Dealing with evidence

Between consistency and flexibility

Based on the provisions of SGB V, evidence-based medicine (EBM) forms the basis for the work of the Federal Joint Committee (G-BA) and the Institute for Quality and Efficiency in Health Care (IQWiG) in their medical assessments and decisions on the specification of care content such as disease management programmes (DMP) and the services to be financed by the statutory health insurance (SHI). The evaluation of available evidence is thus directly linked to pricing, reimbursement and also to care management. Greater flexibility in dealing with the evidence requirements has been exercised in the past, but without consistency and certainly also when a lowering of the requirement level led to a financial advantage for the statutory health insurance (SHI). At the same time, there are also examples in which the clinical relevance of an intervention was so obvious that the G-BA used its room for manoeuvre in the interests of healthcare.

>> While the origins of EBM always favoured the use of the best available evidence, the G-BA focuses on the best achievable evidence. This tendency towards a rigid approach often does not do justice to specific therapy situations, especially when the requirement for corresponding studies is not appropriate. There is no robust set of rules for dealing with such situations. Therefore, if the fundamental feasibility of greater flexibility in dealing with the assessment of evidence exists, the establishment of a transparent and comprehensible framework for this flexibilisation would be necessary so that corresponding decisions do not appear inconsistent or even arbitrary. This necessity arises in particular from innovative developments in precision medicine, gene and cell therapies - and in future probably also mRNA technologies - with new challenges for evidence generation and pricing.

In the future, an adapted evaluation taking into account the entire spectrum of evidence could support price negotiations in order to agree on an appropriate reimbursement amount in a free negotiation involving new reimbursement models.

Dealing with evidence - flexible when needed?

Based on the provisions of SGB V, EBM forms the basis for the work of the G-BA and IQWiG in their medical assessments and decisions on the specification of care content such as DMPs and the services to be financed by SHI. The assessment of available evidence is therefore directly linked to pricing, reimbursement and also to care management. This also raises questions as to how strict or flexible the handling of evidence should be. Depending on the context, different views are permitted. However, it should be undisputed that such handling must be transparent and consistent. The past shows that at least this can be questioned.

For example, it is actually taboo for health insurance companies to cover the costs of off-label therapies as a rule. The reason given by health insurance companies is always the lack of evidence and thus uncertainty regarding efficacy and safety, which is necessary to protect patients. Proceedings relating to disputes over benefits for off-label therapies have been taken to the Federal Social Court in several cases and once even to the Federal Constitutional Court [1]. However, it can also happen that health insurance companies explicitly promote off-label therapies if they are financially more advantageous, even if alternative treatments approved for the specific use are available. The example of ‚Avastin' (bevacizumab), which is used off-label for the treatment of eye diseases even though ‚Lucentis' (ranibizumab) has official approval for this purpose, is a prominent example of this inconsistent approach. In contracts with ophthalmologists, complex flat rates were specifically agreed that were not adequate for the authorisation-compliant injection of ‚Lucentis', but enabled above-average remuneration for the use of the cancer drug ‚Avastin'. Irrespective of the arguments otherwise put forward by the SHI, off-label use was not only tolerated here, but actively incentivised [2]. Such examples can favour the perception of a double standard in which economic considerations are placed above scientific evidence and personal claims.

Another example is the inclusion of the first biomarker-based test (gene expression test) for deciding on chemotherapy for breast cancer as a standard SHI benefit. This was preceded by a six-year evaluation process. The long duration was mainly due to the high evidence requirements of the G-BA and IQWiG, which demanded randomised studies with a 10-year observation period. Of several available tests, only the Oncotype DX® test was able to demonstrate the required evidence and was consequently included in the SHI benefits catalogue. For a follow-up assessment of the other tests, IQWiG was commissioned by the Federal Joint Committee - irrespective of all previously used criteria for the suitability of a study and including studies with a low level of evidence - to merely research and present the current state of knowledge instead of a benefit assessment. The original requirement for 10-year data was also explicitly dropped. Nevertheless, IQWiG demonstrated the lack of evidence of benefit and the lack of concordance with the ‚Oncotype DX' test in its report for the other tests. If the assessment standards that otherwise apply to the introduction of new methods into standard care had been applied, the situation would have been clear. Nevertheless, all available biomarker tests were transferred to standard care - they were significantly more favourable than the ‚Oncotype DX' test, which was assessed according to stricter evidence criteria. In summary, a method that had been proven to be beneficial in the meantime was withheld from care for many years in order to exclude a possible risk to patients and subsequently spark competition with products that were not held to the same standards. The motivation of manufacturers to carry out such a costly study is likely to have been permanently damaged by this decision.

A current example can be seen in the discussions about the designation of so-called combination drugs by the G-BA as a result of the GKV-FinStG. In the interim decisions, the G-BA not only named those combinations that were explicitly authorised as combination therapies, but also generally concluded that their use in combination therapy always complied with the authorisation if there was

a fundamental agreement in the area of application. Ultimately, it also expressed a departure from the previous understanding of authorisation, according to which the use of a medicinal product is generally only made possible on the basis of the official assessment of clinical trials [3]. The review of the state of scientific knowledge for such combinations was completely dispensed with. As a result, in many decisions the G-BA not only named fictitious combinations for which there was no explicit authorisation, but which also lacked any medical rationale. The expected negative effects on healthcare and the benefit assessment procedure itself were initially ignored. The G-BA only relented on the basis of an intensely critical discussion and finally announced that it would only name drugs for which at least some information on use as a combination therapy was available in the Summary of Product Characteristics [4], not only in the future but also retrospectively. However, an evidence-based review of the plausibility of the designations was also subsequently dispensed with.

Fortunately, there are also examples in which the G-BA and IQWiG have shown the necessary flexibility when obviously beneficial innovations threatened to get tangled up in the otherwise rigid framework of the benefit assessment. A series of direct-acting antivirals (DAAs), which came onto the market from the beginning of 2014, revolutionised the treatment of hepatitis C and demonstrated their efficacy in studies by achieving a sustained virological response (SVR) for the majority of affected patients. Achieving the endpoint means, for example, that no HCV RNA can be detected in the blood for 24 weeks (SVR24) or longer. If the causative agent of an infectious disease is no longer detectable after a few weeks, a cure can be assumed. However, in the benefit assessment procedure for ‚Sovaldi' (sofosbuvir), IQWiG found that the endpoint SVR24 was not a patient-relevant endpoint. Although not formally validated, SVR24 could be understood as a sufficiently valid surrogate endpoint for the occurrence of hepatocellular carcinomas; however, it was not possible to quantify the added benefit. For several genotypes of the disease, only historical comparisons were possible in this procedure, none of which were recognised by IQWiG and consequently no added benefit was recognised for the majority of the target population. Intense disputes in the commenting procedure and in the expert community then led to the usual methodological requirements being abandoned. As a result, SVR was recognised as a directly patient-relevant endpoint in the assessment of drugs for the treatment of chronic hepatitis C and the uncontrolled studies were also used to assess the added benefit [5].

Despite the high costs of the therapy, which clearly worried the health insurance companies at the time, an additional benefit was recognised for at least some of the patients. The recognition of single-arm studies in this case represented a relevant deviation from the otherwise usual assessment criteria, which can best be explained by the attempt to avoid burdening the assessment procedure with medical credibility problems. This exception also illustrates that there is no recognisable systematic approach for dealing with special situations and that the decision was inconsistent as a result. Against the background of such examples, it should be noted that the AMNOG procedure should recognise inconsistencies and failures of the past as lessons that can help to make the system more transparent and flexible in the future. This means that the evaluation process requires more openness and transparency so that patients, physicians and the public can better understand how decisions are made. Greater flexibility could also enable the system to respond to new scientific findings and changing health needs.

Framework conditions for the future handling of evidence

Between best available evidence and best achievable evidence

The founders of evidence-based medicine (EBM) envisioned that treatment decisions should be based on the best available scientific evidence. Probably the best-known definition comes from David L. Sackett: ‚EBM is the conscientious, explicit and judicious use of current best external scientific evidence to inform decisions in the medical care of individual patients.' In practice, this means ‚combining individual clinical expertise with the best available external clinical evidence from systematic research' [6].

Treatment decisions have to be made constantly in day-to-day care, because even the non-application of an existing treatment option constitutes such a decision. Doctors can therefore do no more than incorporate the ‚best available' knowledge into the therapeutic decision. With the introduction of EBM into the Fifth Book of the German Social Code (SGB V), its application was extended beyond medical care and decision-making to questions of benefit law and the pricing of interventions. For example, the terminology can be found in the specifications for the work of the IQWiG (Section 139a SGB V), the proof of a therapeutic improvement in the formation of reference price groups (Section 35 SGB V) and the additional benefit of drugs (Section 35a SGB V).

At the same time, EBM was reinterpreted in such a way that it was no longer based on the best available evidence, but on the best achievable evidence. As a result, the original idea of EBM was lost sight of: If the ideal evidence is not available, all other available studies are generally ignored and are precisely not included as ‚best available evidence' in the (benefit) decision. But if a therapeutic decision is possible on the basis of existing knowledge, why should this not apply to reimbursement in SHI? IQWiG's concept of requiring so-called ‚dramatic effects' is the only methodologically legitimised exception to the acceptance of historical comparisons, for example. IQWiG's definition of a dramatic effect is based on an observed relative risk of 10. The associated requirement for a reliably documented fatal course of the disease is explained by IQWiG on the basis of insulin administration in type 1 diabetes mellitus [7]. It is obvious that although this requirement describes the handling of a special case that is conceivable in principle, in the reality of medical progress it is not a suitable concept for doing justice to the actual idea of EBM.

Data collection accompanying application as an opportunity for special cases

The Act for More Safety in the Supply of Medicinal Products (GSAV) introduced the possibility of requiring so-called application-related data collection (AbD) for medicinal products for the treatment of rare diseases (orphan drugs) and for certain forms of authorisation. The Federal Ministry of Health (BMG) explicitly demanded ‚that randomised blinded clinical trials are not included'. At the same time, the BMG demanded that the ‚results of an application-accompanying data collection can be used to quantify an additional benefit within the scope of the benefit assessment' [8]. These regulations in the AMNOG should lead to a rethink in the handling of evidence in several key areas:

- Dynamisation of evidence generation: data from clinical studies, which are conventionally collected in the run-up to approval and market launch, will no longer be the

only pillar of evidence that benefit assessments can refer to and rely on .

- Evidence generation close to care: Data collected during the actual use of a medicinal product reflect the reality of care. The mandatory inclusion of such data would therefore improve the chances of utilising both types of evidence (RCT and real-world evidence).
- Individualisation of evidence: With data collected close to the application, differences in the response of different patient groups can be better recorded. Ideally, care-relevant knowledge is generated that promotes a shift towards individualised medicine in which treatments are tailored more closely to individual patients.

At the same time, the AbD could be a step towards the harmonisation of HTA and approval processes, in which the regulatory authorities carry out accelerated procedures with the requirement of subsequent data collection (e.g. in registries), so that despite the differences between approval and HTA, a meaningful interlocking of both processes remains possible. The AbD should actually have a catalysing function here, because the fact that an additional benefit is to be quantified in future with non-randomised registry studies means that the G-BA and IQWiG are being encouraged to rethink their methods. The legally prescribed inclusion of non-randomised, healthcare-related data in the benefit assessment makes it necessary to develop the traditional, often rigid concept of evidence-based medicine into a more dynamic, adaptable model. Despite this conceivable catalytic function of the AbD in making the overall benefit assessment procedure more flexible, it is premature at this stage to draw a definitive conclusion. At the same time, it is foreseeable that the AbD can only be a solution in very specific cases and is not the answer to all the challenges of drug assessment.

Key points for the consistent and flexible handling of evidence

In order to create a universally suitable framework for a more consistent handling of evidence, the following key points would be conceivable. The assessment of the added benefit of new drugs is based on a comparison with an appropriate comparator therapy (ACT). This is a fictitious study comparator, which is often formed differently for individual subpopulations. As this determination is made with knowledge of the existing authorisation study, the G-BA often creates questions for which it already knows in advance that they cannot be answered with the available study data. The vast majority of procedures with unrecognised added benefit are based on the fact that the G-BA requested data that was simply not available. Virtually all existing evidence in which the study comparator deviates from the definition of the ZVT is not included in the assessment.

Even the point enshrined in Section 5 AM-NutzenV that it may be ‚impossible or inappropriate' to ‚conduct or require studies of the highest level of evidence' does not lead to anything in the assessment reality, as this is not systematically checked. Instead of the necessary review, IQWiG and the G-BA regularly argue exclusively on methodological grounds that studies of lower evidence levels than RCTs should not be included in the assessment. The question of whether ‚the new' is better than ‚the previous' is obvious and justified in principle. However, in order to rectify the described flaw, the benefit assessment would have to be based on an extension of this question, namely the additional questions ‚What information is available on the benefit of the new product?' and ‚What evidence would be possible and appropriate?'

The consequence of the expanded question would be a departure from the exclusive inclusion of only the highest quality of evidence. Instead, there would be a graduated assessment taking into account the entire spectrum of evidence. It would be left to the contractual partners in accordance with Section 130b SGB V to agree on an appropriate reimbursement amount in a free negotiation, taking into account the available assessed evidence, its certainty of results and other factors.

A further fundamental need for adjustment arises from the obviously strong focus of the German HTA on avoiding the first type of error - i.e. the wrongly recognised additional benefit. The question of what harm could be caused by an unjustifiably unrecognised additional benefit is regularly included in decisions by the regulatory authorities. In particular, the authorisation of new medicinal products whose evidence is assessed by the G-BA and IQWiG as insufficient or immature is often the result of a consideration by the European Medicines Agency (EMA). Conditional approval, for example, is based on weighing up the advantage of rapid availability of the medicinal product against the risk of less extensive data than would normally be required. In addition, the requirement to fulfil a previously unmet medical need must be met. As the G-BA does not systematically carry out such a balancing of interests, the products concerned are likely to fail the evidence requirements of the added benefit assessment.



Another option would be to promote the dynamisation of evidence generation and evidence assessment. To this end, the establishment of suitable structures for systematic evidence-generating care would be helpful. A promising step has already been taken with the draft Health Data Utilisation Act (GDNG) and the possibility it provides for linking SHI routine data with cancer registry

data. The establishment of a solid registry structure in other indication areas and the subsequent linkage with the research dataset would be a logical further development. These data are intended and also suitable for research purposes; the previous restriction to a narrow circle of authorised users should be lifted. It is to be hoped that the resulting research results, together with the study data that the G-BA and IQWiG have so far not considered worthy of consideration, will in future form a more reliable basis for assessing the value of a drug therapy.

The Digital Act (DigiG) is intended to introduce mandatory performance-related remuneration for digital health applications (DiGA). However, dynamic reimbursement models cannot be implemented in the pharmaceutical price negotiations with the National Association of Statutory Health Insurance Funds. However, these may be more appropriate than rigid pricing, especially for therapies that are authorised due to the high therapeutic need and special medical features with evidence below the highest evidence level. To this end, the actual therapeutic success achieved in the provision of care should be included in pricing as supplementary evidence. The contracting parties can be trusted to identify suitable therapy situations and success parameters and to develop intelligent contract models.

Conclusion

The adaptation of evidence assessment in the German healthcare system is of crucial importance in order to meet the challenges of modern medical developments. The current practice, which focusses on ‚best achievable evidence', can lead to an inadequate evaluation of innovations and thus ultimately to restrictions in the care of patients. In order to meet this challenge, a transparent framework should be established that defines clear criteria and procedures for a desired flexibilisation in the handling of evidence, in which the entire spectrum of evidence can be taken into account. This should be accompanied by the development of dynamic reimbursement models that enable appropriate remuneration for innovative therapies. <<

by: Hans-Holger Bleß, fbeta

Citation

Bleß, H.-H.: ‚Between consistency and flexibility', in ‚Monitor Versorgungsforschung' (06/23), pp. 42–45. http://doi.org/10.24945/MVF.06.23.1866-0533.2562

Literature

1. Bundesverfassungsgericht. BVerfG, Beschluss des Ersten Senats vom 06. Dezember 2005 [Internet]. 2005 Dez [zitiert 2. November 2023]. Verfügbar unter: https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/2005/12/rs20051206_1bvr034798.html
2. Deutsches Ärzteblatt. Off Label Use in der Augenheilkunde: Streit auf dem Rücken von Patienten und Ärzten [Internet]. 2007 [zitiert 2. November 2023]. Verfügbar unter: https://www.aerzteblatt.de/archiv/57221/Off-label-Use-in-der-Augenheilkunde-Streit-auf-dem-Ruecken-von-Patienten-und-Aerzten
3. Gemeinsamer Bundesausschuss. Zusammenfassende Dokumentation über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XIIa – Kombinationen von Arzneimitteln mit neuen Wirkstoffen nach § 35a des Fünften Buch Sozialgesetzbuch (SGB V): Ergänzung zur Benennung von Kombinationen gemäß § 35a Absatz 3 Satz 4 SGB V [Internet]. 2023 Okt [zitiert 2. November 2023]. Verfügbar unter: https://www.g-ba.de/downloads/40-268-9837/2023-10-05_AM-RL-Anlage-XIIa_Ergaenzung-Benennung-Kombinationen_ZD.pdf
4. Gemeinsamer Bundesausschuss. G-BA regelt Benennung von Arzneimittel-Kombinationen [Internet]. 2023 Okt. Verfügbar unter: https://www.g-ba.de/presse/pressemitteilungen-meldungen/1134/
5. Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Sofosbuvir [Internet]. 2014 Jul [zitiert 2. November 2023]. Verfügbar unter: https://www.g-ba.de/downloads/40-268-2899/2014-07-17_AM-RL-XII_Sofosbuvir_2014-02-01-D-091_TrG.pdf
6. Sackett DL. Evidence-based medicine. Seminars in Perinatology [Internet]. 1997 Feb;21(1):3–5. Verfügbar unter: https://doi.org/10.1016/s0146-0005(97)80013-4
7. IQWiG. Allgemeine Methoden Version 7.0 [Internet]. IQWiG. 2023 Sep [zitiert 2. November 2023]. Verfügbar unter: https://www.iqwig.de/methoden/allgemeine-methoden_version-7-0.pdf
8. Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses gemäß § 91 SGB V vom 16. Juli 2020 hier: Änderung der Verfahrensordnung (VerfO): Änderung im 5. Kapitel – Verfahren zur Forderung einer anwendungsbegleitenden Datenerhebung nach § 35a Absatz 3b SGB V. 2020 Okt.