

Mindestmengen und Schwellenwertmodelle bei der Knie-Totalendoprothese (Knie-TEP)

Grouven U, Bender R

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, Köln, Deutschland

ulrich.grouven@iqwig.de

Einleitung und Fragestellung

In den letzten drei Jahrzehnten wurden für verschiedene Leistungen der Gesundheitsversorgung eine Vielzahl von Studien zu einer möglichen Assoziation zwischen Behandlungsmenge und Ergebnisqualität publiziert [1]. Vor diesem Hintergrund wurde das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) vom Gemeinsamen Bundesausschuss (G-BA) damit beauftragt, auf der Basis deutscher Daten für den Leistungsbereich Knie-TEP Modelle für die Berechnung von Schwellenwerten zu entwickeln und anzuwenden. Im vorliegenden Beitrag sollen die wesentlichen Ergebnisse präsentiert werden.

Material und Methoden

Grundlage für die Berechnungen waren Daten der BQS Bundesgeschäftsstelle Qualitätssicherung aus dem Bundesdatenpool 2004. Die Daten wurden für den vorliegenden Auftrag von der BQS aufbereitet und dem IQWiG so zur Verfügung gestellt, dass auf der Beobachtungseinheit "Patient" regressionsanalytische Berechnungen möglich waren. Die Fallzahl der behandelten Patienten bezog sich jeweils auf ein Krankenhaus. Die Daten wurden in pseudonymisierter Form unter Beachtung der amtlichen Datenschutzrichtlinien aufbereitet und statistisch ausgewertet. Als relevante Zielgrößen wurden die binären Qualitätsindikatoren „Unbeweglichkeit“ und „Wundinfektion“ sowie die für eine Risikoadjustierung zu betrachtenden Einflussgrößen (Alter, Geschlecht, ASA Score, präoperativer Status) betrachtet.

Zur Modellierung des Zusammenhangs zwischen Fallzahl und Qualitätsindikatoren unter Berücksichtigung potenzieller Confounder wurden multiple logistische Regressionsmodelle sowie Generalized Estimating Equations (GEE) [2] zur Berücksichtigung möglicher Cluster-Effekte verwendet. Es wurde die Methode der Fractional Polynomials (FP) eingesetzt, die eine flexible Modellierung auch nicht monotoner Zusammenhänge ermöglicht [3]. Zur Überprüfung der Modellanpassung wurden der Hosmer-Lemeshow-Test, das adjustierte Likelihood-Ratio R^2 , sowie die c-Statistik (Fläche unter der ROC-Kurve) verwendet [4]. Zur Untersuchung der Ergebnisstabilität wurden die gleichen Modelle auch auf die Daten aus dem Jahr 2003 angewendet. Schließlich wurden regressionsdiagnostische Verfahren sowie Sensitivitätsuntersuchungen hinsichtlich fehlender Werte durchgeführt.

Schwellenwerte wurden mit Hilfe des VARL-Ansatzes nach Bender [5] basierend auf dem logistischen Regressionsmodell hergeleitet. Zur Bewertung der klinischen Relevanz wurden zum einen die Differenz (δ) des mittleren Risikos über alle Patienten (π) und dem mittleren Risiko der Patienten aus High-Volume-Krankenhäusern (π_{HV}), d.h. $\delta = \pi - \pi_{HV}$, und zum anderen die gewöhnliche Differenz γ der Risiken zwischen Low-Volume- (π_{LV}) und High-Volume-Krankenhäusern, d.h. $\gamma = \pi_{LV} - \pi_{HV}$, berechnet, wobei der Trennwert für die Unterscheidung zwischen High-Volume- und Low-Volume-Krankenhäusern durch die betrachtete Mindestmenge gegeben ist. Der Kehrwert von δ wird als Population Impact Number (PIN) [6], der Kehrwert von γ als Number Needed to Treat (NNT) bezeichnet [7]. δ und γ beleuchten somit unterschiedliche Sichtweisen bei der Darstellung von Interventionseffekten (mit und ohne Bezug zur Grundgesamtheit).

Ergebnisse

Es wurden 91714 Datensätze für den Indikator „Unbeweglichkeit“ und 110349 Datensätze für den Indikator „Infektion“ im Rahmen einer Complete-Case-Analyse ausgewertet. Das mittlere Risiko für „Unbeweglichkeit“ betrug 26.9%. Ein deskriptiver Risikoplot zeigte eine U-förmige Struktur, so dass FPs zur Modellierung des Zusammenhangs verwendet wurden. Die beste Anpassung lieferte ein FP 2. Ordnung mit linearem und kubischem Term. Der geschätzte Risikoverlauf basierend auf einem logistischen Regressionsmodell mit FPs ist in Abb. 1 dargestellt. Der Einfluss der Fallzahl ist hoch signifikant ($p < 0.0001$). Eine Adjustierung bzgl. potenzieller Confounder im Rahmen multifaktorieller Modelle führte zu vergleichbaren Ergebnissen bzgl. der Fallzahl. Eine Mindestmengenberechnung war aufgrund des U-förmigen Risikoverlaufs nicht sinnvoll durchführbar.

Das mittlere Risiko für den Indikator „Infektion“ betrug 0.8%. Die Risikokurve in Abhängigkeit von der Fallzahl basierend auf einem einfachen logistischen Regressionsmodell zeigt einen statistisch signifikanten ($p < 0.0001$), aber sehr flachen, leicht fallenden Verlauf (Abb. 1). Die Berechnung eines Schwellenwertes nach der VARL-Methode bei Vorgabe eines akzeptablen Risikos von 1% lieferte einen Wert von VARL=116 (95% KI: 90-141). Die zugehörige Risikoreduktion betrug $\delta = 0.15\%$ (95% KI: 0.11-0.19%) (entspricht einer PIN von 664 (95% KI: 529-872)) und $\gamma = 0.53\%$ (95% KI: 0.41-0.67%) (entspricht einer NNT von 187 (95% KI: 149-246)).

Die Berechnungen mit den Daten aus dem Jahr 2003 führten zu konsistenten Ergebnissen.

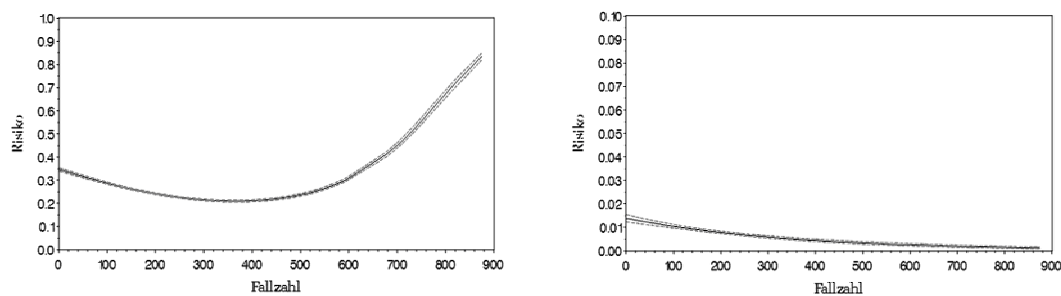


Abb.1: Risikoverlauf mit 95%-Konfidenzbändern für die Indikatoren „Unbeweglichkeit“ (links) und „Infektion“ (rechts) basierend auf logistischen Regressionsmodellen

Schlussfolgerung

Die gefundenen Ergebnisse aus den verfügbaren deutschen Daten zeigen zwar einen Zusammenhang zwischen Leistungsmenge und Ergebnisqualität, erlauben aber keine Herleitung eines klaren und einheitlichen Grenzwertes bzgl. der Fallzahl pro Krankenhaus für die Unterscheidung zwischen höherer und niedrigerer Qualität. Diese Ergebnisse stehen in Übereinstimmung mit entsprechenden Resultaten basierend auf internationalen Daten [8].

Literatur

- [1] Rathmann W, Windeler J. Zusammenhang zwischen Behandlungsmenge und Behandlungsqualität. Evidenzbericht, Fachbereich Evidenz-basierte Medizin des MDS, Essen, 2002.
- [2] Kleinbaum DG, Klein M. Logistic Regression: A Self-Learning Text (2nd Ed.). New York: Springer, 2002.
- [3] Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol 1999; 28: 964-974.
- [4] Hosmer DW, Lemeshow S. Applied Logistic Regression (2nd Ed.). New York: Wiley, 2000.
- [5] Bender R. Quantitative risk assessment in epidemiological studies investigating threshold effects. Biom J 1999; 41: 305-319.
- [6] Heller RF, Dobson A. Disease impact number and population impact number: Population perspectives to measures of risk and benefit. BMJ 2000; 321: 950-952.
- [7] Cook RJ, Sackett DL. The number needed to treat: A clinically useful measure of treatment effect. BMJ 1995; 310: 452-454.
- [8] Schröder P, Rath T. Mindestmengen in der Kniegelenkendoprothetik. Orthopäde 2005; 34: 198-209.