



IQWiG im Dialog 2016

*"Sind strengere Kriterien zur
Evidenzbewertung notwendig?"*

Abstracts der Vorträge



Dr. Stefanie Thomas & Prof. Dr. rer. biol. Hum. Ralf Bender (IQWiG)

Neue Vorschläge zur Evidenzbewertung

Die Hauptaufgabe des IQWiG ist es, in Hinblick auf Systementscheidungen möglichst zuverlässig die Frage zu beantworten, ob unter Verwendung der international anerkannten Standards der evidenzbasierten Medizin ein Nutzen oder Schaden durch eine Maßnahme nachzuweisen ist. Bekanntermaßen unterliegen "Standards" generell, so auch die der evidenzbasierten Medizin respektive der Erstellung systematischer Reviews bzw. HTAs als Entscheidungsgrundlage, einer permanenten Weiterentwicklung. In den vergangenen Jahrzehnten lag der Schwerpunkt darauf, die notwendigen Verfahren zur systematischen Beschaffung, Bewertung und (meta-analytischen) Zusammenfassung von Daten immer weiter zu explizieren mit dem Ziel, eine wissenschaftlich fundierte und transparente Methodik zu gewährleisten – nicht zuletzt seitens des IQWiG in seinem Methodenpapier.

In der jüngeren Vergangenheit scheinen jedoch grundsätzlichere Diskussionen bestimmter Standards (neu) geführt zu werden. Aktuell erkennbaren Tendenzen, noch "schnellere" Nutzenbewertungen für Systementscheidungen auf knappen und gegebenenfalls unsicheren Datengrundlagen basieren zu wollen, steht aktuell ein Strauß unterschiedlicher Forderungen seitens verschiedener Wissenschaftler entgegen, strengere Kriterien für die Evidenzbewertung anzulegen. Mitglieder des Core Editorial Teams der Cochrane-Injuries-Group fordern z. B. bei systematischen Übersichten nicht registrierte Studien auszuschließen, härtere Vorgehensweisen zur Überprüfung der Datenqualität sowie die Durchführung von Fallzahlplanungen und die Anwendung sequenzieller Methoden für Meta-Analysen und eine damit verknüpfte Beurteilung der Relevanz von Effekten (Roberts et al., 2015). Ähnliche, teils noch darüber hinaus gehende Vorschläge wurden von Jakobsen et al. (2014) formuliert. Für die Praxis relevante Forderungen sind hierbei insbesondere die Adjustierung für multiple Endpunkte sowie die Berechnung sowohl von p -Werten als auch von Bayes-Faktoren zur Beurteilung geschätzter Interventionseffekte. Mithilfe Bayes'scher Verfahren begründet wiederum auch Johnson (2013) seine Forderung, generell bei Anwendung statistischer Signifikanztests strengere Signifikanzniveaus zu verwenden (z. B. 0,005 oder 0,001 anstelle von 0,05).

Ziel des einführenden Vortrags ist es, den Strauß aktueller Vorschläge zu sortieren und für die folgenden Vorträge und Diskussion vorzubereiten. Konkret sollen ihre jeweilige Position innerhalb der Nutzenbewertung respektive das Problem, welchem Abhilfe geschaffen werden soll, fixiert und dabei soweit möglich, den aktuellen Verfahrensweisen im Institut sowie ersten Erfahrungen und Diskussionen aus der Literatur gegenüber gestellt werden.

Literatur

Jakobsen, J.C., Wetterslev, J., Winkel, P., Lange, T. & Gluud, C. (2014): Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med. Res. Methodol.* 14, 120.

Johnson, V.E. (2013): Revised standards for statistical evidence. *Proc. Natl. Acad. Sci. U.S.A.* 110, 19313-19317.

Roberts, I., Ker, K., Edwards, P., Beecher, D., Manno, D. & Sydenham, E. (2015): The knowledge system underpinning healthcare is not fit for purpose and must change. *BMJ* 350, h2463.

Dr. Simon Wandel (Novartis, Basel)

Bayes'sche Evidenzsynthese – führt dies zu strengeren Kriterien in der Nutzenbewertung?

Statistische Kriterien zur Bewertung von Studienergebnissen sind immer wieder Gegenstand von Kontroversen. So argumentiert z. B. Johnson¹, dass generell stringenter Signifikanzniveaus (< 0.005 oder < 0.001) verwendet werden sollten. Die zentralen Probleme vermag jedoch auch dieser Ansatz nicht zu lösen: Wie kann Evidenz aus unterschiedlichen Quellen (randomisiert, nicht-randomisiert, Register-Daten) statistisch adäquat kombiniert werden, um die Wirksamkeit von Medikamenten zu quantifizieren?

In den letzten Jahren hat sich gezeigt, dass Bayes'sche Evidenzsynthese (z.B. Spiegelhalter², Welton³, Dias⁴) zur formalen Integration von verschiedenen Datenquellen äusserst hilfreich ist. Sie ermöglicht einerseits Wahrscheinlichkeits-Aussagen über klinisch relevante Grössen, andererseits einen statistisch korrekten Umgang mit der Unsicherheit aller Modell-Parameter, so auch der Heterogenität zwischen Studien (z.B. Welton³, Friede⁵).

Der Nutzen Bayes'scher Evidenzsynthese wird im Zusammenhang mit Zulassungen und Nutzenbewertungen aufgezeigt, wobei öffentlich zugängliche Beispiele von FDA und NICE verwendet werden. Ein besonderes Augenmerk wird dabei auf die Kriterien gerichtet, welche zur Nutzenbewertung dienlich sind.

Literatur

¹ Johnson VE (2013): Revised standards for statistical evidence. *Proc Natl Acad Sci USA* 110, 19313-19317.

² Spiegelhalter DJ, Abrams, KR & Myles, JP (2004): *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Wiley, Chichester.

³ Welton NJ, Sutton AJ, Cooper N, Abrams KR & Ades AE (2012): *Evidence Synthesis for Decision Making in Healthcare*. Wiley, Chichester.

⁴ Dias S, Welton NJ, Sutton AJ & Ades AE (2011): *NICE DSU Technical Support Document 1: Introduction to evidence synthesis for decision making* (last updated April 2012).

⁵ Friede T, Röver C, Wandel S & Neuenschwander B (2016): Meta-analysis of few small studies in orphan diseases. *Res Synth Meth* (in press).

Dr. Erik von Elm (Cochrane Schweiz)

Evidenzpolizei – Welche Rolle sollen systematische Reviews haben?

Autoren systematischer Reviews befinden sich in einer besonderen Rolle, da sie die Methodik und Ergebnisse klinischer Forschung kritisch bewerten und dabei auch auf Unregelmässigkeiten bis hin zu wissenschaftlichem Fehlverhalten stossen können. Neben dem Nachfragen bei den Autoren der Originalarbeiten wird zunehmend auch Information aus Quellen wie Studienprotokollen, Registern oder Zulassungsdokumenten herangezogen, die über den Informationsgehalt von wissenschaftlichen Publikationen weit hinausgehen können. Es werden Ergebnisse einer neuen empirischen Studie (Elia et al., 2016) vorgestellt, in der der Umgang von Autoren systematischer Reviews mit wissenschaftlichem Fehlverhalten in den Originalstudien untersucht wird. Es wird diskutiert, welche Rolle systematische Reviews über die Synthese von Evidenz hinaus haben sollen.

Literatur

Elia, N., von Elm, E., Chatagner, A., Popping, D.M. & Tramer, M.R. (2016): How do authors of systematic reviews deal with research malpractice and misconduct in original studies? A cross-sectional analysis of systematic reviews and survey of their authors. *BMJ Open* 6, e010442.

Prof. Dr. Tim Friede (Universitätsmedizin Göttingen)

Was ist im Rahmen von Evidenzbewertungen relevant und wie wird es bewertet?

In der evidenzbasierten Medizin kommt der Evidenzbewertung eine zentrale Rolle zu. Hier stellt sich zunächst die Frage, welche Studien überhaupt für eine gegebene Fragestellung relevant sind und wie ihre Qualität zu bewerten ist. Zum Beispiel werden bei der Nutzenbewertung von Arzneimitteln in der Regel nur randomisierte kontrollierte Studien mit bestimmten Endpunkten und Kontrollgruppen als relevant angesehen. Wenn man zusätzlich sogar fordert, dass diese prospektiv registriert gewesen sein müssen (Roberts et al, 2015), dann kann die Evidenzbasis recht schnell sehr dünn werden, insbesondere bei der Betrachtung von Subgruppen. Andererseits werden Methoden der Meta-Analyse auch bei sehr wenigen Studien (Friede et al, 2016) und generalisierten Evidenzsynthese (Spiegelhalter et al, 2004) propagiert, die es ermöglichen, Daten aus Studien mit durchaus verschiedenen Designs zu integrieren und damit eine umfassendere Datenbasis schaffen. In dem Vortrag werden wir auf die Rolle von klinischen Registern und Routinedaten eingehen und Qualitätsanforderungen an diese diskutieren. Zudem kommt der Offenlegung von Daten in diesem Zusammenhang eine größere Bedeutung zu (Wegscheider und Friede, 2016).

Derzeit werden Forderungen lauter, Prinzipien, die sich im Bereich der Planung und Auswertung klinischer Studien etabliert haben, nun auch für die Evidenzsynthese zu übernehmen. Hierzu zählen die Adjustierung des Signifikanzniveaus für multiple Endpunkte oder wiederholte Auswertungen bei kumulativen Meta-Analysen (sogenannte Trial Sequential Analyses) sowie Fallzahlplanungen und die Beurteilung der klinischen Relevanz von Effekten. Im Vortrag werden wir die Eignung dieser Kriterien im Zusammenhang mit Meta-Analysen diskutieren und dabei insbesondere auf Kriterien der klinischen Relevanz eingehen (Kieser et al, 2004, 2013).

Abschließend werden wir argumentieren, dass die Wahl der Kriterien im Zusammenhang mit der Wahl der Evidenzbasis zu sehen ist und dass hier Risiken der Verzerrung in der Effektschätzung und die Präzision der Schätzung balanciert werden sollten.

Referenzen

Friede T, Röver C, Wandel S, Neuenschwander B (2016) Meta-analysis of two studies in the presence of heterogeneity with applications in rare diseases. (Submitted)

Kieser M, Röhm J, Friede T (2004) Planning considerations when assessing the clinical relevance of trial results by responder analyses. *Statistics in Medicine* 23: 3287-3305.

Kieser M, Friede T, Gondan M (2013) Assessment of statistical significance and clinical relevance. *Statistics in Medicine* 32: 1707-1719.

Roberts I, Ker K, Edwards P, Beecher D, Manno D, Sydenham E (2015) The knowledge system underpinning healthcare is not fit for purpose and must change. *BMJ* 350, h2463.

Spiegelhalter DJ, Abrams KR, Myles JP (2004) *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Wiley, Chichester.

Wegscheider K, Friede T (2016) Do we consent to rules of consent and confidentiality? (Submitted)