

IQM Hearing

„Qualitätsmessung in der Medizin - Risikoadjustierung und Risikoadjustierungsmodelle“

**Freitag | 01. Februar 2019 | 10:00 Uhr - 16:00 Uhr
Hotel Rossi | Lehrter Straße 66 | 10557 Berlin**

Inhalt

Risikoadjustierung von Qualitätsindikatoren mit Routinedaten

Christian Günster, Dipl.-Math.; Dr. Elke Jeschke
Wissenschaftliches Institut der AOK (WIdO)

Risikoadjustierung mit multipler logistischer Regression*

Dr. Marcus Kutschmann, Dipl.-Stat.
BQS Institut für Qualität und Patientensicherheit

Risikoadjustierung auf Basis multipler Regressionsanalysen zur Ermittlung von individuellen Erwartungswerten

Raphael Graf; Julia Schmidt
3M Health Information Systems

Anforderungen an eine Risikoadjustierung zur Qualitätsmessung medizinischer Behandlungsergebnisse

Dr. Ulrike Nimptsch
Technische Universität Berlin, Fachgebiet Management im Gesundheitswesen

Stratifizierung und Risikoadjustierung - Vorschlag eines evidenzbasierten Algorithmus zur Erstellung von interpretierbaren und zielgerichteten Qualitätsindikatoren

Prof. Dr. Jochen Schmitt; Martin Röbler, M.Sc.; Dr. Olaf Schoffer
*Zentrum für evidenzbasierte Gesundheitsversorgung (ZEGV),
Universitätsklinikum Carl Gustav Carus und Medizinische Fakultät, TU Dresden*

Die Reliabilität der Inpatient Quality Indicators in der vergleichenden Darstellung der Qualitätsleistung von Kliniken

Dr. med., Dr. sc. nat. Michael Havranek
Department Health Sciences & Health Policy, Universität Luzern

Zur Bedeutung der Risikoadjustierung in der Qualitätsmessung

PD Dr. Günther Heller

Der BAA-Risikoscore der DGG als Entscheidungshilfe zur Therapiewahl bei Versorgung des intakten Bauchaortenaneurysmas

Prof. Dr. med. Reinhart T. Grundmann; M. Steffen; Univ.-Prof. Dr. med. Thomas Schmitz-Rixen
*Deutsches Institut für Gefäßmedizinische Gesundheitsforschung (DIGG)
der Deutschen Gesellschaft für Gefäßchirurgie und Gefäßmedizin*

Risikoadjustierte Qualitätsvergleiche im Deutschen Qualitätsbündnis Sepsis auf Basis der §21-Daten

Dr. Daniel Schwarzkopf; Dr. Daniel Oliver Thomas-Rüddel; Dr. Carolin Fleischmann; Dr. Hendrik Rüddel; Prof. Dr. Konrad Reinhart
*Deutsches Qualitätsbündnis Sepsis - icosmos / Center for Sepsis Control and Care,
Universitätsklinikum Jena*

Weiterentwicklung von Indikatoren zur Ergebnisqualität in der Frauenheilkunde aus Routinedaten

Prof. Dr. Birgit Seelbach-Göbel
Deutsche Gesellschaft für Gynäkologie und Geburtshilfe e. V. (DGGG)

*Beitrag noch nicht vom Autor zur Veröffentlichung freigegeben

Das Wissenschaftliche Institut der AOK (WIdO)

Das Wissenschaftliche Institut der AOK ist eine 1976 gegründete Forschungs- und Entwicklungsstelle in der AOK-Gemeinschaft. Am Institut werden zentrale Fragestellungen der Gesundheits- und Pflegeversorgung und ihrer Finanzierung bearbeitet. Das WIdO versteht sich als Bindeglied zwischen Wissenschaft und Praxis. Das heißt einerseits, wissenschaftliche Erkenntnisse für die Praxis nutzbar zu machen, und andererseits, Fragen aus der Praxis auf die wissenschaftliche Agenda zu bringen.

Entsprechend seinem Statut hat das WIdO zwei zentrale Aufgabenschwerpunkte. Zum einen unterstützen die durchgeführten Forschungs- und Entwicklungsarbeiten insbesondere die AOK-Gemeinschaft dabei, eine qualitativ hochwertige und wirtschaftliche Gesundheitsversorgung sicherzustellen. Zum anderen geht es um Forschung und Lehre, die Grundlagen, Probleme und die Weiterentwicklung der gesetzlichen Krankenversicherung und mit ihr zusammenhängender Gebiete in den Blick nimmt. Das WIdO engagiert sich auch in internationalen Forschungskooperationen, etwa mit der Weltgesundheitsorganisation (WHO), der „European Surveillance of Antimicrobial Consumption“ (ESAC) oder der Piperska-Gruppe.

Forschungsschwerpunkte des WIdO sind empirische und ordnungspolitische Fragestellungen zu Inanspruchnahme und Ergebnissen der Gesundheits- und Pflegeversorgung sowie zu ihrer Finanzierung. Die Analysen folgen zum Teil der sektoralen Struktur der Versorgungslandschaft. Das gilt beispielweise für Analysen zum Arzneimittelmarkt oder zur ambulanten und stationären Versorgung. Vielfach sind es aber auch sektorenübergreifende Fragestellungen aus Patientensicht, die auf Basis einer integrierten Datengrundlage angegangen werden, wie zum Beispiel bei der Qualitätsforschung. Das WIdO befasst sich zudem mit ordnungspolitischen Fragen des Gesundheitssystems. Dazu zählen zum Beispiel die Ausgestaltung der Finanzierung des Krankenversicherungsschutzes sowie vertragswettbewerbliche Steuerungsmöglichkeiten der Versorgung.

Das WIdO hat Zugang zu anonymisierten Routinedaten der 26 Mio. AOK-Versicherten. Am Institut wurde in den letzten Jahrzehnten eine Vielzahl von Studien im Bereich der Qualitäts- und Versorgungsforschung auf der Grundlage von Routinedaten durchgeführt. Sekundärdatenanalysen des WIdO sind maßgeblicher Bestandteil der Publikationsreihen Krankenhaus-, Arzneiverordnungs-, Versorgungs-, Fehlzeiten-, Heilmittel- und Pflege-Report, die vom WIdO herausgegeben bzw. mit Beteiligung des WIdO entstehen.

Organisatorisch ist das WIdO eine selbstständige Einheit innerhalb des AOK-Bundesverbandes ohne eigene Rechtspersönlichkeit. An der Spitze des Instituts steht eine dreiköpfige Geschäftsführung. Ein eigener Vorstand berät über die Aufgaben des Instituts. Ein wissenschaftlicher Beirat unterstützt dabei, den hohen wissenschaftlichen Ansprüchen gerecht zu werden. Im WIdO sind Mitarbeiterinnen und Mitarbeiter aus Epidemiologie, Statistik, Pharmakologie, Ökonomie u.a. beschäftigt.

Risikoadjustierung von Qualitätsindikatoren mit Routinedaten

Christian Günster (Dipl.-Math.), Dr. Elke Jeschke

Wissenschaftliches Institut der AOK (WiDO), Rosenthaler Str. 31, 10178 Berlin

Einführung

Beim indikatorengestützten Vergleich von Kliniken oder Therapien ist eine Risikoadjustierung erforderlich, wenn sich der Patientenmix der zu vergleichenden Gruppen im Hinblick auf Risikofaktoren unterscheidet, die den Qualitätsendpunkt beeinflussen. Mögliche Ursachen für Unterschiede im Patientenmix sind Populationsunterschiede, Spezialisierung und Risikoselektion.

Ziel der Risikoadjustierung ist die Begründung eines fairen Gruppenvergleichs. Sie ist von besonderer Bedeutung beim Vergleich von medizinischen Versorgungseinrichtungen. Die Adjustierung sollte dabei auf patientenseitig vorbestehenden Risikofaktoren wie Krankheitsschwere, Begleiterkrankungen und Vorbehandlungen beruhen (Iezzoni 2013). Nach Faktoren, die durch die Einrichtung beeinflusst werden, sollte nicht adjustiert werden, da sie die zu messende Qualität der medizinischen Versorgung konstituieren, z.B. OP-Zeitpunkt, Prothesentyp, Auftreten von Komplikationen (IQTIG 2018). Bei der Auswahl der Faktoren stehen die klinische und prädiktive Validität des Risikomodells im Vordergrund (Bottle/Aylin 2017).

Die am weitesten verbreiteten Verfahren der Risikoadjustierung sind Stratifizierung und Regressionsmodelle. Bei der Stratifizierung werden die Patienten anhand eines Risikofaktors in Gruppen (Strata) unterteilt oder auch ausgeschlossen. Demgegenüber verbleiben bei der Regression alle Patienten in einer Gruppe, wobei aber nach dem Einfluss jeden Risikofaktors adjustiert wird. In der Regel liegen eine Vielzahl von Einflussfaktoren vor, so dass die Stratifizierung wegen der daraus resultierenden Vielzahl von Strata – zudem oft mit geringer Fallzahl – nicht mehr praktikabel ist.

Krankheitsschwere und Routinedaten

Routinedaten in Form von Klinikabrechnungsdaten wie die Daten nach § 21 Krankenhausentgeltgesetz (KHEntgG) beinhalten Angaben zum Patienten, dem Verlauf des Krankenhausaufenthaltes, zu Haupt- und Nebendiagnosen sowie Codes der durchgeführten Operationen und Prozeduren. Für die Adjustierung von Begleiterkrankungen mittels Routinedaten sind spezielle Komorbiditäts-Klassifikationen und Indizes entwickelt worden. Die bekanntesten darunter sind die Elixhauser Conditions und der Charlson Comorbidity Index (Stausberg/Hagn 2015). Der Vergleich der verschiedenen Klassifikationen, die zumeist im Hinblick auf den Endpunkt Sterblichkeit entwickelt worden sind, bei ihrer Anwendung zur Adjustierung weiterer Endpunkte und Therapien ist Gegenstand aktueller Forschung. Beispielsweise zeigt sich die Diskriminationsfähigkeit (Trennschärfe) der Elixhauser-Klassifikation bei der Vorhersage von unerwünschten Ereignissen und Sterblichkeit nach

totalem Hüftgelenkersatz gegenüber anderen Systemen zur Beschreibung der Krankheitsschwere und Komorbidität überlegen (Ondeck et al. 2018, Inacio et al. 2016). Routinedaten von Krankenkassen umfassen zusätzlich Angaben über ambulante und stationäre Vorbehandlungen, die eine noch detailliertere Darstellung der Risikosituation ermöglichen (Swart et al. 2014).

Qualitätssicherung mit Routinedaten (QSR)

Die Berücksichtigung der Schwere der therapierten Erkrankung, der Begleiterkrankung und unter Umständen auch der Therapieform ist ein zentraler Bestandteil des vom wissenschaftlichen Instituts der AOK durchgeführten Verfahrens Qualitätssicherung mit Routinedaten (QSR). Im QSR-Verfahren erfolgen Klinikvergleiche für 20 medizinische Leistungsbereiche mit 92 Qualitätsindikatoren. Die Ergebnisse werden in Form eines Klinikberichtes den Kliniken und der IQM-Gruppe bereitgestellt. Das Verfahren basiert im Wesentlichen auf Abrechnungsdaten aus der Klinikabrechnung gemäß § 301 SGB V von Krankenhausbehandlungen bei AOK-Patienten. Es erfasst rund 25 Prozent aller AOK-Krankenhauspatienten. Für acht Leistungsbereiche werden risikoadjustierte Indikatoren im Kliniksuchportal AOK-Krankenhausnavigator Einrichtungsbezogen veröffentlicht (WIdO 2018a).

Identifikation von Risikofaktoren

Die Definition der Modelle zur Risikoadjustierung im QSR-Verfahren beginnt mit der Identifikation der einzubeziehenden Faktoren. Potentielle Risikofaktoren mit Relevanz für die Qualitätsindikatoren werden erstens im Zuge einer Literatur- und Indikatorenrecherche ermittelt, zweitens durch Expertenkonsens in einer beratenden Panelgruppe von klinischen Experten benannt. Schließlich wird in einem dritten Schritt eine explorative, empirische Analyse (möglichst als Longitudinalanalyse) durchgeführt, um weitere Faktoren mit signifikantem Einfluss auf den Endpunkt zu detektieren.

Konkret erfolgt – sofern geboten – für jeden QSR-Indikator eine Risikoadjustierung mindestens nach den Einflussgrößen Alter, Geschlecht und Begleiterkrankungen. Die Definition der Begleiterkrankungen wird gemäß den 31 Elixhauser Conditions vorgenommen. Hinzukommen mögliche Faktoren wie Schwere des Behandlungsanlasses, OP-Verfahren oder eine antithrombotische Medikation im Vorjahr. Als Kriterien zur Auswahl von Risikofaktoren für das Risikoadjustierungsmodell werden angelegt:

- hohe Konstruktvalidität der Risikodefinition, d.h. dass die Definition die intendierte Risikosituation möglichst genau erfasst
- prädiktive Validität des Risikofaktors
- keine Beeinflussbarkeit des Risikofaktors durch die Einrichtung (beim Einrichtungsvergleich)
- keine Adjustierung nach in der Einrichtung erworbener Morbidität oder Komorbidität
- keine Adjustierung nach Prozessvariablen (z.B. Verlegung auf die Intensivstation)

- ausreichende Homogenität der Dokumentation des Risikofaktors
- Operationalisierbarkeit des Risikofaktors mittels Routinedaten.

Dabei ist zu berücksichtigen, dass ein Mehr an Variablen in den statistischen Modellen und bessere Fitmaße nicht gleichbedeutend mit einer besseren Adjustierung sind, sondern die Vergleichbarkeit von Kliniken sogar verschlechtern können (Nicholl 2007, Dimick et al. 2010, Heller/Schnell 2007).

Modellierung

In QSR werden risikoadjustierte Analysen mittels logistischen Regressionen durchgeführt. Logistische Regressionen stellen bei der Untersuchung binärer Outcomes eine weit verbreitete Analyseverfahren dar, die eine flexible Modellierung unterschiedlichster inhaltlicher Annahmen erlauben.

Zur Modellierung von zusammengesetzten Indikatoren, die verschiedenartige Endpunkte enthalten (z.B. Gesamtindikator), werden multinomiale logistische Regressionen verwendet. Die Indikatorenbestandteile werden dazu nach Schwere des Endpunkts hierarchisiert. Der Indikatorbestandteil Sterblichkeit hat dabei den höchsten Schweregrad.

Die Modellierung erfolgt stets mit robusten logistischen Regressionen. Der Grund hierfür ist, dass eine Gruppe von Patienten aus demselben Krankenhaus sich untereinander (in aller Regel) ähnlicher ist als eine Patientengruppe, die aus unterschiedlichen Krankenhäusern stammt. Auch die Kodiergewohnheiten innerhalb eines Krankenhauses werden ähnlicher sein als zwischen verschiedenen Krankenhäusern (so genanntes „clustering of patients“). Dieses Phänomen entspricht einer mangelnden statistischen Unabhängigkeit der analysierten Krankenhausfälle. Um diesem Sachverhalt Rechnung zu tragen, werden robuste logistische Regressionsmodelle verwendet (Rousseeuw/Leroy 2003).

Für jeden Qualitätsindikator werden zunächst einzeln Risikoadjustierungsmodelle mit allen identifizierten Risikofaktoren berechnet. Dann werden die Modelle durch Ausschluss nicht signifikanter Risikofaktoren reduziert und neu berechnet. Bei einfachen logistischen Regressionen wird der stepwise backward-Algorithmus angewendet. Die Modelle werden auch auf mögliche Multi-Kollinearität der Risikoadjustierungsvariablen geprüft, indem Varianz-Inflations-Faktoren (VIF) berechnet wurden (Tabachnick/Fidell 1996; Chatterjee/Hadi 2012). Alle Analysen werden mit dem Programmpaket Stata durchgeführt.

Überprüfung der Modellgüte

Die Güte der geschätzten Modelle wird mittels des Hosmer-Lemeshow-Goodness-of-Fit-Tests und über die C-Statistik der Receiver-Operator-Characteristic (ROC) beurteilt. Dabei wird in einer ROC-Kurve über alle Schwellenwerte hinweg die Sensitivität gegen die (Un-)Spezifität abgetragen und die Fläche unter der Kurve ausgewiesen (Metz 1978; van Erkel/Pattynama 1998; Hosmer et al. 2013; Schneeweiss et al. 2003).

Schließlich werden zum Test auf systematische, unerklärte Varianz Vergleiche von Klinikgruppen durchgeführt. Zum Beispiel ist denkbar, dass Unikliniken, die verstärkt ältere und multimorbide Patienten oder Patienten mit besonders schweren Verläufen der Indexerkrankung behandeln, trotz Adjustierung nach individuellen Patientenrisiken systematisch schlechter abschneiden als andere Kliniken. In diesem Fall könnte eine unvollständige Risikoadjustierung vorliegen. Die zu betrachteten Klinikgruppen werden durch die Panelgruppe festgelegt. Üblicherweise werden Unikliniken, Maximalversorger und Fachkliniken betrachtet. Die Relevanz signifikanter Gruppeneinflüsse wird durch die Panelgruppe bewertet.

Ergebnisdarstellung

Die so erhaltenen Regressionsmodelle werden genutzt, um fallspezifisch erwartete Ereigniswahrscheinlichkeiten zu berechnen. Die über alle Patienten einer Klinik kumulierten Ereigniswahrscheinlichkeiten ergeben dabei die erwarteten Ereignisse der Klinik ($E = \text{expected}$), die für die im Folgenden erläuterte Berechnung der SMR (standardized morbidity/mortality ratio) benötigt werden. SMRs werden berechnet, um risikoadjustierte Vergleiche von krankenhausspezifischen Qualitätsindikatoren zu ermöglichen. Dabei wird für jede Klinik die Rate der beobachteten Ereignisse ($O = \text{observed}$) durch die Rate an erwarteten Ereignissen dividiert.

Sind O und E gleich groß, ergibt sich eine SMR von 1. Dies wäre gleichbedeutend mit der Aussage, dass die risikoadjustierte Ereignisrate der durchschnittlichen Ereignisrate der Population entspricht. Eine SMR von 1,5 bedeutet, dass die risikoadjustierte Ereignisrate um 50 % über dem Durchschnitt liegt, während eine SMR von 0,7 anzeigt, dass das Risiko, ein Ereignis zu erfahren, in der betrachteten Klinik vermindert ist und risikoadjustiert 70 % der durchschnittlichen Ereignisrate des Vergleichskollektivs beträgt. Als Vergleichskollektiv werden alle Behandlungsfälle im Leistungsbereich herangezogen.

Risikoadjustierung mit Routinedaten am Beispiel Hüftgelenkersatz

In der folgenden Analyse wurde anhand bundesweiter Routinedaten von 133.367 Tumorfreien AOK-Patienten mit einer Implantation einer Hüftgelenks-Endoprothese bei Coxarthrose im Zeitraum von 2014 bis 2016 untersucht, welche Komorbiditäten mit einem erhöhten Risiko für eine Hüftprothesenrevision innerhalb eines Jahres sowie für schwere Allgemeinkomplikationen (maschinelle Beatmung über 24h, Reanimation, Sepsis, Herzinfarkt, Schlaganfall, Pneumonie, SIRS, Transfusion ≥ 6 TE u.a.) assoziiert sind. Die Definitionen entsprechen den Festlegungen im QSR-Verfahren (WIdO 2018b), wobei der Indikator für schwere Allgemeinkomplikation in Anlehnung an die Funktion für schwere Komplikationen des G-IQI-Handbuchs 5.1 definiert wurde. Als Risikofaktoren wurden neben Alter und Geschlecht die Elixhauser-Komorbiditäten (ohne Tumor und Übergewicht), BMI (30-34, 35-39, 40+) und eine antithrombotische Medikation im Vorjahr betrachtet.

Die Revisionsrate innerhalb eines Jahres lag bei 2,64 %, eine schwere Allgemeinkomplikation trat bei 2,44 % auf. Signifikante Risikofaktoren ($p < 0,05$) sind in Tabelle 1 dargestellt. Das Risiko einer Hüftrevision steigt mit zunehmendem BMI. Das volle Modell ist bzgl. seiner Diskriminationsfähigkeit einem reinen Alter-Geschlecht-Modell deutlich überlegen. Die Fläche unterhalb der Receiver Operator Characteristic Curve (AUC) beträgt 0,6162 vs. 0,5214 (Revision) bzw. 0,8269 vs. 0,6771 (Allgemeinkomplikation). Ein Wert von 1,0 wird erreicht, wenn alle Komplikationen korrekt vorhergesagt würden. Die Differenzierung nach BMI sowie die Berücksichtigung der antithrombotischen Medikation erhöht die Modellgüte geringfügig (Abbildung 1).

Die Adjustierung mittels Routinedaten erfasst beim Hüftgelenkersatz relevante Risikofaktoren. Die Diskriminationsfähigkeit des Revisionsmodells ist mäßig, was erwartet werden kann, da die im Modell nicht zu berücksichtigende Implantationsoperation (Prothesentyp, Operateur, etc.) den größten Einflussfaktor darstellen dürfte. Das Modell für schwere Allgemeinkomplikationen hat eine gute Passung. Eine BMI-Differenzierung und eine Erweiterung des Elixhauser-Klassifikation ist klinisch angezeigt. Sie kann für einzelne Häuser im Klinikvergleich bedeutsam sein. Schließlich gibt die Analyse Hinweise, welche Begleiterkrankung in der Therapie zur Vermeidung von Komplikationen besonders adressiert werden müssen.

Fazit und Empfehlungen

Für eine Risikoadjustierung mittels Routinedaten liegen spezielle Klassifikationssysteme zur Erfassung von Begleiterkrankungen vor. Robuste logistische und multinominale Regressionen stellen etablierte Modellierungsmethoden dar. Die Einbeziehung klinischer Expertise sowie die empirische Analyse bei Identifikation von Risikofaktoren und Bewertung der statistischen Modellierung zeigt sich im QSR-Verfahren als essentiell für die Entwicklung der Risikomodelle. Sofern die relevanten Einflussgrößen in Routinedaten abgebildet werden können, lassen sich mittels Routinedaten Risikomodelle mit guter und sehr guter prädiktiver Validität bilden. Durch den Vergleich mit klinischen Modellen oder Modellen unter Verwendung von Longitudinaldaten, wie sie in den Daten der Krankenkassen vorliegen, können die Eigenschaften von Modellen auf Grundlage der Daten nach § 21 KHEntgG zusätzlich überprüft werden.

Literatur

Bottle A, Aylin P (2017): Statistical Methods for Healthcare Performance Monitoring. Boca Raton, FL: CRC Press.

Chatterjee S, Hadi AS (2012). Regression Analysis by Example. New York, John Wiley Sons.

Dimick JB, Osborne NH, Hall BL, Ko CY, Birkmeyer JD (2010), Risk adjustment for comparing hospital quality with surgery: how many variables are needed? Journal of the American College of Surgeons 210 (4), 503-508. doi:10.1016/j.jamcollsurg.2010.01.018.

Heller G, Schnell R (2007), Hospital mortality risk adjustment using claims data. *Jama* 297 (18), 1983; author reply 1984. doi:10.1001/jama.297.18.1983.

Hosmer DW, Lemeshow S, Sturdivant RX (2013). *Applied Logistic Regression*. Wiley.

Iezzoni, LI (2013): Reasons for Risk Adjustment. Chapter 1. In: Iezzoni, LI; Hrsg.: *Risk Adjustment for Measuring Healthcare Outcome*. 4th edition. Chicago, US-IL: Health Administration Press, 1-14.

Inacio MCS, Pratt NL, Roughead EE, Graves SE (2016). Evaluation of three co-morbidity measures to predict mortality in patients undergoing total joint arthroplasty. *Osteoarthritis and Cartilage* 24: 1718-1726.

IQTIG (2018): Methodische Grundlagen V 1.1s;

https://igtig.org/dateien/dasiqtig/grundlagen/IQTIG_Methodische-Grundlagen-V1.1s-2018-11-15.pdf.

Metz CE (1978), Basic principles of ROC analysis. *Semin Nucl Med* 8(4), 283–98.

Nicholl J (2007), Case-mix adjustment in non-randomised observational evaluations: the constant risk fallacy. *Journal of epidemiology and community health* 61 (11), 1010-1013. doi:10.1136/jech.2007.061747.

Ondeck NT, Bohl DD, Bovonratwet P, McLynn RP, Cui JJ, Grauer JN (2018): Discriminative Ability of Elixhauser's Comorbidity Measure is Superior to Other Comorbidity Scores for Inpatient Adverse Outcomes After Total Hip Arthroplasty. *J Arthroplasty* 33(1): 250-257.

Rousseeuw PJ, Leroy AM (2003). *Robust Regression and Outlier Detection*. Wiley.

Schneeweiss S, Wang PS, Avorn J, Glynn RJ. Improved comorbidity adjustment for predicting mortality in Medicare populations. *Health Serv Res*. 2003 Aug;38(4):1103-20.

Stausberg J, Hagn S (2015), New Morbidity and Comorbidity Scores based on the Structure of the ICD-10. *PLoS One* 10 (12), e0143365. doi:10.1371/journal.pone.0143365.

Swart E, Ihle P, Gothe H, Matusiewicz D (Hrsg.) (2014). *Routinedaten im Gesundheitswesen. Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven*. Hogrefe-Verlag.

Tabachnick BG, Fidell LS (1996). *Using Multivariate Statistics* (3rd ed.). New York: Harper Collins.

van Erkel AR, Pattynama PM. Receiver operating characteristic (ROC) analysis: basic principles and applications in radiology. *Eur J Radiol*. 1998 May;27(2):88-94.

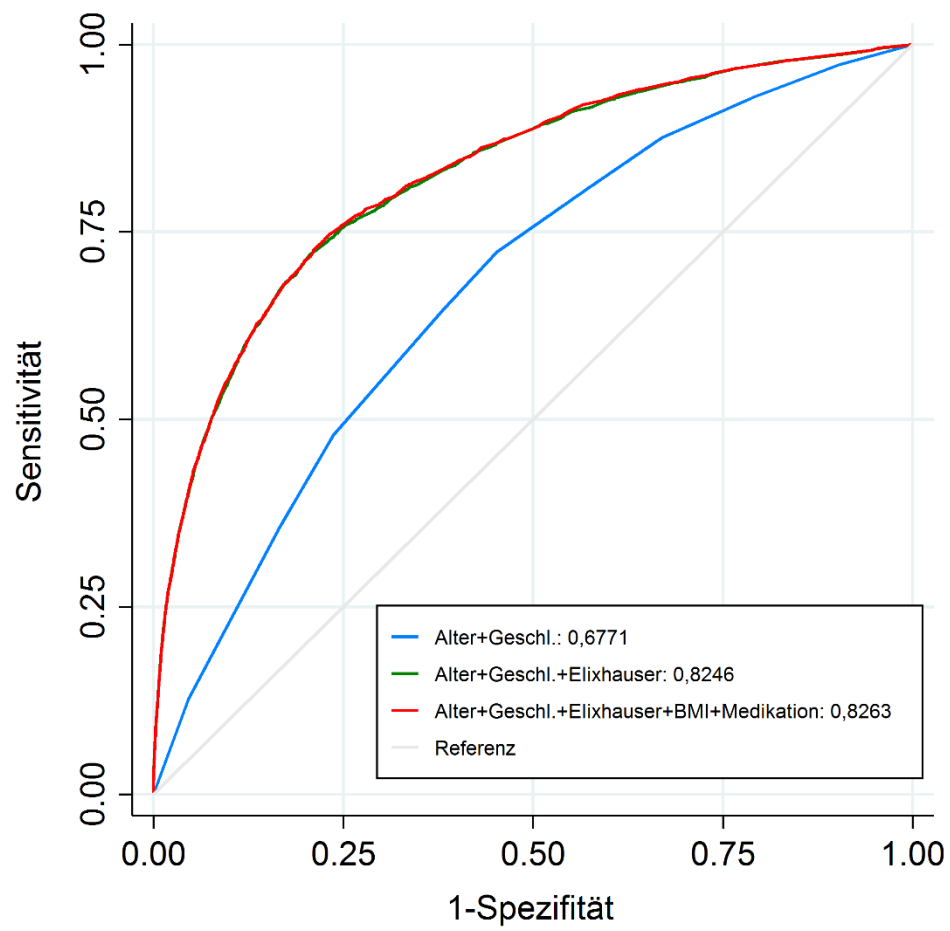
Wissenschaftliches Institut der AOK (WIdO) (2018a). <http://www.qualitaetssicherung-mit-routinedaten.de>.

Wissenschaftliches Institut der AOK (WIdO) (2018b). *QSR-Verfahren: Indikatorenhandbuch – Verfahrensjahr 2018*; <http://www.qualitaetssicherung-mit-routinedaten.de/downloads>.

Tabelle 1: Ergebnis der multivariaten Regressionsanalyse für die QSR-Indikatoren Revisions-OP und schwere Allgemeinkomplikation nach elektivem Hüftgelenkersatz (AOK-Daten 2014-2016)

Risikofaktor	Revisions-OP innerhalb von 365 Tagen	Schwere Allgemein- komplikation im Aufenthalt
	Odds Ratio (95%-KI)	Odds Ratio (95%-KI)
Alter (<i>Referenz: unter 60 Jahre</i>)		
60 bis 68 Jahre	-	1,39 (1,18-1,64)
69 bis 74 Jahre	-	1,77 (1,49-2,10)
75 bis 79 Jahre	-	2,47 (2,11-2,90)
über 79 Jahre	-	2,99 (2,55-3,51)
Geschlecht (<i>Referenz: männlich</i>)		
weiblich	0,88 (0,82-0,95)	0,67 (0,62-0,73)
BMI (<i>Referenz: BMI unter 30</i>)		
30 bis 34	1,25 (1,11-1,41)	-
35 bis 39	1,57 (1,38-1,79)	-
ab 40	2,40 (2,10-2,74)	1,50 (1,23-1,82)
Elixhauser-Begleiterkrankungen		
Alkoholabusus	1,85 (1,33-2,57)	-
Hypertonie, mit Komplikationen	-	1,39 (1,17-1,65)
Hypertonie, ohne Komplikationen	-	1,30 (1,17-1,44)
Blutungsanämie	-	2,00 (1,28-3,12)
Chronische Lungenerkrankung	1,22 (1,10-1,36)	1,25 (1,11-1,41)
Defizienzanämie	-	1,67 (1,30-2,15)
Depression	1,55 (1,37-1,76)	1,29 (1,12-1,49)
Diabetes, mit Komplikationen	1,38 (1,14-1,67)	1,42 (1,20-1,67)
Diabetes, ohne Komplikationen	1,11 (1,01-1,22)	1,17 (1,07-1,28)
Herzklappenerkrankung	-	1,75 (1,52-2,01)
Gewichtsverlust	1,81 (1,39-2,36)	2,21 (1,72-2,85)
Kardiale Arrhythmie	1,13 (1,02-1,25)	-
Koagulopathie	2,31 (1,95-2,73)	-
Herzinsuffizienz	1,26 (1,11-1,42)	2,50 (2,23-2,81)
Lähmung	1,67 (1,22-2,28)	9,92 (8,02-12,28)
Lebererkrankung	-	3,18 (2,51-4,03)
Nierenversagen/-insuffizienz	1,23 (1,10-1,37)	1,74 (1,57-1,93)
Periphere Gefäßerkrankung	-	1,59 (1,37-1,83)
Psychose	1,95 (1,23-3,08)	1,89 (1,15-3,13)
Lungenkreislaufstörungen	1,42 (1,05-1,92)	-
Rheumatoide Arthritis/Kollagenose	1,27 (1,05-1,54)	-
Flüssigkeits- und Elektrolytstörungen	1,76 (1,60-1,94)	4,24 (3,86-4,65)
Andere neurologische Erkrankungen	1,25 (1,01-1,54)	2,23 (1,88-2,65)
Antithrombotische Medikation im Vorjahr	-	1,21 (1,11-1,33)

Abbildung 1: ROC-Kurven der Modelle zur Prädiktion schwerer Allgemeinkomplikationen



3M Deutschland GmbH - Health Information Systems

Unternehmensporträt

3M ist ein 1902 gegründeter multinationaler Konzern mit Sitz in St. Paul, Minnesota, USA (<http://www.mmm.com>). 3M ist mit mehr als 90.000 Mitarbeitern in 200 Ländern vertreten und erzielte 2017 einen Umsatz von über 31 Mrd. US-Dollar. Grundlage für seine Innovationskraft ist die vielfältige Nutzung von 46 eigenen Technologieplattformen. Heute umfasst das Portfolio mehr als 55.000 verschiedene Produkte für fast jeden Lebensbereich.

Seit mehreren Jahrzehnten engagiert sich 3M weltweit auch im Bereich Informationstechnologie im Gesundheitswesen. Der Geschäftsbereich 3M Health Information Systems (HIS) entwickelt international anerkannte Patientenklassifikationssysteme für den stationären und ambulanten Sektor..

Die von 3M eingeführten Systematiken und Technologien gelten weltweit als anerkannter Standard und Benchmark. Neben den Schwerpunkten Patientenklassifikation und Qualitätsmanagements, werden von 3M HIS Wissensdatenbanken und Personenidentifikationssystemen (Master Patient Index) entwickelt. Weltweit arbeiten mehr als 1.400 Experten im Geschäftsbereich HIS.

In Deutschland ist 3M Health Information Systems seit 1996 als Geschäftsbereich der 3M Deutschland GmbH mit Sitz in Berlin und Neuss (<http://www.3m-drg.de/>) vertreten. In einer ca. 60 Personen zählenden Organisation finden sich Mediziner, Informatiker und Ökonomen, die aufgrund ihrer praktischen Erfahrung im Bereich des Gesundheitswesens und insbesondere der DRG Systematik und der Qualitätssicherung / des Qualitätsmanagements qualifiziert sind. Neben dem deutschen Markt werden auch Softwarelösungen für die Bereiche Patientenklassifikation (3M KODIP/360 Encompass) und DRG (Grouper) für den Schweizer Markt entwickelt. Fast 2.000 Kliniken in Deutschland und der Schweiz setzen Softwarelösungen oder Services von 3M HIS ein.

Ergänzend zu inhaltlichen Konzepten und Softwarelösungen bietet 3M im stationären Sektor Benchmark-Projekte mit den Schwerpunkten DRG und Qualität an. Insgesamt werden im Rahmen dieser Projekte jährlich über 7 Mio. Patientendatensätze ausgewertet. Die größten Projekte werden derzeit in Kooperation mit der Initiative Qualitätsmedizin (IQM) und dem QKK e. V. (Qualitätsindikatoren für kirchliche Krankenhäuser) durchgeführt.

Die 3M Deutschland GmbH ist Mitglied im Bundesverband der Hersteller von IT-Lösungen für das Gesundheitswesen (bvitg e. V.).

IQM Hearing „Qualitätsmessung in der Medizin – Risikoadjustierung und Risikoadjustierungsmodelle“

Risikoadjustierung auf Basis multipler Regressionsanalysen zur Ermittlung von individuellen Erwartungswerten

- Schmidt J., Graf. R. - 3M Health Information Systems

Abstract

Die vergleichende Qualitätsmessung hat zunehmend an Bedeutung gewonnen. Damit jedoch der Outcome möglichst der Behandlungsqualität entspricht ist, es wichtig mithilfe einer Risikoadjustierung die patientenindividuellen Risikofaktoren weitestgehend zu isolieren. Aus diesem Grund möchte die IQM auf Basis aktueller Methoden eine robuste Form der Risikoadjustierung von Indikatoren für eine vergleichende Darstellung identifizieren.

In diesem Abstract wird eine variable und flexible Methode der Risikoadjustierung vorgestellt. Die Flexibilität wird durch eine stufenweise Anpassung der Risikoadjustierung unter Berücksichtigung des medizinischen Kontextes des Indikators gewährleistet. Außerdem wird dessen Aussagekraft im Verhältnis zu dem zugrunde liegenden Aufwand in die Auswahl der Methode einbezogen.

Bei den Indikatoren, welche in der IQM verwendet werden, liegt vornehmlich eine binäre Outcomegröße vor. Aus diesem Grund wird eine Modellbildung mithilfe der multiplen logistischen Regression (MLR) durchgeführt. Diese macht es möglich, neben dem Einfluss der Risikofaktoren auf die Zielgröße, auch deren Interaktion miteinander zu berücksichtigen [1]. Die relevanten Risikofaktoren werden unter Beachtung des Modells des patientenindividuellen Schweregrades nach Valderas et al. [2] ermittelt. Neben der Krankheitslast finden auch weitergehende Aspekte Berücksichtigung, damit bei einem Vergleich das krankenhausindividuelle Patientenspektrum Würdigung findet.

Ziel ist es, ein robustes, medizinisch nachvollziehbares und transparentes Modell je Krankheitsbild / Qualitätsindikator zu entwickeln. Die Komplexität der möglichen Modelle geht dabei mit einem steigenden Aufwand einher. Abbildung 1 zeigt den stufenweisen Aufbau immer komplexer werdender Modellvarianten mit zunehmend spezifischer Berücksichtigung der Indexerkrankungen.

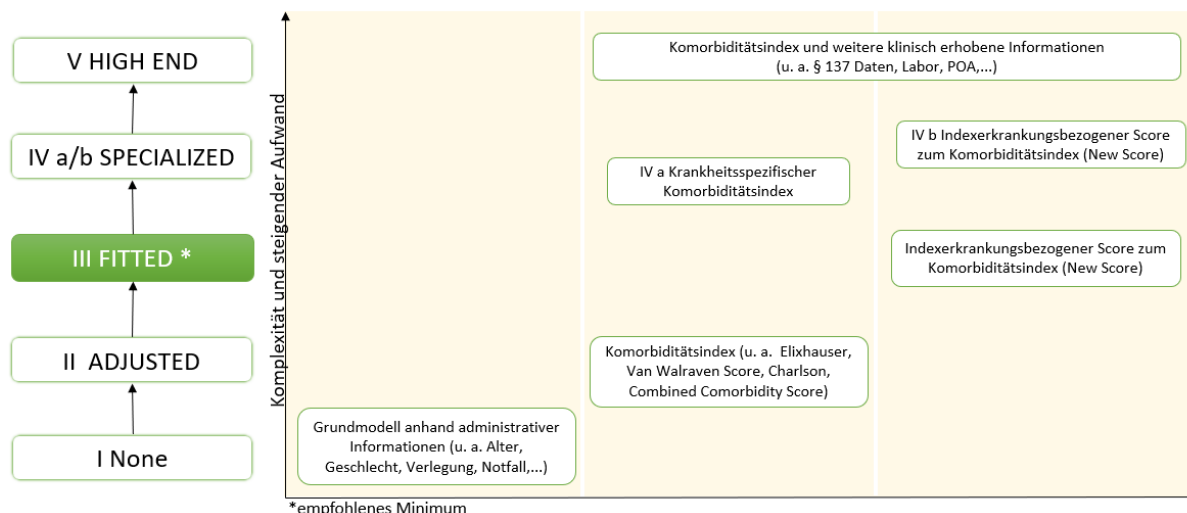


Abbildung 1: Modellvarianten

Die relevanten Risikofaktoren werden nach Empfehlungen gültiger Leitlinien und auf Basis aktueller Literatur ausgewählt. Anschließend werden die Einzelfaktoren anhand der relativen und absoluten Häufigkeiten analysiert und die Signifikanz des Einflusses auf die Zielgröße getestet. Es können u.a. Variablen wie Alter, Geschlecht, Verlegung und Notfall eingehen. Der Schwerpunkt bei der Modellentwicklung liegt jedoch in der Wahl und Anpassung von Komorbiditätsscores an die Indexerkrankung. Berücksichtigt werden u.a. der Charlson Index [3], Elixhauser Score [4], van Walraven Score [5] und Combined Comorbidity Score [6]. Es werden je nach Modellvariante verschiedene valide Komorbiditätsindizes als Summenscore getestet oder speziell angepasste Scores zur Indexerkrankung nach methodischen Vorgehen von Quan entwickelt [7]. Nach sukzessiver Einbeziehung der

Risikofaktoren wird das abschließende Modell geprüft und validiert. Hierfür wird als globales Gütemaß das Pseudo R^2 von Nagelkerke herangezogen [8]. Als Maß der Diskrimination wird die c-Statistik verwendet und anschließend visualisiert als ROC Kurve [9, 10]. Die Modellentwicklung erfolgt zunächst auf einem eigenen Entwicklungsdatensatz. Das Modell selbst wird anschließend an einem Validierungsdatensatz evaluiert.

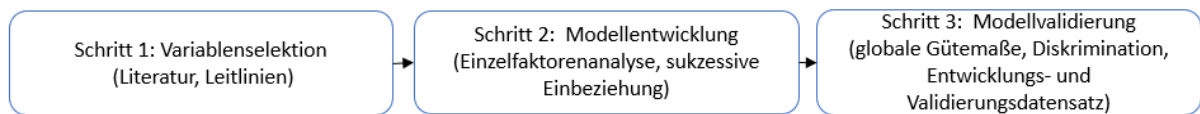


Abbildung 2: Methodisches Vorgehen

Das methodische Vorgehen (Abb.2) erfolgt bei jeder Modellvariante. Anhand des Indikator GIQI- 01.1 „Hauptdiagnose Herzinfarkt (>19Jahre), Anteil Todesfälle“ wurden bis Stufe IVa alle Modelle beispielhaft entwickelt und miteinander verglichen. Grundlage dafür sind 65131 Fälle, die unter die Nennerdefinition fallen. Die Krankenhaussterblichkeit liegt bei 7.95% (n=5181 Fälle). Die Modelle werden zusätzlich mit einem Modell aus den Variablen Alter und Geschlecht verglichen. Die Tabelle 1 und Abbildung 3 zeigen die Modellvarianten im Vergleich. Die Modellvarianten II bis IVb befinden sich nach Ash und Schwartz et al. [11] bei der Beurteilung der Diskriminationsfähigkeit mit Werten zwischen 0.8 und 0.9 im exzellenten Bereich. Die höchste prognostische Güte und Diskrimination erreichen die Stufen III und IVb. Während in der Stufe III die 30 Diagnosegruppen des Elixhauser bewertet wurden, wurde bei der Stufe IVb ein Score auf die neun Gruppen der Ontario AMI prediction rule (OAMIPR) entwickelt [12]. Die OAMIPR ist ein speziell für Routinedaten entwickelter Komorbiditätsindex mit dem Ziel die Sterblichkeit nach Myokardinfarkt vorherzusagen.

Modellvariante / Stufen	R ² *	AUC	KI 95%**
Alter/Geschlecht	0.051	0.654	[0.650 - 0.657]
I None Grundmodell	0.199	0.779	[0.775 - 0.784]
II Adjusted ECI mit van Walraven Score	0.295	0.850	[0.847 - 0.854]
III Fitted ECI mi New Score	0.348	0.870	[0.866 - 0.874]
IVa Specialized OAMIPR	0.325	0.860	[0.856 - 0.864]
IVb Fitted OAMIPR mit New Score	0.451	0.895	[0.891 - 0.899]

Tabelle 1: Vergleich der Modellvarianten

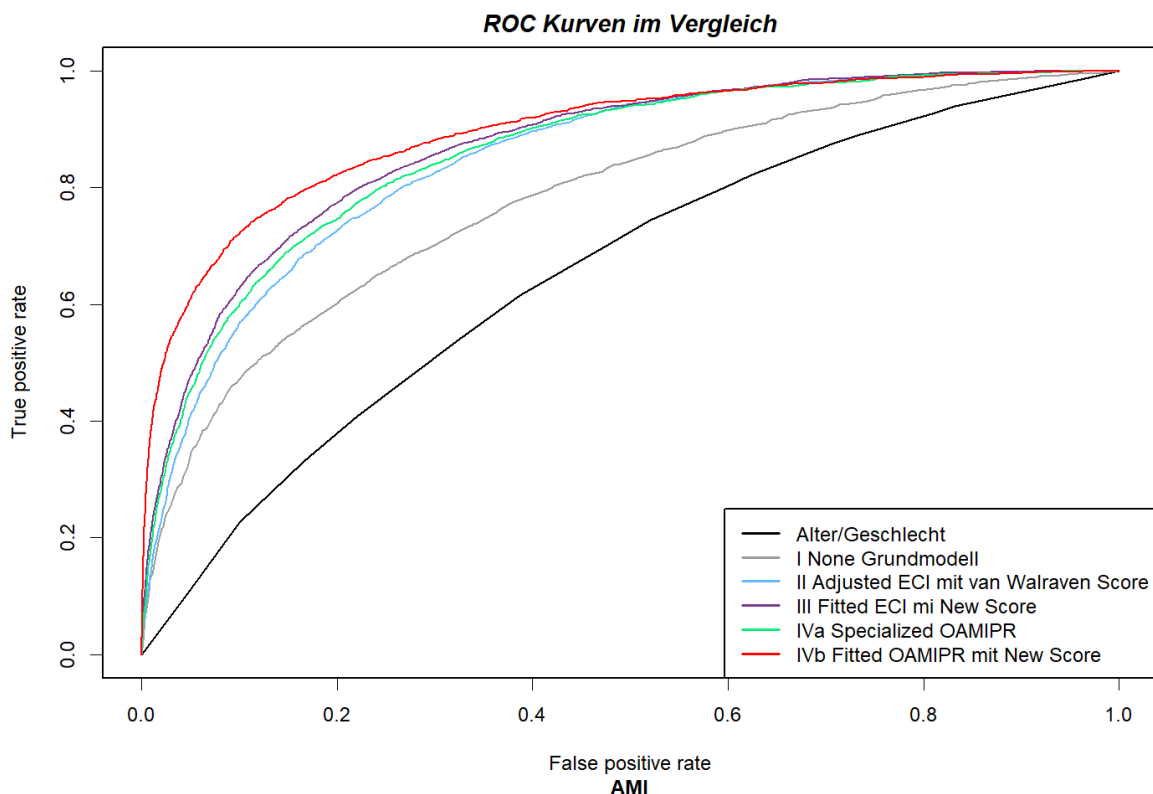


Abbildung 3: ROC Kurven im Vergleich

Eine Risikoadjustierung mittels MLR basierend auf Routinedaten kann immer nur zu einem gewissen Grad die Zielgröße schätzen. Dabei enthalten Routinedaten nicht alle Einflussgrößen. Ebenso stellen Komorbiditätsindizes nie die vollständige Krankheitslast des Patienten dar. Hier können klinische Daten zu einer deutlichen Modellgüte beitragen [13]. Neben fehlender Information können zusätzlich die Unterschiede zwischen Begleiterkrankung und Komplikation nicht ausreichend identifiziert werden, um mögliche Störgrößen auszuschließen. Bei den in dieser Arbeit verwendeten Daten liegt eine Besonderheit bei Fällen mit einer Verweildauer von einem Tag vor. Sie unterscheiden sich hinsichtlich der Kodierung von denen mit längerer Verweildauer ($VWD=1 \pm 6.03$ ND, $VWD > 1 \pm 9.94$). Ein Grund kann sein, dass mögliche Erkrankungen dieser Patienten durch das frühzeitige Versterben keinen Aufwand im Krankenhaus verursacht haben bzw. dort nicht bekannt waren. Auch ist die beobachtete Letalität (29.92%) bei diesen Patienten höher als bei Patienten mit einer Verweildauer > 1 (5.34%). Durch die mögliche Unterkodierung dieser Fälle und die hohe Letalität ergibt sich eine beobachtbare Störgröße welche für diese beispielhafte Berechnung bewusst im Modell gelassen wurde, um den Indikator in der jetzt gültigen Spezifikation verwenden zu können.

Bei einer Risikoadjustierung ist zu empfehlen, dass für jeden Indikator ein eigenes Risikomodell erstellt wird. Wenn verfügbar sollte ein krankheitsspezifischer Score mit individueller Anpassung auf die Indexerkrankung verwendet werden. Aufwand und Nutzen der individuellen Anpassung von Scores sollte je nach Indikator und vorhandener Literatur abgewogen werden. Auch gilt es solche Indikatoren zu identifizieren, für die ein Risikomodell mit Komorbiditäten keinen Mehrwert bietet. Das können Indikatoren sein, welche Patienten mit vielen unterschiedlichen Grund- und Begleiterkrankungen als Basis haben, z.B. Beatmung > 24 h. Bei diesem Indikator besteht die Gefahr, dass durch die Verwendung von Komorbiditätsscores Komplikationen und Komorbiditäten vermischt werden. Dagegen trägt bei einem Indikator mit gleicher Indexerkrankung die Verwendung eines passenden Scores dazu bei, die Ungerechtigkeit bei Vergleichen deutlich zu mindern. Ebenso kann die Modellgüte durch das zusätzliche Erfassen von klinischen Faktoren gesteigert werden.

Es ist anzuraten nach Prüfung der Einzelfaktoren und genaueren Analyse zu überlegen, ob bestimmte Störgrößen aus dem Indikator vorab auszuschließen sind. So kann es sinnvoll sein bei einigen Indikatoren nur Fälle mit einer bestimmten Mindestverweildauer in die Nennerpopulation einzubeziehen [14].

Um eine Entscheidung treffen zu können, welches Modell zu verwenden ist, sollten die berechneten Risikomodelle jeder Stufe in einer Übersicht aufbereitet und die Ergebnisse in Form eines Berichts beschrieben werden. Dieser muss auch mögliche Erkenntnisse aus der aktuellen Literatur zu diesem Krankheitsbild und den Auswirkungen und Möglichkeiten durch die Verwendung von Routinedaten beinhalten.

Literatur

- 1 Kreienbrock L, Pigeot I, Ahrens W. Epidemiologische Methoden. 5. Aufl. Heidelberg: Spektrum Akademischer Verlag; 2012
- 2 Valderas JM, Starfield B, Sibbald B et al. Defining comorbidity: implications for understanding health and health services. *Annals of family medicine* 2009; 7: 357–363
- 3 Charlson ME, Pompei P, Ales KL et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of chronic diseases* 1987; 40: 373–383
- 4 Elixhauser A, Steiner C, Harris DR et al. Comorbidity measures for use with administrative data. *Medical care* 1998; 36: 8–27
- 5 van Walraven C, Austin PC, Jennings A et al. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. *Medical care* 2009; 47: 626–633
- 6 Gagne JJ, Glynn RJ, Avorn J et al. A combined comorbidity score predicted mortality in elderly patients better than existing scores. *Journal of clinical epidemiology* 2011; 64: 749–759
- 7 Quan H, Li B, Couris CM et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *American journal of epidemiology* 2011; 173: 676–682
- 8 Nagelkerke NJD. A note on a general definition of the coefficient of determination. *Biometrika* 1991; 78: 691–692
- 9 Sing T, Sander O, Beerenwinkel N et al. ROCR: visualizing classifier performance in R. *Bioinformatics (Oxford, England)* 2005; 21: 3940–3941

- 10 *Kundu S, Aulchenko YS, van Duijn CM et al.* PredictABEL: an R package for the assessment of risk prediction models. *European journal of epidemiology* 2011; 26: 261–264
- 11 *Ash AS, Schwartz M.* Empirically Evaluating Risk Adjustment Models. Risk adjustment for measuring health care outcomes. Chicago: Health Administration Press; 2013: 249–300
- 12 *Tu JV, Austin PC, Walld R et al.* Development and validation of the Ontario acute myocardial infarction mortality prediction rules. *Journal of the American College of Cardiology* 2001; 37: 992–997
- 13 *Castro-Dominguez Y, Dharmarajan K, McNamara RL.* Predicting death after acute myocardial infarction. *Trends in cardiovascular medicine* 2018; 28: 102–109
- 14 *Becker A, Stausberg J, Fischer B et al.* Risikoadjustierung von Qualitätsindikatoren. Eine Positionsbestimmung der DKG-Expertengruppe „Qualitätsmessung und Risikoadjustierung“. *das Krankenhaus* 2016: 954–963

Anforderungen an eine Risikoadjustierung zur Qualitätsmessung medizinischer Behandlungsergebnisse

Ulrike Nimptsch, Technische Universität Berlin, Fachgebiet Management im Gesundheitswesen

Verfahren zur Risikoadjustierung werden im Rahmen von Qualitätsmessungen eingesetzt, um Behandlungsergebnisse unabhängig von solchen Faktoren beurteilen zu können, die nicht in der Verantwortung des Leistungserbringers liegen. Dabei ergeben sich spezifische Anforderungen, die auch vom Anwendungszweck der Qualitätsmessung abhängig sind.

Für Anwendungszwecke, die primär auf die einrichtungsinterne medizinische Ergebnisverbesserung ausgerichtet sind, dient die Risikoadjustierung vornehmlich dazu, Auffälligkeiten mit hoher „Trefferquote“ erkennen zu können. Dies ist besonders wichtig wenn Qualitätskennzahlen als Aufgreifkriterien für weitergehende Fallanalysen, z.B. im Rahmen von Peer Reviews, genutzt werden. Bei Qualitätsmessungen, die dem Public Reporting dienen, zielt die Risikoadjustierung darauf ab, einen fairen Vergleich zwischen den Leistungserbringern und damit eine faire Bewertung der Qualität sicherzustellen [1].

Eine Risikoadjustierung kann mit unterschiedlichen Verfahren durchgeführt werden, beispielsweise als Risikodifferenzierung oder Risikostratifizierung oder anhand von Erwartungswerten, die auf der Grundlage von Standardisierungen oder Regressionsverfahren berechnet werden können. Unabhängig davon welches Verfahren zur Anwendung kommt, muss bei der Auswahl der Merkmale, die in der Risikoadjustierung berücksichtigt werden, deren inhaltliche Eignung vor dem Hintergrund des Ziels der Qualitätsmessung bewertet werden [2].

Dies betrifft beispielsweise Begleiterkrankungen. Hierbei ist zu vermeiden, dass mögliche Komplikationen fälschlicherweise als Begleiterkrankung gewertet werden. Eine Adjustierung nach Komplikationen kann im Rahmen von Qualitätsmessungen aus naheliegenden Gründen nicht zielführend sein. Auch der Schweregrad einer Erkrankung ist eine mögliche Einflussgröße für das Behandlungsergebnis. Hier muss beachtet werden dass die Erkrankungsschwere (z.B. im Zusammenhang mit operativen Eingriffen) auch die Qualität der Indikationsstellung widerspiegeln kann. Der Einbezug von Merkmalen des Behandlungsprozesses (z.B. bestimmte Operationsmethoden) in die Risikoadjustierung kann problematisch sein, sofern diese als Bestandteil der Behandlungsqualität zu werten sind [3]. Die Berücksichtigung sozioökonomischer Merkmale (z.B. Einkommen oder Bildungsstatus) in der Risikoadjustierung einer medizinischen Qualitätsmessung muss vor dem Hintergrund ethischer Überlegungen kritisch diskutiert werden [2].

Bei allen Risikoadjustierungsfaktoren ist auch die Verlässlichkeit und Vollständigkeit der Dokumentation von Bedeutung. Ein unterschiedliches Dokumentationsverhalten verschiedener Einrichtungen kann die Vergleichbarkeit risikoadjustierter Messwerte beeinträchtigen. Auch unerwünschte Verhaltensweisen bei den Leistungserbringern (im Sinne von strategischer Dokumentation) sind möglich [1].

Die Auswahl von Faktoren für eine Risikoadjustierung muss unter Berücksichtigung der Ziele der Qualitätsmessung sorgfältig abgewogen werden. Wenn ein Public Reporting zur Unterstützung der Auswahlentscheidung von (potentiellen) Patienten oder Einweisern angestrebt wird muss zusätzlich beachtet werden, dass auch risikoadjustierte Messwerte unsicher sind wenn sie auf kleinen Fallzahlen beruhen. Funnel-Plots können helfen, die

statistische Unsicherheit der Messwerte zu visualisieren [4]. Die Verwendung von empirischen Bayes-Verfahren, wie beispielsweise eine Reliabilitätsadjustierung von Messwerten [5], erscheint dagegen diskussionswürdig.

- [1] Mansky T, Nimptsch U. Medizinische Qualitätsmessung im Krankenhaus – Worauf kommt es an? Z Evid Fortbild Qual Gesundheitswes 2014; 108(8-9): 487-94.
- [2] Lux G, Biermann J, Dahl H, Matusiewicz D, Mostardt S, Nimptsch U, Wasem J, Walendzik A. Risikoadjustierung und Komorbiditäten. In: Swart E, Ihle P, Gothe H, Matusiewicz D (Hrsg.) Routinedaten im Gesundheitswesen. Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven. Bern: Verlag Hans Huber 2014, 411-23.
- [3] Nimptsch U, Spoden M, Mansky T. Variablendefinition in fallbezogenen Krankenhausabrechnungsdaten – Fallstricke und Lösungsmöglichkeiten. Gesundheitswesen (eingereicht).
- [4] Spiegelhalter D. Funnel plots for comparing institutional performance. Statist Med 2005; 24: 1185-1202.
- [5] Dimick JB, Staiger DO, Birkmeyer JD. Ranking Hospitals on Surgical Mortality: The Importance of Reliability Adjustment. Health Serv Res 2010; 46 (6 p1): 1614-29.

Zentrum für evidenzbasierte Gesundheitsversorgung (ZEGV)

Das ZEGV dient als gemeinsame Einrichtung der Medizinischen Fakultät und des Universitätsklinikums Carl Gustav Carus an der TU Dresden der Integration von klinischer und methodischer Expertise und der Stärkung des Profilschwerpunktes Prävention und Versorgungsforschung der Dresdner Universitätsmedizin. Ziel des ZEGV ist es, wirksame Wege zur Verbesserung der Qualität, Effektivität, Effizienz und des Nutzens medizinischer Interventionen unter Routinebedingungen aufzuzeigen. Das ZEGV hat sich somit die systematische Untersuchung des Transfers von medizinischer Innovation in die Versorgungspraxis unter Berücksichtigung der Wirksamkeit, Wirtschaftlichkeit und Sicherheit wie auch von Patienten- und Populationspräferenzen zur Aufgabe gemacht. Forschungsschwerpunkte des ZEGV liegen in den Bereichen Versorgungsepidemiologie, Outcomes-Forschung, Effectiveness-Forschung, Versorgungsforschung und Nutzenforschung. Am ZEGV sind über 30 Mitarbeiter unterschiedlicher Profession beschäftigt, darunter auch Biometriker, die über umfangreiche Erfahrungen in der erfolgreichen Durchführung und Analyse größerer Studien der Evidenzbasierten Medizin und Versorgungsforschung verfügen.

Der gemeinsame Schwerpunktbereich „Qualitäts- und Patientensicherheitsforschung“ des ZEGV und des Zentralbereichs Qualitäts- und Medizinisches Risikomanagement des Universitätsklinikums Dresden untersucht systematisch Maßnahmen und Instrumente des Qualitäts- und Medizinischen Risikomanagements hinsichtlich des Nutzens für den Patienten, der Wirksamkeit für den Behandlungsablauf und der Effizienz für das medizinische Personal. Dabei besteht seit längerer Zeit eine erfolgreiche Zusammenarbeit mit IQM. Derzeit bearbeitet das ZEGV in Zusammenarbeit mit IQM ein Projekt der Qualitätsforschung. Im Rahmen der IMPRESS-Studie wird die Wirksamkeit des IQM Peer Reviews anhand einer multizentrischen randomisierten kontrollierten Interventionsstudie untersucht.

Der Direktor des ZEGV, Prof. Dr. med. Jochen Schmitt, ist Facharzt für Dermatologie, Epidemiologe und seit 2011 Inhaber des Lehrstuhls für Sozialmedizin und Versorgungsforschung an der Medizinischen Fakultät Carl Gustav Carus der TU Dresden. Prof. Schmitt hat langjährige Erfahrung in der Konzeption, Durchführung und Analyse von klinischen, versorgungsepidemiologischen und gesundheitsökonomischen Studien verschiedenster Art und ist Autor von über 200 Publikationen in Peer-review Zeitschriften.

Stratifizierung und Risikoadjustierung

Vorschlag eines evidenzbasierten Algorithmus' zur Erstellung von interpretierbaren und zielgerichteten Qualitätsindikatoren

Prof. Dr. Jochen Schmitt; Martin Rößler, M.Sc.; Dr. Olaf Schoffer

Zentrum für Evidenzbasierte Gesundheitsversorgung (ZEGV)
Universitätsklinikum Carl Gustav Carus und Medizinische Fakultät, TU Dresden

Hintergrund: Ein werte- und qualitätsorientiertes Gesundheitssystem setzt eine faire, interpretierbare und zielgerichtete Bewertung der Leistungserbringer voraus. Für alle Zwecke der Qualitätsmessung ist das Thema Risikoadjustierung zentral und wird entsprechend kontrovers diskutiert. Durch Zugang zu umfangreichen Routinedaten, die Repräsentation verschiedener Krankenhäuser sowie klinische und methodische Expertise bietet IQM die Möglichkeit, ein evidenzbasiertes Vorgehen zur Risikoadjustierung bei der stationären Qualitätsmessung zu entwickeln und zu implementieren.

Ziel: Ziel des Beitrags ist es, ausgehend von den epidemiologischen Grundkonzepten Confounding und Effektmodifikation einen Algorithmus zur Bildung von validen und vergleichbaren Qualitätsindikatoren zu entwickeln.

Methodik: Aufbauend auf theoretischen Überlegungen werden zunächst Hypothesen über patientenindividuelle und klinikbezogene Risikofaktoren und Effektmodifikatoren generiert. Diese werden auf Basis eines geeigneten statistischen Modells empirisch überprüft. Aus der Synthese von Theorie und Datenevidenz wird ein Set relevanter Variablen für die Risikoadjustierung und Stratifizierung abgeleitet. Als Datengrundlage dienen die gemäß §21 KHEntgG erhobenen Routinedaten von 237 IQM-Mitgliedskliniken, die bei der vom Innovationsfonds geförderten IMPRESS-Studie teilnehmen. Der Algorithmus wird am Beispiel des Indikators Sterblichkeit von Patienten mit Beatmung > 24 Stunden illustriert. Alle regulatorischen und datenschutzrechtlichen Voraussetzungen für die Durchführung der präsentierten Analysen sind erfüllt.

Ergebnisse: Potentielle Risikofaktoren neben Alter und Geschlecht sind Grunderkrankung/Indikation, Komorbiditäten, Umstände der Aufnahme und des Krankenhausaufenthaltes sowie Struktureigenschaften des behandelnden Krankenhauses. Bei den Komorbiditäten ist kritisch zu hinterfragen, ob beobachtete Assoziationen plausibel und kausal sind und ob Informationen ohne Angabe zu „present on admission“ (POA) verwertbar sind. Plausible, zu explorierende Proxies für die Schwere und Komplexität der Grunderkrankung sind der Patientenzugang (Notfall/Verlegung), die Behandlung in mehreren Fachabteilungen eines Krankenhauses sowie die Verlegung in ein anderes Krankenhaus. Das Merkmal Universitätsklinikum sollte als möglicher Effektmodifikator exploriert werden.

Diskussion: Für den Indikator Mortalität bei Beatmung > 24 Stunden konnten auf Basis empirischer Analysen relevante Risikofaktoren bestätigt und Effektmodifikatoren identifiziert werden. Für Effektmodifikatoren auf Klinikebene wird im Sinne einer fairen, interpretierbaren und zielgerichteten Bewertung eine stratifizierte Darstellung empfohlen.

Department Health Sciences & Health Policy der Universität Luzern

Institutsbeschreibung: Das Department Health Sciences & Health Policy der Universität Luzern betreibt Forschung und Ausbildung in den Bereichen Gesundheit und Behinderung unter Berücksichtigung bio-psycho-sozialer Gesichtspunkte. Als Institut mit umfassendem Interesse am gesamten Versorgungs- und Gesundheitssystem sind wir nicht nur wissenschaftstiftend in der Forschung und Lehre tätig, sondern machen die erworbenen Kenntnisse und Ergebnisse auch der Öffentlichkeit, den politischen Entscheidungsträgern und den Gesundheitsberufen zugänglich. Nebst unseren Master- und Doktoratsstudiengängen in Health Sciences und Medizin, betreiben wir beispielsweise mit dem Swiss Learning Health System (SLHS) eine Plattform zur Vernetzung der Gesundheitssystem- und Versorgungsforschung mit Politik und Praxis. Unsere Aktivitäten, wie Policy Briefs, Stakeholder-Dialoge und die Entwicklung von Gesundheitsmanagement-Tools, dienen dazu, Wissen aufzubauen, Interaktionen zu erleichtern sowie die beteiligten Akteure des Gesundheitssystems zu engagieren, um gemeinsam Veränderungen herbeizuführen, das Gesundheitssystem in der Schweiz zu stärken und die Gesundheit und das Wohlbefinden in der Bevölkerung zu verbessern.

Reliabilität der Inpatient Quality Indicators in der vergleichenden Darstellung der Qualitätsleistung von Kliniken¹

Dr. med., Dr. sc. nat. Michael Havranek, Department Health Sciences & Health Policy, Universität Luzern

Mittels Qualitätsindikatoren sollen Qualitätsleistungen von Kliniken gemessen werden. Als Massangaben beinhalten solche Indikatoren jedoch stets einen Messfehler (z.B. durch unkorrigierte Risikofaktoren der Patienten oder generell durch Random Noise), weshalb die tatsächliche Qualitätsleistung der einzelnen Kliniken nie mit Sicherheit bestimmt, sondern lediglich geschätzt werden kann. Verwendet man Qualitätsindikatoren nun zunehmend zur vergleichenden Darstellung von Kliniken (anstatt nur zur Identifikation von Auffälligkeiten), nimmt auch die Frage nach der Reliabilität (Zuverlässigkeit) dieser Indikatoren zur Schätzung und zum Vergleich der tatsächlichen Qualitätsleistung der Kliniken an Bedeutung zu.

Als Bestandteil eines aktuellen Forschungsprojektes untersuchten wir deshalb die Reliabilität von knapp 30 Swiss Inpatient Quality Indicators (CH-IQIs) anhand retrospektiver, anonymisierter Daten aller Schweizer Spitäler und Kliniken (beginnend im Jahr 2010), indem Qualitätsleistungen der Spitäler und Kliniken aus vorhergehenden Jahren verwendet wurden, um Qualitätsleistungen in Folgejahren vorherzusagen. Dabei wurde einerseits untersucht, welche der Indikatoren unabhängig von der gewählten Risikoadjustierungsmethode zufriedenstellende Reliabilitäten aufweisen und welche nicht. Andererseits wurde geprüft, ob durch Verwendung der im englischsprachigen Raum angewandten Shrinkage Estimators (Schrumpfungsgewichte) als Bestandteil des Risikoadjustierungsmodells (z.B. in Form von hierarchischen Regressionsmodellen mit einem Random Effect oder aber in Form einer sogenannten Reliabilitätsadjustierung [1]) und/oder durch gezielte, datengestützte Selektion der zu verwendenden Risikofaktoren (wie z.B. Alter, Geschlecht, Komorbiditäten etc.) die Reliabilität der Indikatoren verbessert werden kann.

Die vorläufigen Resultate dieser Studie werden nun im Rahmen des öffentlichen Hearings der IQM (Initiative Qualitätsmedizin) erstmals vorgestellt. Einerseits wird aufgezeigt, dass verschiedene der rund 30 geprüften CH-IQIs unabhängig von der Wahl der Risikoadjustierungsmethode zu tiefe Reliabilitäten für Qualitätsvergleiche aufweisen. Andererseits wird demonstriert, dass die im englischsprachigen Raum angewandten Shrinkage Estimators zu signifikanten Reliabilitätsverbesserungen der Indikatoren führen. Schliesslich wird ein datenbasiertes Selektionsvorgehen vorgestellt, um bei der Wahl der Risikoadjustierungsfaktoren nicht nur die Area under the Receiver Operator Characteristic Curve (AUC), sondern auch den Einfluss der Anzahl und Auswahl der Risikoadjustierungsfaktoren auf die Reliabilität der Indikatoren als Kriterium zu berücksichtigen.

¹ aktualisiert am 25.01.2019

Damit zielt unser Beitrag nicht primär darauf ab, bestimmte Risikofaktoren vorzuschlagen oder ein konkretes Risikoadjustierungsmodell zu präsentieren. Sondern es sollen die Relevanz der Reliabilität von Vergleichsindikatoren angesprochen und statistische Methoden zur Reliabilitätsverbesserung demonstriert werden. Zusammenfassend argumentieren wir dafür, dass bereits bei der Wahl der zu Vergleichszwecken verwendeten Indikatoren deren Reliabilität berücksichtigt werden sollte, dass eine Reliabilitätsadjustierung mit Shrinkage Estimators Bestandteil des Risikoadjustierungsmodells sein sollte und dass bei der Wahl der Risikoadjustierungsfaktoren nicht nur die AUC, sondern auch deren Einfluss auf die Reliabilität berücksichtigt werden sollte.

Literatur: [1] Dimick JB, Staiger DO, Birkmeyer JD. Ranking Hospitals on Surgical Mortality: The Importance of Reliability Adjustment. Health Serv Res 2010; 45 (6 Pt 1): 1614-29.

Kontakt: Dr. med., Dr. sc. nat. Michael Havranek, Department Health Sciences & Health Policy, Universität Luzern, Frohburgstrasse 3, 6002 Luzern, Schweiz, +41 79 632 99 75, michael.havranek@unilu.ch.

PD Dr. med. Günther Heller

Studium und akademischer Werdegang

Studium der Soziologie und Medizin in Mannheim, Berlin, Frankfurt und Heidelberg

1992 Staatsexamen in Humanmedizin,

1996 Dissertation am Deutschen Krebsforschungszentrum in Heidelberg.

2006 Habilitation und Ernennung zum Privatdozent, Fakultät für Medizin der Universität Marburg.

Berufliche Tätigkeit

1992-1994 Chirurgischen Klinik am Heinrich-Lanz-Krankenhaus in Mannheim

1995-1996 Zentrum für Epidemiologie und Gesundheitsforschung in Berlin

1996-2002 Institut für Sozialmedizin und Medizinische Soziologie der Universität Marburg

2002-2010 Wissenschaftliches Institut der AOK (WIdO)

2010-2015 AQUA-Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen

seit Oktober 2015 Institut für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTIG).

Die hier dargestellte Arbeit wurde nicht im Rahmen der beruflichen Tätigkeit für das IQTIG verfasst

Zur Bedeutung der Risikoadjustierung in der Qualitätsmessung

Günther Heller

Hintergrund: Mit einer gerechten Erfassung und Darstellung der Ergebnisqualität wird seit jeher die Forderung nach einer ausreichenden Risikoadjustierung verknüpft und allgemein breit akzeptiert. So findet sich bei der Umschreibung der Aufgaben des Institutes nach § 137a SGB V die Aufgabe „abgestimmte risikoadjustierte Indikatoren“ zu entwickeln und „einrichtungsbezogen vergleichende risikoadjustierte Übersichten über die Qualität in maßgeblichen Bereichen der stationären Versorgung“ zu entwickeln. Allerdings stellt die Risikoadjustierung nur eines von mehreren Problemen bei der Messung von Ergebnisqualität dar.

Ziel dieser Arbeit ist es:

- 1) häufige Missverständnisse bei der Entwicklung von Risikoadjustierungsmodellen zu adressieren,
- 2) die Risikoadjustierung in den Kontext anderer Messprobleme bei der Erhebung von Versorgungsqualität zu stellen.
- 3) Lösungen vorzuschlagen, wie im Kontext einer freiwilligen Klinikinitiative die Qualitätsmessung im Allgemeinen und die Risikoadjustierung im Besonderen verbessert werden kann.

1a) Zahlreiche Arbeiten nutzen bei der Entwicklung und Bewertung von Risikoadjustierungsmodellen diverse Fitmaße, um die „Güte des Modells“ zu beurteilen (Becker et al. 2016). Anschließend wird dies implizit mit der Güte einer Risikoadjustierung gleichgesetzt. Dies scheint nur begrenzt sinnvoll.

Einerseits muss beachtet werden, dass nur Risikofaktoren für einen Einrichtungsvergleich bestimmte Eigenschaften erfüllen müssen. Sie dürfen bspw. nicht durch die Einrichtung beeinflussbar sein, so dass ein Risikoadjustierungsmodell für einen Einrichtungsvergleich ein unvollständiges Prognosemodell darstellt und je nach betrachtetem Outcome sehr unterschiedliche Fitmaße resultieren (Heller 2017, Günster & Jeschke 2019).

Andererseits ist zu beachten, dass zahlreiche dieser Maße umso bessere Ergebnisse zeigen, je mehr Risikofaktoren im Modell verwendet werden. Einfache Modellüberlegungen können allerdings zeigen, dass dies auch zu einer Verschlechterung der Vergleichbarkeit führen kann, wenn entsprechende Risikofaktoren nicht einheitlich kodiert werden (Heller & Schnell 2007).

Es macht demnach nur Sinn, derartige Maße zu vergleichen, wenn die gleichen Endpunkte in vergleichbaren Settings betrachtet werden und sicher gestellt werden kann, dass das Verständnis und die Kodierung der Risikofaktoren ausreichend homogen ist. Grundsätzlich sollte man sich bei der Einführung bzw. Nutzung eines potentiellen Risikofaktors immer die Frage stellen, ob dies zu einer Besserung oder Minderung der Vergleichbarkeit führt.

1b) Es ist üblich die Entwicklung von Risikoadjustierungsmodellen mit einer Suche / Identifikation möglichst aller relevanten Risikofaktoren zu beginnen. Unglücklicherweise kann danach immer noch bezweifelt werden, ob die Personen innerhalb der so definierten Gruppen vergleichbare Risiken aufweisen („constant risk fallacy“, Nichols 2007). Eine Alternative wäre, Gruppen mit konsentiertem vergleichbarem Risiko zu identifizieren (Heller 2017). Dies würde die Annahmen der zu entwickelnden Risikoadjustierung deutlicher herausstellen und gleichzeitig die Annahmen der „constant risk fallacy“ klarer adressieren.

2) Grundsätzlich wird in der Messtheorie oder Messtechnik bei einer Messung davon ausgegangen, dass ein gemessener Wert sich aus einem wahren Wert und einem Messfehler zusammensetzt. Dabei können die Messfehler in zufällige und systematische Messfehler unterteilt werden. Die Größe

der zufälligen Messfehler lässt sich mit Verfahren der schließenden oder Inferenzstatistik errechnen bzw. schätzen. Die Risikoadjustierung ist dagegen ein Versuch, systematische Fehler auszugleichen, die durch unterschiedlich erkrankte Populationen in den betrachteten Einrichtungen entstehen. Allerdings sind Probleme der Datenvalidität ebenfalls systematischen Fehlern zuzuordnen (Heller 2017).

Wenn nun Einrichtungen mit Fallzahlen von 20 oder 30 Fällen bei einer Prävalenz eines Ergebnisindikators von unter 10% betrachtet werden, ist davon auszugehen, dass Probleme der Datenvalidität bzw. eines zufälligen Fehlers (Dimick et al. 2004) ggf. eine größere Rolle spielen können als diejenigen die mit einer Risikoadjustierung adressiert werden (Heller 2008, Heller et al. 2011).

3) Eine Verbesserung der Qualitätsmessung sollte sich dem entsprechend auf alle Problembereiche der Qualitätsmessung beziehen: Ein homogenes Verständnis von Outcomes und Risikofaktoren und deren Kodierung sollte angestrebt werden. Im Rahmen der Entwicklung von Risikoadjustierungsmodellen könnten in Expertengruppen die Identifikation homogener Risikogruppen, z. B. in einem Fall-Kontroll-Ansatz im Rahmen der stattfindenden Peer Reviews konsentiert und weiter entwickelt werden. Dabei kann die Anzahl der verwendeten Risikofaktoren ggf. eingespart werden (Dimick et al. 2010). Inwieweit eine „Shrinkage“ der Ergebnisse in Richtung des Gesamtmittelwertes, wie in „empirical Bayes Modellen“ (Dimick et al. 2010), in allen Anwendungen eine sinnvolle statistische Lösung darstellt, kann kontrovers diskutiert werden (Nimptsch 2019). Ggf. wäre es stattdessen sinnvoll, diese Ergebnisse schlicht nicht zu berücksichtigen, oder aber unterschiedliche Outcomes zu einem Qualitätsindex oder einer Gesamtbewertung zusammen zu fassen (Heller 2008, Heller et al. 2011, IQTIG 2018, Günster & Jeschke 2019).

Referenzen:

- Becker A, Stausberg J, Fischer B, Carstanjen D, Weyermann M (2016)
Risikoadjustierung von Qualitätsindikatoren
Das Krankenhaus 11.2016: 954-963.
- Dimick JB, Welch HG, Birkmeyer JD (2004)
Surgical mortality as an indicator of hospital quality: the problem with small sample size.
JAMA; 292: 847-51.
- Dimick JB, Staiger DO, Birkmeyer JD (2010)
Ranking Hospitals on Surgical Mortality: The Importance of Reliability Adjustment.
Health Serv Res; 46: 1614-29.
- Dimick JB, Osborne NH, Hall BL, Ko CY, Birkmeyer JD (2010)
Risk adjustment for comparing hospital quality with surgery: how many variables are needed? Journal of the American College of Surgeons; 210 (4): 503-508.
- Günster C, Jeschke E (2019)
Risikoadjustierung von Qualitätsindikatoren mit Routinedaten
IQM Hearing „Qualitätsmessung in der Medizin - Risikoadjustierung und Risikoadjustierungsmodelle“, Berlin. 01.02.2019.
- Heller G, Schnell R (2007)
Hospital mortality risk adjustment using claims data.
Jama; 297: 1983 [Letter to the editor].
- Heller G (2008)
Zur Messung und Darstellung von medizinischer Ergebnisqualität mit administrativen Routinedaten in Deutschland.
Bundesgesundheitsbl - Gesundheitsforsch – Gesundheitsschutz; 51: 1173–1182.
- Heller G, Thomas T, Konheiser S, König T (2012)
Qualitätsindizes. In: Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen (AQUA), (Hrsg.). Qualitätsreport 2011. Göttingen: Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen: 212–213.
- Heller (2017)
Methoden der Risikoadjustierung. In Geraedts M, Drösler S, Döbler K, Eberlein-Gonska M, Heller G, Kuske S, Manser T, Sens B, Stausberg J, Schrappe M. DNVF-Memorandum III „Methoden für die Versorgungsforschung“, Teil 3: Methoden der Qualitäts- und Patientensicherheitsforschung. Gesundheitswesen; 79: e104-e106.
- Hosmer DW, Lemeshow S, Sturdivant RX (2013)
Applied Logistic Regression.
Wiley.
- IQTIG (2018)
Qualitätsindex zum kritischen Outcome bei Reifgeborenen.
In: Beschreibung der Qualitätsindikatoren für das Erfassungsjahr 2017 Geburtshilfe. Indikatoren 2017. Stand: 06.04.2018.
URL: https://iqtig.org/downloads/auswertung/2017/16n1gebh/QSKH_16n1-GEBH_2017_QIDB_V01_2018-04-06.pdf
- Nicholl J (2007)
Case-mix adjustment in non-randomised observational evaluations: the constant risk fallacy. Journal of epidemiology and community health; 61: 1010-1013.

Nimptsch U (2019)

Anforderungen an eine Risikoadjustierung zur Qualitätsmessung medizinischer Behandlungsergebnisse.
IQM Hearing „Qualitätsmessung in der Medizin - Risikoadjustierung und Risikoadjustierungsmodelle“,
Berlin. 01.02.2019



Das Deutsche Institut für Gefäßmedizinische Gesundheitsforschung gGmbH (DIGG) wurde von der Deutschen Gesellschaft für Gefäßchirurgie und Gefäßmedizin (DGG) 2009 gegründet, um Versorgungsforschung und Qualitätssicherungs-Programme in der Gefäßchirurgie zu fördern und zu realisieren.

Das DIGG ist eine 100%-ige Tochter der Deutschen Gesellschaft für Gefäßchirurgie und Gefäßmedizin e.V. und wird von einem ehrenamtlichen Geschäftsführer geleitet. Das oberste Entscheidungsgremium in wissenschaftlichen Fragen ist die sog. Projektsteuerungsgruppe (PSG). Die Amtszeit der PSG-Mitglieder beträgt 3 Jahre. Die PSG sichtet und bearbeitet Forschungsanträge und entscheidet nach Beratung mit dem Vorstand der DGG e.V. über die Art der Förderung. Sie kann in Eigeninitiative und Zusammenarbeit mit der Qualitätssicherungskommission und Forschungskommission der DGG e.V. eigene Studien und Forschungsvorhaben initiieren.

Inhalte, Ziele und Aufgaben des DIGG

- Qualitätssicherungsprogramme und Versorgungsforschung
- Förderung wissenschaftlicher Aufgaben in der Gefäßmedizin
- Erfassung epidemiologischer Daten arterieller und venöser Gefäßerkrankungen in Deutschland
- Erfassung ambulanter und stationärer Versorgungsstrukturen bei Gefäßerkrankungen
- Erfassung der Spezialisierung und Qualifikation von Ärzten/-innen im amb. und stat. Bereich
- Dokumentation und Bewertung von Qualitätsindikatoren der gefäßmedizinischen Versorgung

- Durchführung von Beobachtungsstudien zu gezielten Fragestellungen der vaskulären Versorgung
- Erfassung gesundheitsökonomischer Daten der ambulanten und stationären vaskulären Versorgung
- Forschung im Rahmen regionaler oder nationaler Screening-Programme
- Datenakquise für Publikationen und wissenschaftliche Vorträge.

Qualitätssicherung Bauchortenaneurysma (QS BAA)

- Wofür steht QS BAA?

Analyse der Ergebnisse der verschiedenen Therapieformen (EVAR vs. OS) für juxta- und infrarenale Bauchortenaneurysmen

- Wer kann teilnehmen?

Jede Klinik, die die Behandlung von Bauchortenaneurysmen durchführt, kann als Gefäßzentrum zertifiziert werden und damit Fälle eingeben.

- Wie kann ich Studienzentrum werden?

Homepage www.digg-dgg.de

Geschäftsstelle des DIGG /DGG (Tel.: 030-2800 4390)

E-Mail: sekretariat@gefaesschirurgie.de

- Datenmanagement

Teilnehmer geben die Daten über einen gesicherten webbasierten Zugang ein. Die Eingabe der Daten ist jederzeit möglich und kann auch in Teilabschnitten abgespeichert werden. Die Installation eines speziellen Programms auf dem Rechner des Nutzers ist nicht erforderlich. Die QS-Projekte werden von der BQS betreut. Die Versorgungsregister werden selbstinitiiert von Ärzten und Studynurses deutscher Universitätskliniken betreut.

Deutsches Institut für Gefäßmedizinische Gesundheitsforschung gGmbH Robert-Koch-Platz 9 10115 Berlin Tel: 030-280 990 99-0 Fax: 030-280 990 99-9 E-Mail: sekretariat@gefaesschirurgie.de

Der BAA- Risikoscore der DGG als Entscheidungshilfe zur Therapiewahl bei Versorgung des intakten Bauchaortenaneurysmas

R.T. Grundmann, M. Steffen, Th. Schmitz-Rixen

Zielsetzung: Zur Versorgung des intakten Bauchaortenaneurysmas (iAAA) stehen das endovaskuläre Vorgehen (EVAR) und die offene Versorgung (OR) zur Verfügung. Die Entscheidung, welches Vorgehen zu wählen ist, wird wesentlich durch die zu erwartende Operationsletalität beeinflusst. Ziel der vorliegenden Analyse war es, einen Risikoscore zu entwickeln, der die Operationsletalität in Abhängigkeit von der individuellen Risikokonstellation des Patienten für beide Vorgehensweisen vorhersagen lässt.

Methodik: Zur Modellerstellung wurden die Daten des BAA-Registers des Deutschen Instituts für Gefäßmedizinische Gesundheitsforschung (DIGG) der Deutschen Gesellschaft für Gefäßchirurgie und Gefäßmedizin (DGG) der Jahre 2013-2015 herangezogen [1-3]. Es handelte sich um 10.404 Patienten, die entweder mit EVAR (n= 7878) oder mit OR (n= 2534) wegen eines iAAA versorgt wurden. Für beide Subgruppen wurden alle präoperativ vorliegenden Risikovariablen auf eine signifikante Korrelation zur Letalität untersucht. Diese Items wurden in eine binäre logistische Regressionsanalyse eingesetzt und anschließend alle Variablen aus der Regressionsanalyse entfernt, bis nur noch Variablen inkludiert waren, die auch im Kontext der Regression ihren signifikanten Korrelationseffekt nachweisen ließen. Es folgte ein additives Risikovorhersagemodell für EVAR, welches aus den Modellfaktoren abgeleitet wurde und 7 Risikogruppen erkennen ließ. Die Risikospanne für die Operationsletalität lag hierbei zwischen 0,2% und 44,4%. Analoges geschah für OR und anschließend wurden beide Modelle fusioniert. Um die Vorhersagequalität des Gesamtmodells zu prüfen, wurde der Gesamtdatensatz einer ROC- Analyse unterzogen und die AUC bestimmt. Es fand sich eine AUC von 0,817 ($\pm 0,014$) mit hoher Signifikanz ($p < 0,001$). Das 95% Konfidenzintervall lag zwischen 0,789 und 0,844.

Ergebnisse: Der Risikoscore für EVAR umfasst: Alter > 85 Jahre, Geschlecht weiblich, juxta renale Pathologie, BAA-Durchmesser > 65 mm, Diabetes mellitus, ASA > 3 (alle je 2 Punkte), Kardiale Begleiterkrankung (3 Punkte) und GFR < 30 ml/min (5 Punkte). Für

OR gehen ein: Alter > 80 Jahre, Geschlecht weiblich, juxta- renale Pathologie , Z. n. Myokardinfarkt , Z. n. Apoplex (alle je 2 Punkte), ASA > 3 (3 Punkte), jedwede renale Begleiterkrankung (3 Punkte).

Folgerung: Mit dem DIGG-Vorhersagemodell lässt sich die Operationsletalität bei endovaskulärer und offener Versorgung des intakten Bauchaortenaneurysma für bestimmte Patientenkonstellationen verlässlich vorhersagen, wie eine Validierung des Datensatzes anhand von 3831 Fällen mit intakten AAA des Jahrgangs 2016 demonstrierte.

Literatur:

1. Grundmann RT (2015) Versorgung des abdominalen Aortenaneurysmas (AAA) 2014. Registerbericht des DIGG der DGG. Gefäßchirurgie 20: 376–384
2. Schmitz-Rixen T, Steffen M, Grundmann RT (2017) Versorgung des abdominalen Aortenaneurysmas (AAA) 2015. Registerbericht des DIGG der DGG. Gefäßchirurgie 22: 180–188
3. Schmitz-Rixen T, Steffen M, Grundmann RT (2018) Versorgung des abdominalen Aortenaneurysmas (AAA) 2016. Registerbericht des DIGG der DGG. Gefäßchirurgie [Epub ahead of print]

Korrespondenz:

Prof. Dr. med. Reinhart T. Grundmann, Deutsches Institut für Gefäßmedizinische Gesundheitsforschung (DIGG) der Deutschen Gesellschaft für Gefäßchirurgie und Gefäßmedizin, Berlin. E-Mail: grundmann@medsachverstand.de

M. Steffen, Klinikum Saarbrücken gGmbH, Winterberg 1, 66119 Saarbrücken. E-Mail: misteffen@klinikum-saarbruecken.de

Univ.-Prof. Dr. med. Thomas Schmitz-Rixen, Direktor der Klinik für Gefäß- und Endovascularchirurgie und des Universitären Wundzentrums, Klinikum der Goethe-Universität, Frankfurt/M. E-Mail: Schmitz-Rixen@em.uni-frankfurt.de

Risikoadjustierte Qualitätsvergleiche im Deutschen Qualitätsbündnis Sepsis auf Basis der §21-Daten

D. Schwarzkopf, D. O. Thomas-Rüddel, C. Fleischmann, H. Rüddel, K. Reinhart

Hintergrund: Sepsis ist die häufigste Ursache vermeidbarer Todesfälle im Krankenhaus. Auf Basis der §21-Daten können Krankenhäuser Indikatoren der Versorgungsqualität mit geringem Ressourceneinsatz ermitteln und untereinander vergleichen. Für die erheblichen Unterschiede im Case-Mix muss dabei jedoch adäquat adjustiert werden.

Entwicklung und Validierung eines Risikomodells: Ein Risikomodell für die Krankenhausletalität bei schwerer Sepsis oder septischem Schock wurde anhand der DRG-Statistik des DESTATIS entwickelt (N=134.000 in 2015) [1]. Kandidaten für Risikofaktoren waren u.a. Komorbiditäten gemäß Charlson und Elixhauser Indices, Merkmale der Infektion (z.B. Infektionsfokus, Vorliegen eines sept. Schocks), Aufnahmegrund sowie ausgewählte Prozeduren (z.B. Komplexbehandlung Schlaganfall). Ein Risikomodell auf Basis der Daten von 2013 zeigte bei Anwendung auf Daten in 2015 eine ausreichende bis gute Validität (AUC=0,74; R^2 -Schätzer=0,17; Letalität im niedrigsten und höchsten Risikodezil: 10%, 77%). Das Risikomodell wird jeweils auf Basis der aktuell jüngsten Daten beim DESTATIS aktualisiert.

Anwendung des Risikomodells im Deutschen Qualitätsbündnis Sepsis: Das entwickelte Risikomodell wird als externer Adjustierungsalgorithmus eingesetzt, um die risk-adjusted mortality rate (RSMR) für Sepsispatienten im Deutschen Qualitätsbündnis Sepsis zu ermitteln und zwischen Krankenhäusern zu benchmarken [2]. Krankenhäusern werden die Qualitätsindikatoren für verschiedene Subgruppen über interaktive grafische Darstellungen online verfügbar gemacht. Einzelfallanalysen und Peer-Reviews werden fokussiert auf verstorbene Patienten mit geringem vorhergesagtem Versterberisiko. Für die Veröffentlichung durch die teilnehmenden Häuser wird die RSMR zusätzlich mittels Schrumpfungsgewichten reliabilitätsadjustiert. In einer Online-Befragung bewerteten jeweils 80% der Vertreter der Häuser das Qualitätsreporting als nützlich und als gut verständlich.

Validität der Kodierung von Sepsis: Zur Prüfung der Validität der Sepsiskodierung in §21-Daten wurde ein nach Liegedauer und Vorliegen einer Komplexbehandlung Intensivmedizin stratifiziertes retrospektives Sample von 1000 Krankenakten des Universitätsklinikums Jena gezogen [3]. Durch Sichtung aller verfügbaren Akteninformationen wurden durch Ärzte Sepsisfälle identifiziert. Die explizite Kodierung von schwerer Sepsis in §21-Daten wurde gegen diesen Goldstandard validiert. Die Kodierung mittels der ICD-Codes R65.1 und R57.2 zeigte: Sens=25,1%, Spez=99,6%, PPV=56,1%, NPV=98,5%. Durch die zusätzliche Berücksichtigung mikrobiologischer Sepsiscodes in Kombination mit Organdysfunktionscodes konnte die Sensitivität auf 42% erhöht werden.

Diskussion: Das entwickelte Risikomodell zeigte gute Validitätskriterien die mit US-amerikanischen Modellen auf ähnlicher Datenbasis vergleichbar sind [4]. Risikoadjustierte Qualitätsindikatoren werden von Praktikern als nützlich für das Qualitätsmanagement bewertet. Zugleich zeigen sich erhebliche Einschränkungen der Validität der Kodierung von Sepsis in §21-Daten.

Ausblick: Die etablierte Methodik der Risikoadjustierung ist prinzipiell auf andere Indikatoren übertragbar. Die Untersuchung der klinischen Validität der §21-Daten offenbart jedoch Defizite. Die

Qualität der Kodierung von Risikofaktoren sowie mögliche systematische Unterschiede der Validität der Kodierung zwischen Krankenhäusern müssen untersucht werden.

Literatur:

1. Schwarzkopf D, Fleischmann-Struzek C, Rüdgel H, Reinhart K, Thomas-Rüdgel DO. A risk-model for hospital mortality among patients with severe sepsis or septic shock based on German national administrative claims data. PLOS ONE. 2018;13(3):e0194371. doi: 10.1371/journal.pone.0194371.
2. Schwarzkopf D, Rüdgel H, Gründling M, Putensen C, Reinhart K. The German Quality Network Sepsis: study protocol for the evaluation of a quality collaborative on decreasing sepsis-related mortality in a quasi-experimental difference-in-differences design. Implement Sci. 2018;13(1):15. doi: 10.1186/s13012-017-0706-5.
3. Fleischmann-Struzek C, Thomas-Rüdgel DO, Schettler A, Schwarzkopf D, Stacke A, Seymour CW, et al. Comparing the validity of different ICD coding abstraction strategies for sepsis case identification in German claims data. PLOS ONE. 2018;13(7):e0198847. doi: 10.1371/journal.pone.0198847.
4. Ford DW, Goodwin AJ, Simpson AN, Johnson E, Nadig N, Simpson KN. A Severe Sepsis Mortality Prediction Model and Score for Use With Administrative Data. Crit Care Med. 2016;44(2):319-27. doi: 10.1097/ccm.0000000000001392. PubMed PMID: 00003246-201602000-00010.

Kontakt

Dr. phil. Daniel Schwarzkopf, Center for Sepsis Control and Care, Universitätsklinikum Jena, Am Klinikum 1, 07747 Jena

Daniel.Schwarzkopf@med.uni-jena.de

Tel: +49 3641 9 323195

Prof. Dr. Birgit Seelbach-Göbel

Klinik für Geburtshilfe und Frauenheilkunde der Universität Regensburg – St. Hedwig

Es ist kein Zufall, dass die ersten planungsrelevanten Qualitätsindikatoren Ende 2016 für den Fachbereich der Frauenheilkunde definiert und eingeführt wurden. Blickt doch gerade die Geburtshilfe auf eine über 40 Jährige Tradition in der Qualitätssicherung zurück, als die Münchner Perinatalstudie in den siebziger Jahren des 20. Jahrhunderts begann, die Daten der Geburten in Krankenhäusern zu sammeln. Uns stehen heute eine Vielzahl an Kennzahlen aus der externen Qualitätssicherung zur Verfügung, die einen Vergleich zwischen verschiedenen Krankenhäusern ermöglichen könnten, wenn das behandelte Patientenkollektiv denn vergleichbar wäre. Zum Beispiel rechtfertigten einerseits große Perinatalzentren hohe Kaiserschnittraten mit dem Argument, ihr großer Frühgeburtenanteil mache dies erforderlich. Andererseits haben gerade kleine Abteilungen die höchsten Kaiserschnittraten, aber dennoch auch die höchsten Azidoseraten bei reifgeborenen Kindern. Im Indikatorenkatalog des IQM spielt das perinatale Outcome keine Rolle, obwohl gerade dieses die Ergebnisqualität widerspiegelt. Die meisten gelisteten Indikatoren (GIQI) sind nicht durch Referenzwerte definiert sondern geben lediglich (Mengen-) Informationen. Der externen Qualitätssicherung ist es durch Einführung der Kennzahl „beobachtete zur erwarteten Rate“ an beispielsweise Kaiserschnitten, Azidosen, höhergradigen Dammrissen teilweise gelungen, Unterschiede im Risikoprofil der einzelnen Krankenhäuser zu berücksichtigen und eine Gegenüberstellung zu ermöglichen. Im internationalen Vergleich hat sich die Robson- Klassifikation für Kaiserschnitte zur besseren Differenzierung nach Risikogruppen bewährt. Verschiedentlich hat sie auch in Deutschland in den EQS der Länder Eingang gefunden. Bedauerlicherweise gehört eine niedrige Kaiserschnitttrate bisher nicht zu den planerischen Qualitätsindikatoren, obwohl die Folgen einer großzügigen Indikationsstellung immer häufiger in vermehrten Komplikationen im Zustand nach Kaiserschnitt evident werden. In der Gynäkologie weisen die GIQI lediglich Zahlen zu Prozeduren auf, die Ergebnisqualität wird nicht erfasst. Um Ergebnisqualität abzubilden ist daher die Verknüpfung mit Daten aus der externen Qualitätssicherung unerlässlich.

¹ Die Präsentation des Beitrages beim IQM Hearing am 01.02.2019 musste aus Termingründen leider entfallen.