The following text is a translation of the German article
“Patientenregister für die Nutzenbewertung – Kein Ersatz für randomisierte Studien”
(Dtsch Arztebl 2017; 114(16): A-783 / B-662 / C-648)

Patient registries for benefit assessments – No replacement for randomized trials

Windeler J, Lauterberg J, Wieseler B, Sauerland S, Lange S.
Institute for Quality and Efficiency in Health Care (IQWiG), Cologne, Germany

Patient registry data are unsuitable for investigating cause-and-effect relationships and thus unsuitable for benefit assessments. Their otherwise undisputed potential can only be exploited if sophisticated quality requirements are met.

Lately, it has widely been suggested that questions regarding the benefit and harm of drugs, medical devices, and other medical interventions can be answered faster, more cost-efficiently, or even more credibly by means of analyses of so-called real-world data from routine data sources and medical registries than by means of clinical trials. Correspondingly, in the field of health policy, the number of those favouring greater use of such information and data sources for market authorization, funding, and policy decisions in health care seems to be increasing.

However, a glance at the more recent history of international health research, for example, the failed outcomes research projects in the United States [1,2], already shows that an overly optimistic view does not seem to be appropriate here.

What are registries?

The authors of the first standard guide on patient registries provide the following definition: *“…*a patient registry is an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves one or more predetermined scientific, clinical, or policy purposes” [3]. More precisely, registries are not studies, but patient-related data collections of widely differing kinds and purposes. Correspondingly, patient registries and their analyses are very diverse in practice. They can be roughly distinguished according to registry type, even though in some cases, types overlap. For instance, one distinguishes between disease-related registries [e.g. [4]] and procedure- or product-specific ones [e.g. [5,6]]. Although registries can pursue several aims at the same time, the main purpose can often be identified: for example, registries may largely serve epidemiological purposes (e.g. [7]), the assurance of treatment quality [e.g. [8,9]] or the recording of outcomes such as after exposure to dangerous substances [e.g. [10]].

Great expectations

Historically, disease-related registries (e.g. for tuberculosis) have been used to describe and examine epidemiological relationships and differences in the dissemination and course of specific diseases on a population basis. For some decades, there has also been an

increasing focus on other aims for patient registries. These aims include supplementary contributions to clinical research through (according to the assumption of registry advocates) the unfalsified and realistic representation of the safety and effectiveness of medical interventions in everyday healthcare. Further aims are the collection and analysis of data for the purposes of quality assurance, healthcare planning, and health economic evaluation. All of these aims are supposed to be achieved by the complete and unselected inclusion of patients using long observation periods, so that “natural” courses of diseases can be studied, important prognostic factors identified, and rare or late treatment complications or problems with products detected as safety signals [11]. In addition, through the (ideally unfiltered) broad inclusion of patients, registries are also supposed to offer the options of assessing treatment effects in everyday life, identifying subgroups of patients with particularly good or bad treatment outcomes, or gaining information on patient groups who are often not represented in clinical trials, such as geriatric and/or multi-morbid patients.

Quality criteria

In 2010, the German Network for Health Services Research (DNVF) formulated quality criteria for medical registries in a memorandum [12]. The authors stated that, for the largely voluntary procedures without a legal basis, sufficient registry quality with regard to data completeness and validity is only achievable if there is high acceptance among patients and reporting organizations. The fact that the establishment of high-quality registries [3,12,13] largely requires a run-up of several years and continuous improvement processes is shown by the experience in Scandinavian countries, which have established several well-functioning patient registries. The implementation of the clinical cancer registries in Germany shows the diverse challenges entailed in the creation of a national and complete mandatory registration process in a federal system, even in the event of a legal basis and regulated funding [14].

Data quality requirements not met

Incomplete and invalid data represent a major problem in patient registries. It is therefore not surprising that, for example, for the national complete data collection of the Danish hospital-based patient registry [15], a total of 114 studies have been conducted since 1978 solely to validate the data collected from over 8 million patients (e.g. on diagnoses and treatments). This was largely done by comparing the registry data with patient records, but also with other data sources. Despite clearly increasing data quality over the past years, with regard to validity and completeness, in this model project the consistency rates found between the registry and the sources used for validation ranged from under 15% to 100%, depending on the topic and disease.

Data gaps in patient registries are not random. They can thus cause substantial bias in results, leading to false conclusions. In this context, in registries with a consecutive inclusion of patients and a complete collection of patient data, it must first be asked whether all centres actually and verifiably included all patients. For instance, are severely ill patients underrepresented, as they could not provide the informed consent required for registration? Or are, as in the Berlin Myocardial Infarction Registry, geriatric female patients underrepresented, as instead of being treated in a cardiology ward they are often treated in other internal medical wards [16]? It is also important to ask whether data collection in the follow-up period also includes a careful documentation of drop-outs, comparable to that of clinical trials. This is important to ensure that sicker patients or patients with intolerabilities

and treatment complications do not drop out more often and earlier from the cohorts observed, without being noticed.

Unfortunately, not even a halfway complete overview of German patient registries exists and thus no overview of data management and quality assurance practices. A web-based registry portal [17] was to solve this problem; however, since its establishment unfortunately only one entry has been made. Only some of the registries publish how data collection, recording, storage, transfer, evaluation, and cleaning are performed. The same applies to data processing, analysis, and the presentation of results. In this context, registry experts and statisticians agree that the basic requirements to ensure data quality in registries and the resources needed should not differ from those of clinical trials, but that rather the essential study standards of good clinical practice (GCP) should also apply to registries [18].

No reliable assessment

Even though patient registries can deliver important contributions to epidemiological research, quality assurance and other purposes, as a rule they are unsuitable for the benefit assessment of medical interventions. This is because registries do not allow a valid comparison of results of patient groups receiving different treatments. Such a comparison can only be used to derive unbiased conclusions on the benefit and harm of an intervention if it can be ensured that all characteristics of both groups are equal, except for the test and control intervention applied. Only then can the differences observed actually be interpreted as a causal consequence of treatment. It is theoretically well-substantiated and generally accepted that only a random allocation to treatment groups (randomization) can ensure the required structural equality of the groups to be compared. Only then can the influence of known and unknown confounders, which leads to biased results, be eliminated. Patient groups to be compared that are generated from registry data are strongly susceptible to these types of bias and are thus unsuitable for robust benefit assessments.

This is because the type of treatment the patients included in the registry receive in everyday health care is determined by their physicians and themselves following a number of criteria that are rarely precisely documented, and sometimes documentation is hardly possible. For instance, older patients or patients with major concomitant diseases might be recommended less invasive or less high-risk treatments than younger ones. Or if psychotherapy is required, patients with a lower educational level or insufficient knowledge of the German language might be offered different treatments from those offered to academics. Typically these selection mechanisms are based on characteristics associated with a patient's prognosis (see Box).

In such a situation, which is typical of registries, one can try to apply statistical methods to rectify the results from confounders. But regardless of the fact that relevant confounders with a risk of bias regarding the comparison of interventions are not (yet) known in some clinical fields, highly relevant confounders are not documented in patient registries or documented only very inadequately, even if they are known. For instance, this applies to comorbidities in the mandatory data set of the German clinical cancer registry [19], meaning that later on comorbidities cannot be statistically controlled for as confounders. Putting more effort into the recording of data does not solve the problem either, because with the increasing number of confounders considered, the adjustment methods known would reach their limits (not so much the IT and technical limits as the mathematical and statistical ones). Cases where

registry data show treatment differences so large (“dramatic”) that they can no longer be explained solely by the impact of bias are very rare from an empirical point of view [20]; therefore this does not constitute a desirable aim for a registry.

Besides the main problem of the insufficiently controllable bias of results, important outcomes for a benefit assessment are often missing in registry data sets. For example, hardly any registries routinely collect data on patient-reported outcomes (PROs), such as data on quality of life or the return to normal social activities. But due to a lack of blinding, such data would in any event hardly be suitable for a benefit assessment.

In the debate about the benefit of product and patient registries, their advantage in detecting safety signals through long-term observations is often mentioned. In this context, the terms “signal” and “long-term” are important. To detect a signal, data from good-quality registries are sufficient, and for long-term courses (e.g. durability of prostheses) there is often no other option (and if the prosthesis breaks, the issue of causality is clear). It is then important that at least adjusted comparisons between treatment types and products are conducted as, for example, in the case of the Scandinavian endoprosthesis registries [21,22]. However, as with proof of benefit, RCTs are also the most suitable instrument to demonstrate proof of causality for deficiencies and harm.

Registry-based RCTs

The idea of conducting RCTs in combination with registries in order to bundle their respective advantages and to compensate weaknesses is not new [23-26]. But only recently have there been more practical tests, mainly with Scandinavian registries [27-30], or corresponding plans such as those for the German-Austrian CLIP-ID-AS2 study in the field of infection prevention [31]. It is currently also being proposed to follow this path systematically for post-marketing studies, for example in oncology [32], instead of relying on uncertain “real world data”. In this context, it seems pointless to argue about whether RRCTs are a pragmatic variant of RCTs or a special variant of use for high-quality patient registries. It is more important that in fact a number of strengths of this research approach can be identified. Under the prerequisite that both the patient registry and the RCT meet high quality standards, this includes the options to

- reliably determine the cause-and-effect relationships between treatment and treatment results by means of the RCT population in the registry,
- achieve important practical research simplifications (e.g. for infrastructure, recruitment planning, data usage) and cost savings for the RCTs,
- identify questions requiring further analysis by comparing the RCT population and the remaining registry population,
- conduct long-term observations of intervention effects and risks by means of the RCT population in the registry, as well as
- identify clinically-relevant subgroups in relation to benefit and harm.

For instance, in the TASTE study [27,29], 7244 of about 12,000 registered patients with acute ST-elevation myocardial infarction were included in a randomized trial in a relatively short period of time. On the basis of 30-day and 1-year mortality, an additional catheter-based thrombus aspiration was compared with coronary angioplasty alone – the study

showed equivalent outcomes in both treatment groups. Because in Sweden, all patients with acute myocardial infarction are included in a publicly-funded mandatory registry (SWEDEHEART) with about 150 variables, this industry-independent registry-based RCT could be conducted in 30 centres for less than 10% of the usual costs of a stand-alone RCT [29].

However, although discussed as a potentially ground-breaking technology for clinical research [33], this RRCT approach also has limitations, which are related to the extremely limited availability of suitable patient registries [34]. But they are also related to the fact that all challenges linked to patient randomization basically remain. Even advocates of RRCTs [35] support the conduct of traditional RCTs in cases where, for example, new drugs and medical devices require close and comprehensive safety monitoring with a complex accompanying recording of data and findings. The same applies in cases where a highly reliable recording and confirmation of strictly defined outcomes is required (e.g. by an adjudication committee), which presumably cannot be recorded with sufficient reliability in the routine use of registries. The registry-based RCT approach thus seems to be particularly well-suited for relatively simple clinical questions.

Conclusion

Patient registries are not a convenient short-cut for medical benefit assessments. If robust evidence on the benefit and harm of medical interventions is required and important decisions with an impact on patient safety and health economics are to be made, then medical registries cannot in practice offer a sufficiently reliable information base. This is due to basic methodological reasons and a lack of registry availability and quality. The promotion and conduct of prospective RCTs, which could also be registry-based to answer certain questions, as well as the consideration of their results, thus remain a priority for benefit assessments in the evidence-based health care of the population.

Box

Current example of a patient registry

A US study [36] aimed to answer the question as to whether tracheal intubation during in-hospital emergency treatment for cardiac arrest was associated with the likelihood of survival. For this purpose, for the years 2000 to 2014 the authors retrospectively analysed the data of a large cohort of 108,079 patients collected in a registry established for quality assurance purposes by 668 hospitals. Using a time-dependent propensity score matching procedure, the investigators compared statistically generated pairs of patients with multiple and very similar characteristics who had either been intubated or not in the same minute of the resuscitation procedure (1-15 minutes). With a rate of 16.3%, overall survival was lower among intubated patients than among non-intubated ones (19.4%). Likewise, the results for other outcomes were more favourable for non-intubated patients. Can the recommendation be inferred from these results that (early) tracheal intubation should be dispensed with for in-hospital emergency events involving cardiac arrest? The editor's comment on this study [37] is considerably more cautious in this regard than the authors' conclusion. This is because the study's findings show various problems of interpretation. The findings could also have been caused by "confounding by indication", meaning that patients intubated (early) may have greater severity of illness in the first place, which is not sufficiently depicted by the registry data. Less sick patients in whom the emergency team delay intubation may potentially achieve spontaneous circulation before being subjected to the multiple potential complications of mechanical ventilation. In addition, these analyses of registry data do not allow conclusions on whether emergency physicians who tend to intubate early also tend to use more invasive and potentially more harmful interventions in the further emergency management of patients. Derek Angus, the author of the editorial, therefore concludes that this highly relevant clinical question should be answered prospectively by means of RCTs, potentially in the form of RRCTs, on the basis of the existing registry.

References

1. Anderson C. Measuring what works in health care. Science 1994; 263(5150): 1080, 1082.

2. Deyo RA. Promises and limitations of the patient outcome research teams: the low-back pain example. Proc Assoc Am Physicians 1995; 107(3): 324-328.

3. Gliklich R, Dreyer N, Leavy M. Registries for evaluating patient outcomes: a user´s guide; 3rd edition [online]. 04.2014 [Accessed: 18 November 2016]. URL: http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=1897.

4. Viomedo. DZHK Translationales Kardiomyopathie-Register [online]. 24.03.2016 [Accessed: 14 November 2016]. URL: https://www.viomedo.de/klinische-studien/3686/dzhk-translationales-kardiomyopathie-register.

5. Deutsche Gesellschaft für Allgemein- und Viszeralchirurgie. StuDoQ|NOTES: nationales NOTES-Register der DGAV [online]. [Accessed: 14 November 2016]. URL: http://www.dgav.de/notes/.

6. EPRD: das Endoprothesenregister [online]. [Accessed: 14 November 2016]. URL: http://www.eprd.de/.

7. Kaatsch P, Spix C. German Childhood Cancer Registry: annual report 2015 (1980-2014) [online]. [Accessed: 18 November 2016]. URL: http://www.kinderkrebsregister.de/typo3temp/secure_downloads/22605/0/f474d594c6b5a8805c4e629db249872e05d69ddb/jb2015_s.pdf.

8. Berliner Herzinfarktregister [online]. [Accessed: 14 November 2016]. URL: http://www.herzinfarktregister.de/.

9. Akademie der Unfallchirurgie. TraumaRegister DGU [online]. [Accessed: 14 November 2016]. URL: http://www.traumaregister-dgu.de/de/startseite_tr.html.

10. Deutsche Gesetzliche Unfallversicherung. Zentrale Expositionsdatenbank (ZED) [online]. [Accessed: 14 November 2016]. URL: http://www.dguv.de/ifa/gestis/zentrale-expositionsdatenbank-(zed)/index.jsp.

11. Buhr HJ, Holmer C, Klinger C, Lehmann KS. Nutzen versus Überinterpretation: am Beispiel der Register der GGAV/StuDoQ. Zentralbl Chir 2015; 140(02): 133-138.

12. Müller D, Augustin M, Banik N, Baumann W, Bestehorn K, Kieschke J et al. Memorandum Register für die Versorgungsforschung. Gesundheitswesen 2010; 72(11): 824-839.

13. Zalatel M, Kralj M. Methodological guidelines and recommendations for efficient and rational governance of patient registries. Ljubljana: National Institute of Public Health; 2015.

14. Klinkhammer-Schalke M, Hofstädter F, Gerken M, Benz S. Der Beitrag Klinischer Krebsregister für die Nutzenbewertung: Anforderungen und erste Ergebnisse. Z Evid Fortbild Qual Gesundhwes 2016; 112(Suppl 1): S3-S10.

15. Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sorensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol 2015; 7: 449-490.

16. Maier B, Gothe H, Kieschke J. Registerdaten. In: Swart E, Ihle P, Gothe H, Matusiewicz D (Ed). Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse; Grundlagen, Methoden und Perspektiven. Bern: Huber; 2014. p. 234-244.

17. Universitätsmedizin Mainz. metaRegDe: Verzeichnis der medizinischen Register in Deutschland [online]. [Accessed: 15 November 2016]. URL: https://metareg.imbei.de/home.

18. Wegscheider K. Medizinische Register: Nutzen und Grenzen. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2004; 47(5): 416-421.

19. Bundesministerium für Gesundheit. Aktualisierter einheitlicher onkologischer Basisdatensatz der Arbeitsgemeinschaft Deutscher Tumorzentren e.V. (ADT) und der Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID). Bundesanzeiger Amtlicher Teil 28.04.2014: B2.

20. Pereira TV, Horwitz RI, Ioannidis JP. Empirical evaluation of very large treatment effects of medical interventions. JAMA 2012; 308(16): 1676-1684.

21. Varnum C, Pedersen AB, Kjaersgaard-Andersen P, Overgaard S. Comparison of the risk of revision in cementless total hip arthroplasty with ceramic-on-ceramic and metal-on-polyethylene bearings. Acta Orthop 2015; 86(4): 477-484.

22. Varnum C, Pedersen AB, Makela K, Eskelinen A, Havelin LI, Furnes O et al. Increased risk of revision of cementless stemmed total hip arthroplasty with metal-on-metal bearings. Acta Orthop 2015; 86(4): 491-497.

23. Feit F, Brooks MM, Sopko G, Keller NM, Rosen A, Krone R et al. Long-term clinical outcome in the Bypass Angioplasty Revascularization Investigation Registry: comparison with the randomized trial. Circulation 2000; 101(24): 2795-2802.

24. Olschewski M, Scheurlen H. Comprehensive Cohort Study: an alternative to randomized consent design in a breast preservation trial. Methods Inf Med 1985; 24(3): 131-134.

25. Pocock SJ. The combination of randomized and historical controls in clinical trials. J Chronic Dis 1976; 29(3): 175-188.

26. Sacristan JA, Soto J, Galende I, Hylan TR. Randomized database studies: a new method to assess drugs' effectiveness? J Clin Epidemiol 1998; 51(9): 713-715.

27. Frobert O, Lagerqvist B, Gudnason T, Thuesen L, Svensson R, Olivecrona GK et al. Thrombus Aspiration in ST-Elevation myocardial infarction in Scandinavia (TASTE trial): a multicenter, prospective, randomized, controlled clinical registry trial based on the Swedish angiography and angioplasty registry (SCAAR) platform; study design and rationale. Am Heart J 2010; 160(6): 1042-1048.

28. Gotberg M, Christiansen EH, Gudmundsdottir I, Sandhall L, Omerovic E, James SK et al. Instantaneous Wave-Free Ratio versus Fractional Flow Reserve guided intervention (iFR-SWEDEHEART): rationale and design of a multicenter, prospective, registry-based randomized clinical trial. Am Heart J 2015; 170(5): 945-950.

29. Lagerqvist B, Fröbert O, Olivecrona GK, Gudnason T, Maeng M, Alström P et al. Outcomes 1 year after thrombus aspiration for myocardial infarction. N Engl J Med 2014; 371(12): 1111-1120.

30. Rao SV, Hess CN, Barham B, Aberle LH, Anstrom KJ, Patel TB et al. A registry-based randomized trial comparing radial and femoral approaches in women undergoing percutaneous coronary intervention: the SAFE-PCI for Women (Study of Access Site for Enhancement of PCI for Women) trial. JACC Cardiovasc Interv 2014; 7(8): 857-867.

31. Charité Universitätsmedizin Berlin. Dreiarmige Cluster-randomisierte Dekolonisationsstudie im Rahmen des Gesamtvorhabens "Der Einfluss von Klimaveränderungen auf das Auftreten von Infektionen in Deutschland und Untersuchungen zur Dekolonisation" [online]. In: Deutsches Register Klinischer Studien. [Accessed: 17 November 2016]. URL: https://drks-neu.uniklinik-freiburg.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00010475.

32. Grieve R, Abrams K, Claxton K, Goldacre B, James N, Nicholl J et al. Cancer drugs fund requires further reform. BMJ 2016; 354: i5090.

33. Lauer MS, D'Agostino RB Sr. The randomized registry trial: the next disruptive technology in clinical research? N Engl J Med 2013; 369(17): 1579-1581.

34. Li G, Sajobi TT, Menon BK, Korngut L, Lowerison M, James M et al. Registry-based randomized controlled trials: advantages, challenges and areas for future research. J Clin Epidemiol 20.08.2016 [Epub ahead of print].

35. James S, Rao SV, Granger CB. Registry-based randomized clinical trials: a new clinical trial paradigm. Nat Rev Cardiol 2015; 12(5): 312-316.

36. Andersen LW, Granfeldt A, Callaway CW, Bradley SM, Soar J, Nolan JP et al. Association Between Tracheal Intubation During Adult In-Hospital Cardiac Arrest and Survival. JAMA 2017; 317(5): 494-506.

37. Angus DC. Whether to Intubate During Cardiopulmonary Resuscitation: Conventional Wisdom vs Big Data. JAMA 2017; 317(5): 477-478.