



IQWiG im Dialog 2019

„Liefere Kausalmodelle Belege für kausale
Zusammenhänge?“

Abstracts der Präsentationen



Vanessa Didelez (Leibniz-Institut für Präventionsforschung und Epidemiologie (BIPS),
Fachbereich Mathematik/Informatik der Universität Bremen)

Die Prinzipien und Methoden der kausalen Inferenz

In der kausalen Inferenz geht es darum, Daten anhand geeigneter Modelle und Methoden mit dem Ziel zu analysieren, Schlüsse über Interventionseffekte zu ziehen – für hypothetische (hoher versus niedriger BMI) oder auch ganz konkrete Interventionen (Behandlung mit Medikament A versus B). Gerade bei Beobachtungsstudien ist dieses Ziel gewagt, da solche Studien im Gegensatz zu RCTs die entsprechende Intervention nicht tatsächlich durchführen. Die Beobachtungsdaten selbst sowie die oft standardmäßig angewandten, aber für diese Aufgabe ungeeigneten statistischen Methoden bergen dementsprechend eine Reihe von möglichen Biasquellen, wovon Confounding die offensichtlichste, aber nicht die einzige ist.

In diesem Vortrag gebe ich eine Übersicht und einen Vergleich verschiedener Ansätze der kausalen Inferenz. Dabei werden die folgenden Aspekte beleuchtet:

1. Überlegungen zur Wahl des Maßes des Interventionseffekts ("target of inference" oder "estimand"), z.B. average causal effect, effect of treatment on the treated, principal stratum effect, dynamische Strategien, marginale / bedingte Effekte;
2. welche Art von Annahmen zur prinzipiellen Identifizierbarkeit des gewählten Interventionseffekts erfüllt sein müssen;
3. gängige Modelle: von Strukturgleichungsmodellen zu marginalen strukturellen Modellen;
4. Vor- und Nachteile gängiger Schätzverfahren: G-Computation, Propensity Score Matching, Inverse Probability Weighting, G-Estimation und doppelt robuste Schätzung;
5. spezifische Probleme in (den eigentlich üblichen) Situationen mit zeitabhängiger Behandlung / Exposition bzw. Outcome.

Als Fazit ergibt sich, dass kein Ansatz jemals die Gültigkeit von kausalen Schlüssen garantieren kann, da diese immer von unüberprüfbaren Annahmen abhängt; dass aber explizit kausale Modellierung hilft, eine systematische und in sich stimmige statistische Analyse durchzuführen, die vermeidbaren Bias tatsächlich vermeidet und unvermeidbaren Bias zumindest plausibel minimiert.

Oliver Kuß (Deutsches Diabetes-Zentrum (DDZ), Düsseldorf)

Randomisierte vs. nicht randomisierte Studien: Evidenz aus der Meta-Epidemiologie

Es ist von beträchtlichem klinischen, ökonomischen und erkenntnistheoretischen Interesse, ob nicht randomisierte Studien auch für Therapieentscheidungen herangezogen werden könnten oder sogar sollten. Dieses Interesse widerspiegelnd, liegt eine Reihe von Studiendesigns und die entsprechende Evidenz daraus vor, ob, und falls ja, wie sich Ergebnisse aus randomisierten und nicht-randomisierten Studien unterscheiden. Im Vortrag werden diese Designs, die durchgeführten Studien und die gewonnenen Erkenntnisse daraus vorgestellt und diskutiert.

Ann-Kristin Leuchs (Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM), Bonn)

Kausale Inferenz und Estimands in der Arzneimittelzulassung

Grundlage für die Arzneimittelzulassung ist der kausale Wirksamkeitsnachweis eines neuen Arzneimittels; d. h. ‚Wie wirkt das neue Arzneimittel in einem wohldefinierten Setting im Vergleich zu einer Situation, in der dieses Arzneimittel nicht verschrieben wurde?‘

Um kausale Schlüsse über die Wirksamkeit neuer Arzneimittel zu ziehen, werden als Goldstandard randomisierte klinische Studien durchgeführt. Die Randomisierung gewährleistet eine Vergleichbarkeit der Gruppen zu Baseline, allerdings können nach der Randomisierung auftretende Ereignisse – wie Behandlungsabbrüche (z.B. aufgrund von Nebenwirkungen) oder die Einnahme von Notfallmedikation (z.B. aufgrund von fehlender Wirksamkeit) – zu Imbalancen führen und die Kausalitäts-/Wirksamkeitsbewertung erschweren. In diesem Kontext ist es wesentlich den zu schätzenden Behandlungseffekt (d. h. den Estimand) in Bezug auf potentiell nach Randomisierung auftretende Ereignisse klar zu definieren.

Nicht zuletzt nach Veröffentlichung des Draft ICH E9 Addendums zu Estimands und Sensitivitätsanalysen im Jahr 2017 sind Diskussionen über relevante Estimands in klinischen Studien in den letzten Jahren ein wesentlicher Bestandteil der Bewertung von Arzneimittelzulassungsanträgen und häufiges Thema der wissenschaftlichen Beratung in der Entwicklungsphase neuer Arzneimittel geworden.

In wie weit unterschiedliche Estimands kausal interpretiert werden können und somit Grundlage von Zulassungsentscheidungen sein können wird im Rahmen dieses Vortrages diskutiert. Zudem sollen die zur Schätzung der verschiedenen Estimands nötigen Annahmen und resultierenden Schlussfolgerungen für die Arzneimittelzulassung diskutiert werden.

Lisa Schell, Anke Schulz, Konstanze Angelescu, Dominik Glinz, Marco Knelangen & Stefan Sauerland (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), Köln)

Real Biased Data? Ein Vergleich von "Real World" Daten und RCTs in der Telekardiologie

Sind "Real World" Daten – mit adäquater Confounderkontrolle – genauso valide wie Ergebnisse von randomisierten Studien? Um dies zu untersuchen, ist das Gebiet der Telekardiologie ideal geeignet, da sowohl RCTs als auch große Registerstudien verfügbar sind und so ein Vergleich möglich ist. Daher stellten wir die Mortalitätsergebnisse von nicht randomisierten Studien aus der Telekardiologie den Ergebnissen einer Metaanalyse aus RCTs gegenüber. Dabei zeigten sich stark abweichende Ergebnisse. Wir erläutern mögliche Gründe für diese Diskrepanz und diskutieren Schlussfolgerungen.

Martin Schumacher (Institut für Medizinische Biometrie und Statistik (IMBI), Universitätsklinikum Freiburg); Claudia Schmoor (Zentrum klinische Studien, Universitätsklinikum Freiburg); Jan Beyersmann (Institut für Statistik, Universität Ulm)

Randomisation und kausale Inferenz in klinischen Studien

Seit den grundlegenden Arbeiten von Sir Austin Bradford Hill vor etwa 70 Jahren hat sich die randomisierte kontrollierte Studie (RCT) als Gold-Standard zur Bewertung der Effekte von therapeutischen Maßnahmen etabliert. In den letzten 4 Jahrzehnten sind statistische Methoden zur kausalen Inferenz entwickelt worden, die eine solche Bewertung auch im Rahmen einer nicht randomisierten Studie erlauben. Unter welchen Bedingungen und welchen zusätzlichen Annahmen dies möglich ist, wird am Beispiel einer sogenannten Comprehensive-Cohort-Studie erläutert, in die randomisierte und nicht randomisierte Patienten aufgenommen und unter ansonsten gleichen Bedingungen behandelt und nachbeobachtet wurden. Diese Studie bietet z. B. die Möglichkeit, die Ergebnisse von Propensity-Score-basierten Analysen der nicht randomisierten Teilstudie mit der Intention-to-treat-Analyse der randomisierten Teilstudie direkt zu vergleichen (Schmoor et al., 2008;

Schmoor et al., 2011). Der Regelfall beim Einsatz des Propensity Scores und anderer Methoden der kausalen Inferenz ist jedoch, dass eine randomisierte Teilstudie nicht vorliegt und die Bewertung der therapeutischen Maßnahmen allein auf den Daten einer nicht-randomisierten Studie beruht. Systematische Reviews zum Vergleich der Ergebnisse von randomisierten und nicht-randomisierten Studien ergeben hier bislang auch ein eher diffuses Bild. Im zweiten Teil des Vortrags wird die aktuelle Diskussion zur Bewertung der Effekte von Natrium-Glucose Cotransporter 2 (SGLT₂)-Inhibitoren zur Therapie von Patienten mit Typ 2-Diabetes (Suissa, 2018) aufgegriffen um einerseits die Ergebnisse von randomisierten (Zinman et al., 2015) und nicht randomisierten Studien (Birkeland et al., 2017) zu vergleichen und andererseits den Einsatz von Methoden der kausalen Inferenz bei weitergehenden Analysen, z. B. zur Aufklärung von Wirkungsmechanismen, zu illustrieren (Inzucchi et al., 2018). Randomisierte kontrollierte Studien werden auch in Zukunft den Gold-Standard bei der Bewertung von therapeutischen Maßnahmen darstellen, aber Methoden der kausalen Inferenz bieten eine wichtige Möglichkeit der weiteren Erkenntnisgewinnung, die mehr als bislang genutzt werden sollten.

Literatur

- [1] Schmoor C, Caputo A, Schumacher M. Evidence from nonrandomized studies: a case study on the estimation of causal effects. *Am J Epidemiol* 2008 167(9):1120-9. doi 10.1093/aje/kwn010.
- [2] Schmoor C, Gall C, Stampf S, Graf E. Correction of confounding bias in non-randomized studies by appropriate weighting. *Biom J* 2011 53(2): 369-87. doi: 10.1002/bmj.201000154.
- [3] Suissa S. Lower risk of death with SGLT₂ inhibitors in observational studies: Real or bias? *Diabetes Care* 2018, 41(1):6-10. doi 10.2337/dc17-1223.
- [4] Zinman B, Wanner C, Lachin JM, et al. for the EMPA-REG OUTCOME Investigators. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med* 2015; 373: 2117–28. doi: 10.1056/NEJMoa1504720.
- [5] Birkeland KI, Jørgensen ME, Carstensen B, et al. Cardiovascular mortality and morbidity in patients with type 2 diabetes following initiation of sodium-glucose co-transporter-2 inhibitors versus other glucose-lowering drugs (CVD-REAL Nordic): a multinational observational analysis. *Lancet Diabetes Endocrinol* 2017 5(9): 709-717. doi 10.1016/S2213-8587(17)30258-9.
- [6] Inzucchi SE, Zinman B, Fitchett D, et al. How does Empagliflozin reduce cardiovascular mortality? Insights from a mediation analysis of the EMPA-REG OUTCOME Trial. *Diabetes Care* 2018 41(2):256-363. doi: 10.2337/dc17-1096.