# Real World Evidence in Health Technology Assessments

Dirk Eheberg

IQVIA – Munich, Munich

## Abstract

Health technology authorities often ignore Real World Evidence, even if these data represent the best evidence available for a specific research question. It is important to understand, that data generated by both Real World Evidence and randomized controlled trials have advantages and disadvantages to be considered in a health technology assessment. While data from randomized controlled trials usually have a very high internal validity, these data often lack the external validity and the larger sample sizes that data from Real World Evidence studies can offer. The hallmark of randomized controlled trials is bias control, but with advanced analytical techniques, like Gaussian processes or vector based machine learning, Real World Evidence can overcome this disadvantage. In summary, health technology assessments can benefit from Real World Evidence if strengths and limitations are considered and handled accordingly.

## Zusammenfassung

**Real World Evidence in Arzneimittel-Nutzenbewertungen**
Die für die Nutzenbewertung von Arzneimitteln zuständigen nationalen Behörden ignorieren häufig Daten aus Real-World-Evidence-Studien, selbst dann, wenn diese Daten die beste verfügbare Evidenz in dem spezifischen wissenschaftlichen Kontext darstellen. Dabei ist es wichtig zu verstehen, worin die Vorteile und Nachteile der für Nutzenbewertungen herangezogenen Daten bestehen, je nachdem ob diese über Real-World-Evidence-Studien oder in randomisierten kontrollierten Studien generiert wurden. Daten, die aus randomisierten kontrollierten Studien gewonnen werden, haben meist eine sehr hohe interne Validität, während ihnen üblicherweise externe Validität und ein großer Stichprobenumfang fehlt, welche wiederum typische Stärken von Real-World-Evidence-Studien sind. Ein wesentliches Kennzeichen randomisierter kontrollierter Studien ist das niedrige Verzerrungspotenzial, welches dem Studiendesign inhärent ist. Real-World-Evidence-Studien können diesen Nachteil jedoch über moderne, fortschrittliche Analysetechniken, wie den Gauß-Prozess oder der Support Vector Machine, einer Anwendung des maschinellen Lernens, ausgleichen. Zusammenfassend kann festgestellt werden, dass Nutzenbewertungen von Real-World-Evidence-Daten profitieren können, wenn die Stärken und Limitationen dieser Daten entsprechend berücksichtigt werden.

## Real World Evidence offers valuable insights for research questions alongside RCTs

With recent advances in data availability and more sophisticated methods to combine, process and analyze large amounts of data from different sources, Real World Evidence (RWE) or Real World Data (RWD) offers valuable insights for all kind of research questions alongside the life cycle of a product. However, for market access in general and health technology assessments (HTA) in particular, RWD is often ignored entirely or used only sparsely.

This article highlights reasons for the lack of use of RWD in HTA and addresses why this has a negative impact on the evidence-based decision making process and how RWD can be made fit for the use in HTA.

## HTA agencies across Europe base their decisions on RCT data

Despite the increasing need for data provided under routine settings of health care practice in actual patient groups under real world conditions, HTA agencies heavily and almost exclusively rely on data generated through randomized controlled trials (RCT). Data generated by RCTs and meta-analysis of these data represent the highest category of evidence in evidence-based medicine (EBM), which is the foundation of HTAs. These data build the top of the hierarchical pyramid of evidence, closely followed by data generated by prospective observational cohort studies and pragmatic trials. Case-reports and expert opinions are consentaneously considered the lowest level of evidence. This hierarchical perception of evidence has led to the deceptive impression, that data generated by RCT are the only viable source for a HTA. Consequentially, HTA authorities are unlikely to accept non-RCT data, even if data generated by RCTs are not available for the specific research question. Even though, the evidence requirements in most European countries are rather similar, with a very strong emphasis on basic clinical trial data and comparative clinical analysis for clinical evidence

and basic pricing data and budget impact analysis for economical evidence, the countries differ strongly in the acceptance of RWE data (fig. 1).

## The UK and the Netherlands accept RWE to demonstrate treatment effectiveness in the absence of RCTs

Only the United Kingdom (UK) and the Netherlands (NL) accept RWE to demonstrate treatment effectiveness in the absence of adequate RCT data. A good example for the insights generated by RWD is the case of Zytiga® in the therapeutic area of metastatic castration resistant prostate cancer before chemotherapy in the UK. Initially, Zytiga was rejected by NICE as no improvement in long-term survival and cost-effectiveness could be demonstrated based on the RCT data. To provide a better extrapolation of the treatment effects to clinical practice, an analysis of pharmacy prescription data from five specialist centers in England was conducted. These findings were triangulated against US healthcare insurance claims data, which showed that 14 % of patients were still on treatment after more than 4.4 years. This model based on RWE data provided compelling evidence that some patients in the UK were likely to take Zytiga for a long period of time, which is an indicator that the patients would benefit from the treatment. After the review of this data by NICE, the initial decision was changed and Zytiga considered a cost-effective treatment option for patients with metastatic castration resistant prostate cancer before being treated with chemotherapy when compared to best supportive care.

## In other European countries, RWD is only accepted as supportive evidence

In other countries, RWD are often limited to a supportive role in a HTA process, providing a database for epidemiology, resources usage and cost data. Therefore, RWE data are recommended for pharmacoeconomic analyses in all countries, which include such analyses in the HTA process. The impact of RWD on the HTA submissions should not be neglected, even in countries where RWD are generally not accepted to demonstrate treatment effectiveness. An IQVIA HTA Accelerator analysis on all single drug submissions in France between January 2011 and July 2016 in the therapeutic areas cardiovascular, central nervous and metabolic diseases has demonstrated that RWE data are often included in these HTA submissions and that those HTA submissions including RWE data were more likely to have a better assessment rating (fig. 2).

47 of the 119 assessments (39.5 %) in these therapeutic areas included in the HTA Accelerator analysis made use of RWD. 90 % of the assessments that did not include RWD received an "Improvement of Medical Benefit Assessment" (AMSR) rating of V, which is considered no improvement. Those submission that included RWD were more likely to receive a better ASMR rating, with only 72 % receiving an ASMR rating of V. Additionally, none of the assessments not including RWE data received a better rating than III (significant improvement). However, the conclusion that submissions including RWD receive a better rating because of the inclusion of such data would be hasty. However, a direct hint on the importance of RWD is the circumstance that the lack of RWD is mentioned by HAS as a negative driver for the decision in nine of the 72 assessments without RWD.

## In France, RWD can positively influence the HTA decision making

In conclusion, whether RWD will be accepted for HTA is highly dependent on the country. In some countries, RWD have an important influence on the decision making process even if only used as supportive evidence. At the same time, RWD would have much to offer for the HTA process.

RWD offer insights into a product's performance compared with standard of care in cases in which the RCT does not provide an appropriate comparison.

In general, in any HTA submission there are always four areas, which need to be covered to be suc-

## AUTHOR

***Dirk Eheberg***

holds a degree in biology from the University of Munich. After a few years working in basic research he started post-graduate studies in epidemiology and public health and completed his studies with a master degree in public health from the University of Munich (MPH). He gained experience in epidemiology, biometrics and working with SAS. Furthermore, he participated in special training programs for decision analytic modelling (decision analysis and Markov models in particular).

At IQVIA, Dirk Eheberg works as Senior Project Manager and Strategy Lead. Since 2011 he focuses on benefit dossiers, especially on AMNOG dossiers module 3 and module 4. Dirk gained Health Technology Assessment (HTA) experience in over 50 AMNOG projects and has actively participated in a number of Federal Joint Committee (G-BA) meetings incl. advice requests and oral hearings. Additionally he is specialized on decision analytic modelling in the health care sector (e. g. hemophilia and rare diseases). He was involved in several modelling projects (Markov cohort models, decision tree models and budget impact models) for the German market, and is experienced in international model adaptations and cost researches. Dirk has gained skills in both constructing and adapting user friendly excel-based Markov models.

Since 2012 he is responsible for the calculation and simulation of reference prices for Germany and the IMS Festbetrags-Analyzer®, specific software for reference price calculation.



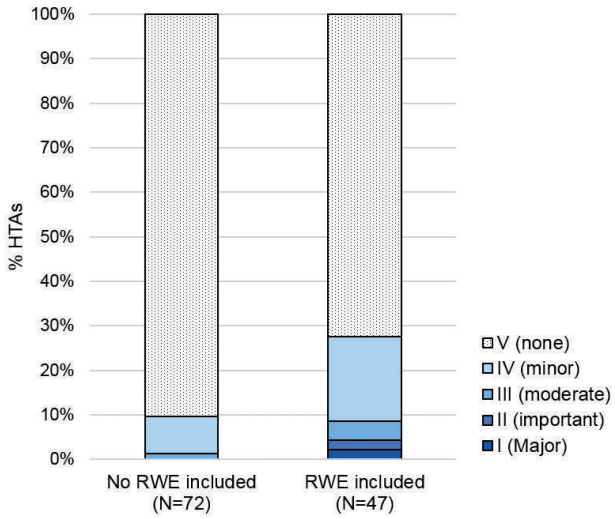
Pharm. Ind. 80, Nr. 4, 476–481 (2018)

■ **Figure 1**

| | | | | | | |
|---|---|---|---|---|---|---|
| RWD accepted for *initial reimbursement discussions* | ✓ | EPI ONLY ✓ | ✓ | ✓ | ✓ | ✓ |
| RWD used to demonstrate treatment effectiveness in the absence of adequate RCT data | ✗ | ✗ | ✗ | ✓ | ✓ | ✗ |
| Epidemiological, resource use or cost data derived from RWD | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| RWD recommended for pharmacoeconomic analyses | ✓ | n/a | ✓ | ✓ | ✓ | ✓ |
| *Preference for RWD in conditional reimbursement schemes* | ✓ | n/a | ✓ | ✗ | n/a | n/a |
| *RWD used in conditional reimbursement scheme decision making if evidence is robust* | ✓ | n/a | ✓ | ✓ | n/a | n/a |

*Acceptance of Real World Data (RWD) for the HTA process among selected European countries (Source: All figures were made by the author/IQVIA Consulting Services).*

■ **Figure 2**




*Distribution of benefit ratings by HAS for submissions with and without RWD.*

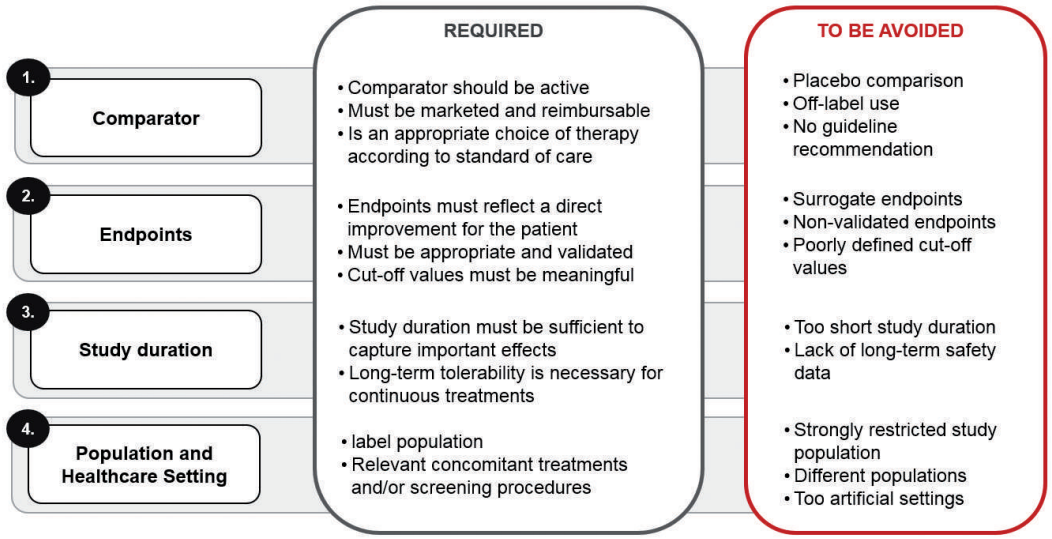
cessful. An HTA is comparative by concept; therefore, evidence against a comparator is inevitable for a successful submission. Ideally, the comparator should be an appropriate choice of therapy according to standard of care. The choice of meaningful endpoints for the assessment is as important as the choice of the comparator. The endpoints should directly reflect an improvement for the patient or be suitable to be used in a cost-effectiveness evaluation, depending on the preferences of the HTA authority. In both cases non validated and inappropriate endpoints will make an assessment impossible.

## RCTs study duration falls short of long-term outcomes, a gap that RWD can close

In order to detect benefits and harms for the patients, the period of observation must be long enough to allow the detection of the events of interest. This is not always possible in a standard trial, especially if the events of interest are extremely rare but highly important or long-term treatment outcomes. Finally, the transferability of the results to the real-life patient population treated under standard healthcare conditions is highly relevant. Therefore, concomitant treatments and screening procedures should be considered. HTA should consider the best evidence available, which will often be data generated by RCTs. These four core factors can be addressed by a RCT if the design of the trial was chosen accordingly ( fig. 3).

Especially the choice of the comparator and the endpoints are normally strong arguments for RCTs as the data basis for a HTA, as RCTs are comparative by design. However, HTA is not limited in its own definition to RCT data. Sometimes RCT data are not available, especially in rare diseases or for medical devices. In this case, the limitation is not that a RCT could not be designed for those research questions on the sketchbook, but often the conduction or enrollment is limited under real life conditions. In these cases, RWD are an appropriate alternative or even superior to RCT data. The advantages and disadvantages of both kind of data is determined by the way these data are generated. On the one hand, RWE data are generated through the evaluation of effectiveness, safety and quality of care in settings and populations that are representative of practice including those not generally captured in traditional clinical trials. Typical study designs for RWE are prospective and retrospective observational studies, pragmatic trials, case-control studies and case series. On the other hand, RCT data are generated by highly artificial trials that are designed to precisely analyze the efficacy of one treatment option against another treatment

**Figure 3**




*The core factors for HTA submissions.*

option. The hallmark of RCTs is the randomization, which is the single most powerful approach to eliminate some of the most limiting bias potentials. Further ways to control different bias potentials are the blinding of study participants, study investigators and personnel analyzing the data and the enhancing of the probability of getting comparable patient collectives in both treatment groups by strict inclusion and exclusion criteria.

In summary, RCTs strife to enhance the internal validity – thereby ensuring that the treatment effect is captured precisely. However, there is a price attached to this increase in internal validity.

## RCTs have often a highly limited external validity

An analysis of the Rheumatoide Arthritis: Beobachtung der Biologika-Therapie (RABBIT) register has demonstrated that only 23–33 % of all patients in the register would be eligible for pivotal trials in the indication of rheumatoid arthritis because of the typical inclusion and exclusion criteria. In real life, patients with rheumatoid arthritis are more likely to be women, older, disabled or suffering from comorbidities than patients enrolled in clinical trials. Therefore, RCTs have often a highly limited external validity (fig. 4).

## RWE is more likely to have a high bias potential but a larger population sample

RWD are usually collected by prospective and retrospective observational cohort studies under real world conditions. Therefore, patients in RWE are often treated according to their individual needs and not according to a randomization scheme. This is reflected by a higher diversity of possible comparators and may even include off-label treatments, if these represent the current standard of treatment. Without strict inclusion and exclusion criteria, patients are more likely to represent real world conditions. Therefore, RWE analyses are usually targeting effectiveness and have a much higher external validity than standard RCTs. The major limitation of RWE analyses remains the bias potential. In general, a bias is any tendency, which prevents unprejudiced consideration of a question. In research, bias occurs when a systematic error is introduced into sampling or testing by selecting or encouraging one outcome or answer over others. Some bias, especially the selection bias and confounding, can be avoided or reduced by randomization. But randomization is not the only way to control a bias potential. In general, there are two methods to control the bias potential in RWE data analysis: stratification or matching. Advanced methods to reduce the bias potential in RWE analysis are, among others, the use of propensity scores, Gaussian processes or vector based machine learning. Each of these methods can be used to avoid a bias potential during the development of the study design or by sophisticated statistical analysis of the collected data. The propensity score is an advanced method based on logistic regression. There are multiple ways to use propensity scores before and after the study is conducted, which are summarized in fig. 5.

Vector based machine learning is an additional advanced approach that can be used. The vector is an arbitrarily complex Boolean combination of code lists, test results, time conditions and demographic characteristics that can be used to match or stratify the patient cohorts accordingly. However,


Pharm. Ind. 80, Nr. 4, 476–481 (2018)

since the vector based machine learning is based on a Boolean approach instead of a logistic regression, it is less prone to non-linearity and non-additivity compared to the propensity score. Therefore, it can be used successfully in cases where the use of the propensity score would be inappropriate or less reliable.

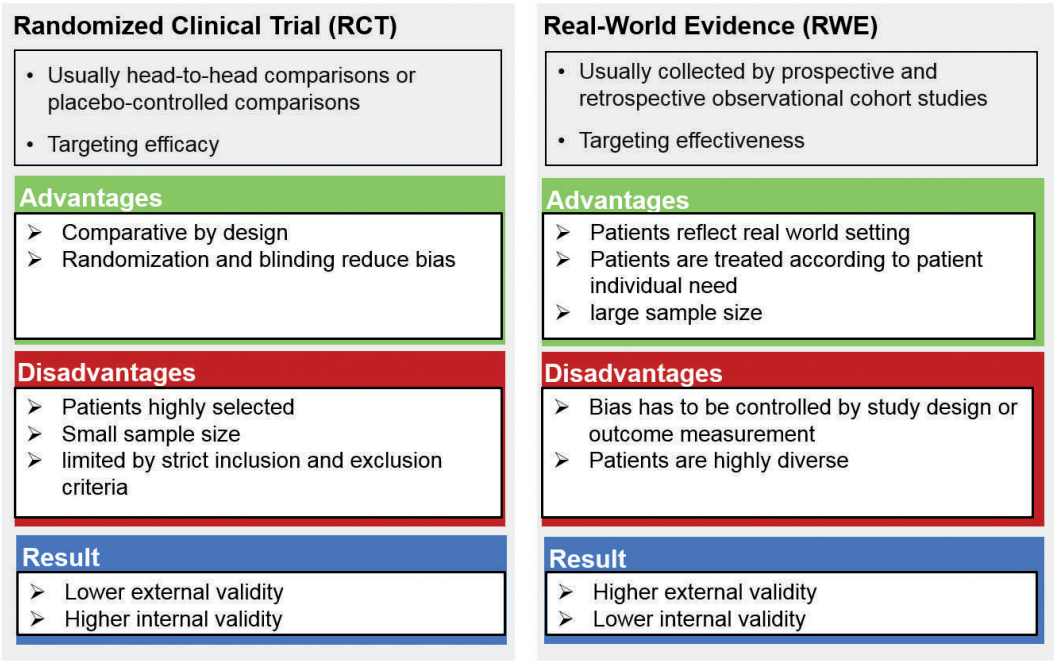
## RWD are suitable to analyze patient characteristics and treatment options under real world conditions

In summary, bias control is still the hallmark of RCTs, but advanced methods to handle known and unknown bias in RWD have greatly increased the reliability of RWE data in this field. Additionally, RWD are more suited to represent patients and treatment options under real world conditions, increasing the external validity of the analysis and by this the transferability of the results. All this is possible using usually much larger databases than standard RCTs.

## The combination of large data bases with advanced analytic techniques can be very powerful to detect benefits

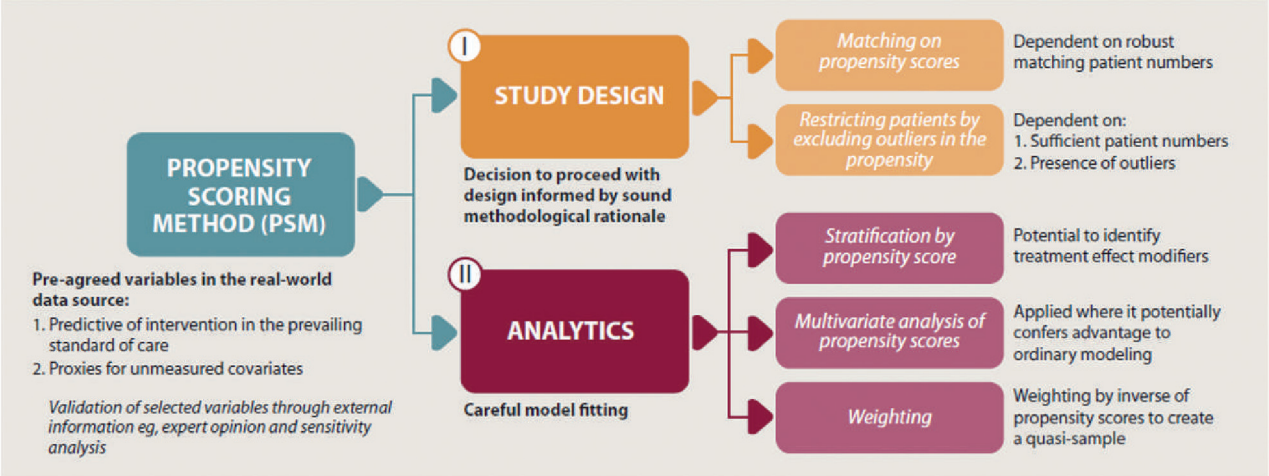
Finally, it has to be addressed which data can be used for RWE in HTA. The initial example of Zytiga has already highlighted, that RWD are not limited to one data source and meaningful conclusions can be derived from the combination of different data sources. Anonymized electronically collected longitudinal patient data – the electronic patient records (EPR) – are often the backbone of RWD analysis. These EPR contain a wide array of data, including demographics, diagnoses, physician's notes to diagnoses, therapies (with or without dosage information), relevant lab values, referrals, hospitalizations and sick notes. EPR are available from different European and non-European countries. In some countries, like the Nordic countries, the EPR can be combined with national register data to provide an even wider data basis. Similar powerful is the combination of EPR with national claims data to provide a bigger database for RWE analysis. In general, nearly any two or more databases can be used to create a data platform that is tailored for a specific research question. Only four components are required to build an evidence platform using different data sources. A cross functional view on the evidence gap marks the beginning of the building process. This view is important to identify the fit for purpose data sources that could be used to fill the data gap. These data sources can be own data, syndicated data or data from a 3rd party. After all relevant data sources are identified, the data need to be harmonized. This is usually achieved by an "anonymized identifier" that allows to link a specific patient in one database with a specific patient in another database. In anonymized data, this is most often accomplished by a mathematical approach. After combining the different data sources in one data mart, the evidence generation can start.

**Figure 4**




*Differences between RCT and RWE data.*

■ **Figure 5**




*Application of propensity score matching.*

## RWE can be generated retrospectively and prospectively. 'Pragmatic trials' are most likely to be accepted RWE in the future

RWE data are not limited to already existing data. An RWE approach can also be used to create new data. Among the different approaches, the pragmatic trial (pRCT) is most likely to be a perfect candidate for the use of RWD in HTA. Essentially, the pRCT combines the advantages of a full-fledged RCT with an observational study. Once physician and patient agreed to participate, treatment is assigned at random according to the protocol in a pRCT. However, the randomization usually only determines the new substance, in case the patient is assigned to the new treatment. If the patient is not assigned to the new treatment, he will be treated according to his individual patient needs. EPR are normally used for patient identification and centralized appointment scheduling for patient recruitment and enrollment. Usually, the patient is not examined in fixed visits during the pRCT. The patient will come according to his own needs. This procedure reduces the costs of a pRCTs greatly and improves the external transferability of the results, as the patient in the trial is most likely examined and monitored like a patient not participating in a trial.

## Direct-to-patient data collection for HTA submissions

Another powerful RWE database is direct-to-patient research. A direct-to-patient approach can be used in recruitment of patients, engagement of patients, data collection and follow-up tracking. For HTA submissions the direct-to-patient data collection is highly interesting, as most HTA authorities are especially interested in patient-reported outcomes and patient-centered outcomes. Patient reported outcomes are defined as reports that come directly from the patient about the status of a patient's health condition without amendment or interpretation of the patient's response by a clinician or anyone else. Patient centered outcomes are defined as outcomes that matter to patients. Usually patient centered outcomes include the improvement in the state of health, the shortening of the duration of illness, the extension of survival, the reduction of side effects or an improvement in the quality of life. Both patient-reported outcomes and patient centered outcomes can be used to enrich a data basis build on EPR.

## RWD use a bigger patient sample size compared to RCT, while providing access to a wide array of information

In summary, RWE data usually have a bigger patient sample size compared to RCT, while providing access to a wide array of information. Demographic, clinical and lab data are often available for the most important pathologies from a wide geographic coverage. RWD are not limited to one data basis and can include either multiple existing data sources or even new generated data. Altogether, this should clearly demonstrate that RWD can be used in HTA submissions, when the strengths and limitations of RWE data are considered and handled accordingly.

**Correspondence:**
Dirk Eheberg
IQVIA Commercial GmbH & Co OHG – HEOR
Landshuter Allee 10
80637 Munich (Germany)
e-mail: Dirk.Eheberg@iqvia.com

Chefredaktion: Claudius Arndt. Sekretariat: Gudrun Below. Verlag: ECV · Editio Cantor Verlag für Medizin und Naturwissenschaften GmbH, Baendelstockweg 20, 88326 Aulendorf (Germany). Tel.: +49 (0) 75 25 94 00, Fax: +49 (0) 75 25 94 01 80. e-mail: redaktion@ecv.de. http://www.ecv.de. Herstellung: rdzGmbH / Holzmann Druck GmbH & Co. KG. 