyman DM, Puzanov I, Subbiah V, Faris JE, Chau I, Blay JY, Wolf J, Raje NS, Diamond EL, Hollebecque A, Gervais R, Elez-Fernandez ME, Italiano A, Hofheinz RD, Hidalgo M, Chan E, Schuler M, Lasserre SF, Makrutzki M, Sirzen F, Veronese ML, Tabernero J,Baselga J (2015). Vemurafenib in Multiple

Nonmelanoma Cancers with BRAF V600 Mutations. N Engl J Med 373:726-36.

Johnson BE, Kris MG, Berry LD, Kwiatkowski DJ, Iafrate AJ, Varella-Garcia M, Wistuba II, Franklin WA, Ladanyi M, Su P-F, Sequist LV, Khuri FR, Garon EB, Pao W, Rudin CM, Schiller JH, Haura EB, Giaccone G, Minna JD,Bunn PA. A multi-centre effort to identify driver mutations and employ targeted therapy in patients with lung adenocarcinomas: The Lung Cancer Mutation Consortium (LCMC). ASCO Annual Meeting; 2013.

National Cancer Institute. Surveillance, Epidemiology, and End Results Program (SEER); URL: http://seer.cancer.gov [last accessed on 8 October 2015].

Pao W,Girard N (2011). New driver mutations in non-small-cell lung cancer. Lancet Oncol 12:175-80.

Pazdur R (2008). Endpoints for assessing drug activity in clinical trials. Oncologist 13 Suppl 2:19-21.

Roche Pharma AG (2015). Gazyvaro® German summary of product characteristics (Fachinformation); URL: http://www.fachinfo.de [last accessed on 8 October 2015].

Rolfo C, Sortino G, Smits E, Passiglia F, Bronte G, Castiglia M, Russo A, Santos ES, Janssens A, Pauwels P,Raez L (2014). Immunotherapy: is a minor god yet in the pantheon of treatments for lung cancer? Expert Rev Anticancer Ther 14:1173-87.

Ruof J, Knoerzer D, Duenne AA, Dintsios CM, Staab T,Schwartz FW (2014). Analysis of Endpoints Used in Marketing Authorisations versus Value Assessments of Oncology Medicines in Germany. Health Policy 118:242-54.

Sgambato A, Casaluce F, Sacco PC, Palazzolo G, Maione P, Rossi A, Ciardiello F,Gridelli C (2015). Anti PD-1 and PDL-1 immunotherapy in the treatment of advanced non-small cell lung cancer (NSCLC): a review on toxicity profile and its management. Curr Drug Saf [Epub ahead of print]; doi: 10.2174/1574886310666150928104210.

Study endpoints, orphan drugs: The methodology debate is in full swing

By Dr. Florian Staeck

Since the introduction of the early benefit assessment in 2011, the selection of trial endpoints and handling of orphan drugs have become controversial and widely-discussed issues in the AMNOG framework. In benefit assessments so far, only ~50 percent of primary endpoints were regarded as relevant for the demonstration of an additional benefit, which prompted several requests during the second meeting of the „Interdisciplinary Platform on Benefit Assessment" (Interdisziplinäre Plattform zur Nutzenbewertung). In particular, the value of surrogate parameters has been elaborately debated. The additional benefit for orphan drugs, which is automatically granted by the legislator, constituted another focus of the discussion.

In addition to the determination of the appropriate comparable therapy, the acceptance of data on endpoints, with which an additional benefit is to be proven, has continually been the subject of controversial debates since the introduction of the Act on the Restructuring of the Medicines Market (AMNOG). The law and the Drug Benefit Assessment Ordinance specify the target criteria of mortality, morbidity and quality of life as patient-relevant endpoints. However, the usability of data with regard to a proof effect is interpreted differently by the various parties within the AMNOG procedure.

The second meeting of the „Interdisciplinary Platform on Benefit Assessment" on 25/26 September 2015 in Kelkheim, was dedicated extensively to this topic, running under the title „Acceptance and patient-relevance of study endpoints". Here it became clear that the methodological debates regarding the acceptance of endpoint results continue and that initial signs are appearing of a consensus being reached in certain aspects. The 30 or so participants engaged in controversial discussions regarding the recognition of surrogate parameters as substitutes for a patient-relevant endpoint.

Here the large number of conditional and time-limited resolutions of the Federal Joint Committee (G-BA) shows how attempts are being made to address the – methodical – uncertainties with regard to proof of an additional benefit. The legally defined exceptional status of orphan drugs was also hotly debated at the meeting. Individual participants described the evidence situation for these new medicines as being particularly unsatisfactory, while others defended the legal status quo and stressed the particular difficulties encountered when measuring the patient-relevant aspects of a treatment with orphan drugs in studies.

Acceptance of endpoints: The focus of criticism according to the pharmaceutical manufacturers was the fact

that the primary study endpoints had not been deemed relevant as a proof of an additional benefit by the Institute for Quality and Efficiency in Healthcare (IQWiG) in around 50 percent of cases.

It was stated that the industry therefore required clear guidelines as to which endpoints should be used to create the studies prior to beginning the Phase III studies. Patients, and in particular test subjects, should have guarantee that the primary study endpoint, which was also determined in cooperation with the approval authorities, is patient-relevant – since they gave their consent to participate in studies on that basis. Other participants of the discussion added that it is disappointing for patients when they participate in a study but their claim to get the treatment is then not answered. It was asserted that the vital question is not whether the IQWiG looks at all data, but rather which data is ultimately accepted

This has consequences, as health insurance companies are increasingly calling prescriptions for drugs where the additional benefit has been classified by the G-BA as „not proven" into question. The same applies if the prescriptions are issued for sub-groups where the Federal Joint Committee has not seen any proof of an additional benefit.

This was countered with the argument that the mere fact that something has been studied does not mean that we are dealing with a patient-relevant endpoint. It is a legitimate goal that money should only be spent on those medicines for which good evidence exist. In particular, the focus of the methodological discussion is how ignorance should be weighted. If unsuitable endpoints have been selected, you do not learn anything about the actually relevant endpoints that have not been studied, argued some participants

The time limitation of benefit assessments, as now imposed in one in five resolutions by the G-BA, should also be viewed in the light of this backdrop. However, the use of this tool is occasionally taken too far: Attendees were reminded of the example of gliptine, which was approved 7 years ago. In summer 2016 the limitation in terms of a preliminary resolution will expire for quite a number of drugs that have previously undergone benefit assessment. It has thus far been unclear what happens if the manufacturers have not submitted any additional studies by the respective date.

Participants at the platform meeting considered this fact to speak strongly in favour of early benefit assessment being supplemented by an ongoing benefit assessment. However, this only makes sense if the G-BA is also given considerable authority to demand also new data from the manufacturers. Other participants did not believe that this suggestion had been well thought out: What would happen if a manufacturer did not submit the requested studies prior to the deadline? The question was asked as to whether the medication would then have to be withdrawn from the market.

The participants also disagreed on possible alternative means of regulation, for example with regard to subsequent post-approval studies. With this requirement the buck doesn't always have to be passed back to the industry. Some participants voted in favour of the commissioning of such studies being handled by scientific institutes rather than the G-BA. A post-approval evaluation must take place where patients are treated, and this is at doctor's practices or health centres. It is the task, for example, of the federal ministries for research and for health to establish a suitable procedural and legal framework for this. In this context, the potential of registry data was controversially assessed. It was suggested that registry data could form the basis for a type of „secondary" or „ongoing" benefit assessment. This was questioned with reference to the

previous, often lacking, usability of registry data. As a rule, the comparison of two interventions is not currently possible. The fact that there are already many registries in Germany but only a very small number of these data collections are independent was subject to criticism. Clinical cancer registries, for example, could not contribute to the benefit assessment to the required extent, as they do not reveal which drugs were used to treat the respective patient. In addition to this, new registries always incur additional document costs, which would lead to a low level of acceptance among doctors. It was therefore requested that the existing tools for benefit assessment should be improved, rather than demanding entirely new instruments such as registry data.

Relevance of surrogate parameters: The question concerning the significance of surrogate parameters in studies and how they can best be validated was subject to extensive debate. Participants stressed the clear differences regarding the recognition of an additional benefit by the G-BA, for instance in the case of lung cancer on the one hand, and in the case of chronic diseases such as diabetes mellitus on the other.

One reason stated for this was that „hard" endpoints are more easily measured in oncology than in the case of chronic conditions. In some types of cancer, however, a disease standstill has been reached in some patients, which comes hand-in-hand with the five-year survival rate of up to 85 percent. In the light of increasing chronification – a long-term latency of the illness without symptoms – it becomes increasingly difficult to measure an endpoint such as overall survival (OS). At the same time, only isolated patient-reported outcomes (PROs) have previously been investigated as primary endpoints.

Taking this into account, morbidity as a patient-relevant endpoint moved to the centre of the debate. In the discussion, classic clinical symptoms such as pain reduction were defined as classical examples of patient-relevant endpoints in morbidity. Results from imaging or laboratory tests were controversially discussed as surrogate parameters. Here differences in the understanding of disease-related morbidity in particular were evident. The following question was asked: Is morbidity to be understood in the sense of a tumour growing or a patient showing symptoms? It was argued that if the focus is on the reduction of symptoms then the discussion of surrogates is out of place, as symptoms can be easily measured. Surrogates may then be worthwhile if the overall survival time cannot be ascertained on account of the short time frame of the clinical study or due to the extended survival time of the patients.

However, it was argued that in this case the determination of the extent of the additional benefit is associated with substantial methodological uncertainty, as it is unclear what the improvement of a surrogate parameter by the factor X actually means. This applies in particular to laboratory or imaging parameters, where the interpretation of the results is by no means unanimous among experts. Doctors actively involved in the provision of care confirmed that a focus on technology and lab values in many cases leads to an overtreatment in patients with a mere deviation from the norm without calling into question the clinically relevance.

Another methodological problem arose from the fact that with ever smaller study collectives, it is not possible to validate a surrogate parameter. It was commented that the potential compilation of data from several research-conducting manufacturers is only possible in individual cases, given the competitive environment in which the pharmaceutical companies operate.

During the discussion regarding the relevant morbidity, the participants requested that the term „therapeutic rele-

vance“ should be used. It was asserted that it allows the option to consider also the disease burden for the patients. However, „real world data“ from practice is required for precisely this reason. Reference was also made to the practice of the US approval authorities, the FDA, who voted that endpoints should be adjusted to the different stages of a disease – in the case of prostate cancer for example. This makes it possible to model new sequential therapy forms in oncology better than it was previously the case, it was argued.

In any case, from the manufacturers‘ perspective it is essential that planning security should be established by means of the early agreement concerning the relevant endpoints with the approval authorities as well as the benefit assessment institutions. It became very clear during the course of the meeting that surrogate parameters without validation have no chance of acceptance in the early benefit assessment.

Status of orphan drugs: The platform participants did not reach a consensus with regard to the question as to whether the current statutory regulation where the additional benefit is demonstrated just because of the EMA´s orphan drug designation should continue to be in force or not. Studies concerning the FDA approval procedure showed that there were a lower proportion of randomised and blind studies for orphan drugs, compared to those for non-orphan drugs.

At the same time, the proportion of orphan drugs among the G-BA's benefit dossiers is increasing: It was reported that in 2014 one on four dossiers was for an orphan drug. Astonishment was expressed at the fact that in spite of the often unsatisfactory data situation, the G-BA had only imposed two time limits in the case of orphan drugs. Supporters of the status quo argued that the limit, whereby no more than five in 10,000 people EU-wide can be affected by the condition, should continue to be accepted. It was emphasised that clinical progress could be generated through good documentation and careful exchange among colleagues. An example in this context are the developments in pediatric oncology in recent years.

Studies can indeed also be conducted with small patient collectives. However, there was doubt as to whether new medicines could continue to be developed in the future without the special status of orphan drugs. If the additional benefit of these medications is not considered to have been proven by the approval, then the price of the appropriate comparable therapy would be set as the upper limit during price negotiations with the National Association of Health Insurance Funds. If there is no evidence that a new product offered an additional benefit, then only a price of the ‚best supportive care' could be achieved. It was warned that, as a result, in many cases orphan drugs would presumably no longer be available. In addition, critics were also operating on false assumptions: For functional reasons, the approval authorities made the decision on the designation of an orphan drug – not the manufacturers.

These arguments were rejected by critics of the current regulations. It was commented that the fact that non-inferiority studies for orphan drugs, for example, were submitted to the G-BA was tantamount to giving up on what was originally intended with the AMNOG. Nowhere in the EU documents is it laid down that weaker study designs are accepted for orphan drugs. On the contrary, it is stipulated in the relevant EU Directive from 2001 that concerning these medicines do patients have a right to assessment methodology that is comparable with that applicable for randomised clinical studies. It was asserted that there is no scientifically discernible justification for differing approaches in the case of medicines for rare conditions. Moreover, the case numbers in the submitted studies were

often not as low as reported. It was demonstrated that if you take the population of the EU as a basis, five in 10,000 residents amounts to 230,000 people. As a result, the legislator should scrutinise the additional benefits that it assumes for these medications and adjust the AMNOG accordingly. On the other hand there was a consensus that the question regarding comparative evidence must be allowed, at least in the cases in which several orphan drugs are approved for the same indication.

In the synopsis of the debate, participants referred to discrepancies in the institutional setting of the early benefit assessment: The methodological complexity of the procedure in the determination of the additional benefit is not reflected in the subsequent price negotiations between the manufacturers and the National Association of Health Insurance Funds. It was criticised that the findings with regard to sub-group-specific additional benefits are levelled out again as the result of mixed pricing. As a result, the instrumental purpose of the AMNOG to enable the fair pricing of a new medicine has been neglected. However, an AMNOG-assessed active substance with positive sub-group results remains in the market and is thus fully prescribable, even in the sub-groups with negative results. It was asserted that it is unrealistic to assume that doctors meticulously apply the results of sub-group assessments on a 1:1 basis when issuing prescriptions on a daily basis.

Given the large number of requests with regard to the amendment of the AMNOG, participants warned against overloading the AMNOG with tasks. It was argued that the complexity of the methodological requirements in benefit assessment and during the subsequent price determination will not be able to be legally addressed. Here codes of procedure in particular are the adequate means of implementing amendments to the AMNOG.

On one point, however, the participants of the second platform meeting were united, agreeing on a need for action: As a result of the AMNOG, five years after the implementation of the AMNOG, there is a host of high-quality information regarding new drugs, available at an early stage. However, up to this point it has not been possible to make this vast amount of information available to treating physicians, in a targeted manner and in an edited format. It was stressed that greater efforts are required within the health insurance system in this regard.

LEGAL NOTICE

ISSUER
Editorial Advisory Board of the
Interdisciplinary Platform:
Prof. Dr. Ulrich Schwabe,
Dr. Harald Herholz

PUBLISHER
Springer Medizin Verlag GmbH
Am Forsthaus Gravenbruch 5
63263 Neu-Isenburg

EDITORIAL PROCESSING
Dr. Florian Staeck
Wolfgang van den Bergh
Helmut Laschet

AUTHORS
Dr. Thomas Kaiser
PD Dr. Michael Kulig
Prof. Dr. Bernhard Wörmann
Prof. Dr. Dietmar P. Berger
Dr. Miriam Walter
Dr. Florian Staeck

IMAGE CREDITS
Title image: Mathias Ernert

LAYOUT / GRAPHICS
Sandra Bahr
Oliver Hippmann

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

Springer Medizin

In cooperation with, and with the friendly support of Roche Pharma AG, DAK Gesundheit, Xcenda GmbH and SpringerMedizin

DISCUSSION PARTICIPANTS

Dr. Pamela Aidelsburger
Silke Baumann
Dr. Jürgen Bausch
Prof. Dr. Dietmar P. Berger
Dr. Margita Bert
Dr. Johannes Bruns
RA Claus Burgardt
Prof. Dr. Eva Susanne Dietrich
Dr. Mathias Flume
Dr. Markus Frick
Prof. Dr. Klaus Genth
Dr. Harald Herholz
Wolfgang Kaesbach
Dr. Thomas Kaiser
PD Dr. med. Michael Kulig
Lothar Kuntz
PD Dr. Stefan Lange
Helmut Laschet
Dr. Ulf Maywald
Prof. Dr. Herbert Rebscher
Prof. Dr. Jörg Ruof
Dr. Vanessa Schaub
Prof. Dr. Ulrich Schwabe
Prof. Dr. Friedrich Wilhelm Schwartz
Dr. Florian Staeck
Wolfgang van den Bergh
Prof. Dr. Bernhard Wörmann

INTERDISCIPLINARY PLATFORM ADVISORY BOARD

Dr. Pamela Aidelsburger
Dr. Jürgen Bausch
Wolfgang van den Bergh
Dr. Harald Herholz
Dr. Rainer Hess
Helmut Laschet
Dr. Ulf Maywald
Prof. Dr. Jörg Ruof
Prof. Dr. Ulrich Schwabe
Prof. Dr. Friedrich Wilhelm Schwartz
Dr. Florian Staeck

INTERDISCIPLINARY PLATFORM ON BENEFIT ASSESSMENT

Clinical Studies – which endpoints count?

Vol. 2
March
2016