Replik zum DÄB-Artikel „Kein Ersatz für randomisierte Studien“ (DÄB 16/2017)

Aufruf zur differenzierteren Betrachtungsweise

„Wenn belastbare Evidenz für Nutzen und Schaden medizinischer Verfahren benötigt wird, und relevante Entscheidungen mit Auswirkungen auf Patientensicherheit und Ökonomie zu treffen sind, können medizinische Register aus prinzipiellen methodischen Gründen und mangels Verfügbarkeit und Qualität auch praktisch keine ausreichend sichere und verlässliche Informationsgrundlage bieten“, schreibt ein Autorenteam rund um IQWiG-Leiter Prof. Dr. Jürgen Windeler im „Deutschen Ärzteblatt“ (16/2017). Dem widersprach das Deutsche Netzwerk Versorgungsforschung (DNVF) auf seiner 22. Mitgliederversammlung in Berlin.

Artikelangaben

Windeler et al., „Kein Ersatz für randomisierte Studien“, in Dtsch Arztebl 2017; 114 (16): A 783–6.

Replik

„Aufruf zur differenzierteren Betrachtungsweise“ in MVF (03/17), S. 16f., doi: 10.24945/MVF.03.17.1866-0533.2016

Literatur

1: DNVF-Memorandum „Register für die Versorgungsforschung“, in „Gesundheitswesen“ 2010; 72(11): 824-839. doi: 10.1055/s-0030-1263132.
2. www.dkrs.de und www.germanctr.
3: Registry of Patient Registries (RoPR, https://patientregistry.ahrq.gov/) der Agency for Healthcare Research
4: www.dgav.de/studoq.html
5: Beteiligt sind die AG Register des DNVF und die AG IT-Infrastruktur und Qualitätsmanagement (AG ITQM) der Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. (TMF)
6. Neugebauer, E.A.M, Stausberg, J.: „Was Register leisten können und was nicht“, in „Unfallchirurg“ 2016, 119:493–500. doi: 10.1007/s00113-016-0176-2
7: Wellner et al. Trials (2017) 18:163, in „Gesundheitswesen 2010“; ebenso bei der AG Register auf der Webseite des DNVF nachlesbar (www.dnvf.de). doi 10.1186/s13063-017-1911-x
8: www.evidencelive.org/manifesto
9: Arunachalam, Lakshmanan MD, FRCS; Hunter, Iain A. FRCS, PhD; Killeen, Shane FRCSI, MD: „Reporting of Randomized Controlled Trials With Statistically Nonsignificant Primary Outcomes Published in High-impact Surgical Journals“, in Annals of Surgery: 06/17. doi: 10.1097/SLA.0000000000001795

>> Windeler und seine Mitautoren pointieren in dem Artikel im Ärzteblatt, dass „in letzter Zeit der Eindruck erweckt“ werde, dass mithilfe von Analysen sogenannter „real world data“ aus Routinedatenbeständen und medizinischen Registern Fragen nach Nutzen und Schaden von Arzneimitteln, Medizinprodukten und anderen medizinischen Interventionen „schneller, kostengünstiger oder gar glaubwürdiger beantwortet werden können als mit klinischen Studien“. Unvollständige und nicht valide Daten stellen nach Ansicht der Autoren „ein Hauptproblem von Patientenregistern“ dar, wobei „Datenlücken in Patientenregistern nicht zufällig“ seien und „in erheblichem Maße zu einer Ergebnisverzerrung und somit falschen Schlussfolgerungen“ beitragen.

Als Fazit kommt Windeler zu der Ansicht, dass „medizinische Register aus prinzipiellen methodischen Gründen und mangels Verfügbarkeit und Qualität auch praktisch keine ausreichend sichere und verlässliche Informationsgrundlage bieten“, wenn „belastbare Evidenz für Nutzen und Schaden medizinischer Verfahren benötigt wird, und relevante Entscheidungen mit Auswirkungen auf Patientensicherheit und Ökonomie zu treffen sind“.

Das ist aber nur eine Seite der Medaille, und wohl auch eine recht dogmenbehaftete. Die Problemlage indes ist vielschichtiger. Der Grund: Es gibt sicher derzeit noch mehr methodisch schlechte Register als methodisch schlechte RCT, was sich einerseits aus der reinen Zeit der Beschäftigung mit diesen Instrumenten begründet, andererseits mit den noch recht jungen Bemühungen, die wissenschaftlichen Grundlagen für qualitativ hochwertige Register zu schaffen und zu publizieren; eine Arbeit, die hauptsächlich vom Deutschen Netzwerk für Versorgungsforschung (DNVF) geleistet wird.

So wurde bereits 2010 ein breit konsentiertes Memorandum „Register für die Versorgungsforschung“ in der Zeitschrift „Gesundheitswesen“ (1) veröffentlicht. Darin steht unter anderem zu lesen, „dass mit der vorliegenden Publikation zwei Ziele verfolgt werden: sie soll einerseits Entwicklern eines Registers als Leitfaden in Hinblick auf eine hohe Qualität des Registers dienen sowie andererseits potenziellen Nutzern von Daten aus Registern die Bewertung der Qualität eines Registers ermöglichen“.

Was indes immer noch fehlt, ist die breite Auffindbarkeit und damit die Vergleichbarkeit von Registern. Dazu sind „international erste Bestrebungen zu erkennen“, ähnliche Angebote wie bei klinischen Studien („Deutsches Register klinischer Studien“ (2) auch für Register (3) zu etablieren. Unter anderem hat die Deutsche Gesellschaft für Allgemein-und Viszeralchirurgie (DGAV) bereits 2012 das Studien-, Dokumentations- und Qualitätszentrum DGAV-StuDoQ eingerichtet (4). Auch haben DNVF und TMF eine Initiative zur Etablierung eines Registerportals in Deutschland (5) ins Leben gerufen, wie Univ.-Prof. Dr. Prof. h.c. Dr. h.c. Edmund A.M. Neugebauer und Prof. Dr. Jürgen Stausberg 2016 in ihrem Beitrag „Was Register leisten können und was nicht“ (6) schrieben. Dort steht aber auch, das bis jetzt – eben durch die fehlenden Registerportale – „die Kenntnis zu abgeschlossenen, laufenden oder geplanten Registern weitgehend auf Zufall“ basiere, zudem auf persönlicher Kenntnis, Veröffentlichungen in wissenschaftlichen und nichtwissenschaftlichen Zeitschriften oder schlichtweg der Internetsuche.

Dieser Mangel, so Neugebauer und Stausberg in ihrem Beitrag, fördere Parallelentwicklungen, behindere die Translation von Erkenntnissen aus Registern in Forschung und Versorgung, erschwere die aktive Kontaktaufnahme von potenziellen Studienzentren und Probanden mit Registern und verhindere einen fachlichen Austausch zwischen Registerbetreibern. Hinzu komme, dass es aufgrund der großen Heterogenität eine große Spannweite bezüglich der Qualität der verschiedenen Register gebe, weshalb „eine globale Aussage zur Frage, was Register leisten können und was nicht, „ohne Berücksichtigung ihrer Qualität nicht möglich“ sei.

Nun muss aber mit der neuen Entwicklung der registerbasierten randomisierten Studien (RRCT) im Besonderen auch die Registerqualität nachgewiesen werden. Darum wurde am Beispiel des nationalen Registers für Qualitätskontrolle, Risikoassessment und Outcomeforschung in der Pankreaschirurgie (DGAV SuDoQ|Pancreas) erstmals eine systematische Bewertung der Registerqualität auf Basis der Empfehlungen des Methodenmemorandums Register des DNVF durchgeführt (7). Dieses Vorgehen ist enorm wichtig, da sich „erst bei Erfüllung qualitativer Mindeststandards“ belastbare Aussagen aus der Analyse von Registerdaten ableiten lassen, wie Neugebauer und Stausberg in ihrem Artikel betonen – was ihrer Meinung im Übrigen „in gleicher Weise auch für randomisierte Studien“ gelte.

Doch deren überlegene Qualität scheinen IQWiG-Leiter Windeler und sein Autorenteam als schlichtweg gegeben hinzunehmen. „There are huge shortcomings in the way that evidence based medicine operates today: bad quality research, evidence that is withheld, piecemeal dissemination, a failure to respect patients priorities and more“, hält dem Carl Heneghan, Professor of Evidence-Based Medicine und Direktor des angesehenen Centre for Evidence-Based Medicine (CEBM) der University of Oxford entgegen, dessen Gründer niemand anderes als der 2015 verstorbene David Sackett war.

Aktuell hat Heneghan das „EBM 2.0 Manifesto“ (8) verfasst, in dem er 20 Gründe aufzählt, warum eine bessere Evidenz nötig sei und warum die bisher verfügbaren Instrumente und Methoden eben nicht ausreichen, die aktuellen Gesundheitsprobleme zu lösen. Auf Heneghans Liste stehen:

- 1. Publication bias
- 2. Poor quality research
- 3. Evidence production
- 4. Research more likely to be false than true
- 5. Reporting bias
- 6. Ghost authorship
- 7. Financial and non financial conflicts of interest
- 8. Estimating costs of new treatments
- 9. Under reporting of harms
- 10. Delayed withdrawal of harmful drugs
- 11. Lack of Shared Decision Making strategies
- 12. Trials lacking external validity
- 13. Regulatory failings
- 14. Criminal behaviour
- 15. Rise of surrogate outcomes
- 16. Unmanageable volume of evidence
- 17. Clinical guidelines beset by major structural problems
- 18. Too much medicine
- 19. Prohibitive costs of drug trials
- 20. Trials stopped early for benefit

Seine 20 Kritikpunkte an EBM und damit auch an RCT führen ihn zu folgendem Fazit: „Equally worrying, the growth and volume of evidence has been accompanied by a corrosion in the quality of evidence, which has compromised medicine's ability to provide affordable, effective, high value care.“

Zu einem ähnlichen Schluss kommt auch ein Autorenteam rund um Lakshmanan Arunachalam MD (9), der in seinem 2017 in „Annals of Surgery“ veröffentlichen Beitrag „Reporting of Randomized Controlled Trials With Statistically Nonsignificant Primary Outcomes Published in High-impact Surgical Journals“ erklärt: „In RCT mit statistisch nicht signifikanten primären Ergebnissen, die in chirurgischen Zeitschriften veröffentlicht wurden, war die Berichterstattung und Interpretation der Befunde häufig mit den Ergebnissen unvereinbar.“

Doch alleine schon die von Heneghan angemahnte „Korrosion in der Qualität der Evidenz“ lässt erahnen, dass der von Windeler et al. konstruierte Gegensatz zwischen randomisierten, kontrollierten Studien und Registern wohl derzeit nicht das wichtigste Thema ist, auch weil – wie Neugebauer und Stausberg konstatieren – diese Herangehensweise „den jeweiligen Stärken und Schwächen nicht gerecht wird“. Vielmehr gehe es doch darum, dass „eine schlechte Qualität der Daten bei beiden Methoden die Ergebnisse verfälschen“ können. Nur sei es eben lückenhaft, „dies nur in Bezug auf Register anzusprechen“. <<

Heneghans Top 10 (www.evidencelive.org/manifesto)

1) Publication bias

In 1995 Sterling reported nothing had changed in the last 30 years when it came to publication bias: in his 1959 study of the four major psychology journals publications he reported that 97% of published studies were statistically significant.

- Publication decisions revisited – the effect of the outcome of statistical tests on the decision to publish and vice-versa. Am Stat 1995; 49:108–12.
- A 2010 HTA systematic review reported that half of all trials had never published results, and positive trials were twice as likely to be published than negative trials.
- Dissemination and publication of research findings: an updated review of related biases. Health Technology Assessment 2010; Vol. 14: No. 8

2) Poor quality research

In 1994 researchers commonly used the wrong techniques, or the right techniques wrongly, misinterpreted results, reported results selectively and often came to unjustified conclusions. All of this meant “we need less research but better research.”

- The scandal of poor medical research BMJ 1994; 308 :283
- In 2009, a Lancet series on avoidable waste in research estimated not much had changed: 85% of research spending currently goes to waste.
- Avoidable waste in the production and reporting of research evidence The Lancet; 374, 9683, 86-89

3. Evidence production problems

Bad Pharma collated countless problems with the production of evidence: many medical tests and trials are profoundly flawed, evidence is often hidden by drug companies to the detriment of patient care. Add to this is a poor regulatory system then it is increasingly clear we need a radically different system than the current defective system for producing evidence.

- How Drug Companies Mislead Doctors and Harm Patients. Macmillan 2012.

4. Research more likely to be false than true

John Ionnadis pointed out research is more likely to be false, particularly when effects are small, when many outcomes are presented and when there are greater financial interests. Moreover, “with increasing bias, the chances that a research finding is true diminish considerably”.

- Why Most Published Research Findings Are False. PLoS Med 2(8): e124.

5. Reporting bias

An empirical analysis of 102 randomised trials found that half of the efficacy and two thirds of the harm outcomes were incompletely reported, with statistically significant outcomes being more than twice as likely to be reported.

- Impirical Evidence for Selective Reporting of Outcomes in Randomized Trials: JAMA. 2004;291(20):2457-2465.
- The COMPare Projects analysis of outcome switching in clinical trials of the top five medical journals reports little has changed: on average each trial in the cohort only reported 62% of its specified outcomes.
- Effects of reporting bias include overestimation of efficacy and underestimation of safety and is a widespread phenomenon.
- Reporting bias in medical research – a narrative review. Trials. 2010 Apr 13;11:37. doi: 10.1186/1745-6215-11-37

6. Ghost authorship

A substantial proportion of trials have ghost authorship, which often go undisclosed, and undermine the validity of the results.

- Guest authorship and ghostwriting in publications related to rofecoxib: a case study of industry documents from rofecoxib litigation. JAMA. 2008, 299 (15): 1800-12. 10.1001/jama.299.15.1800.
- The promotion of gabapentin: an analysis of internal industry documents. Annals of Internal Medicine. 2006, 145 (4): 284-93.
- Constraints on publication rights in industry-initiated clinical trials. JAMA. 2006, 295 (14): 1645-6.
- Ghost authorship in industry-initiated randomised trials. PLoS Med. 2007; 4: e19

7. Financial and non financial conflicts of interest

Financial and non financial conflicts of interests are widespread amongst academic institutions and researchers and is associated with pro industry conclusions, restrictions on publication and data sharing and “private interests”.

- Scope and impact of financial conflicts of interest in biomedical research: a systematic review. JAMA. 2003 Jan 22-29;289(4):454-65.
- Making Sense of Non-Financial Competing Interests The PLoS Medicine Editors September 30, 2008

8. Estimating costs of new treatments

Analysis of 32 Cancer drugs reported 2014 drug costs were on average six times higher than those in 2000: costing on average $11,325 compared to the $1,869 per month in 2000.

- Drug Pricing Trends for Orally Administered Anticancer Medications Reimbursed by Commercial Health Plans, 2000-2014. JAMA Oncol. 2016;2(7):960-961.

9. Under reporting of harms

86% of 92 Cochrane reviews did not include data from the main harm outcome and the primary harm outcome was inadequately reported in 76% of the 931 included trials in these reviews.

- Selective reporting bias of harm outcomes within studies: findings from a cohort of systematic reviews BMJ 2014; 349 :g6501

10. Delayed withdrawal of harmful drugs

Analysis of 462 medicinal products withdrawn from the market found only 43 (9%) were withdrawn worldwide, and the interval between the first reported adverse reaction and year first withdrawal was a median 6 years.

- Post-marketing withdrawal of 462 medicinal products because of adverse drug reactions: a systematic review of the world literature. BMC Med. 2016 Feb 4;14:10.