

Prof. Dr. Andreas Becker/Prof. Dr. Jürgen Stausberg/Burkhard Fischer/Dr. Dirk Carstanjen/  
Prof. Dr. Maria Weyermann

# Risikoadjustierung von Qualitätsindikatoren

Eine Positionsbestimmung der DKG-Expertengruppe „Qualitätsmessung und Risikoadjustierung“

*Qualität ist ein zentraler Topos des Krankenhausstrukturgesetzes. Schlechte Qualität in der Krankenhausversorgung soll künftig mit Vergütungsabschlägen bestraft werden. Qualitätsindikatoren können künftig in der Krankenhausplanung der Länder eine Rolle spielen, Krankenhäuser, die die notwendige Qualität dauerhaft nicht bieten können, sollen darüber hinaus aus dem Krankenhausplan herausgenommen werden können.*

*Ein wichtiger Ausdruck einer Qualitätskultur aber ist ein fairer Vergleich. Dies soll die Risikoadjustierung gewährleisten: indem eine unterschiedliche Verteilung von Störgrößen in den Krankenhäusern ausgeglichen wird. Die Autoren erläutern wichtige Einflussgrößen und skizzieren Kriterien, die ein Verfahren zur Risikoadjustierung im Qualitätsvergleich von Kliniken berücksichtigen sollte.*

## Qualität

Qualität, hier medizinische Behandlungsqualität, ist kein empirisch nachweisbarer Sachverhalt, sondern ein Konstrukt: „Qualität ist der Grad, in dem ein Satz inhärenter Merkmale eines Objekts Anforderungen erfüllt.“<sup>1)</sup> Für Qualitätsbewertungen müssen Anforderungen an die Qualität vorab definiert werden. Die Erfüllung dieser Anforderungen wird in aller Regel numerisch mit Kennzahlen bestimmt. Qualitätsindikatoren sind solche Kennzahlen. In die Berechnung dieser Kennzahlen fließen Ereignisse wie das Versterben eines Patienten oder Messwerte, wie die Dauer einer Beatmung, ein. Die Bestimmung der Qualität ist immer eine Schätzung.

Zum einen werden die Ergebnisse durch den Zufall beeinflusst. So können insbesondere seltene Ereignisse wie eine im Krankenhaus erworbene Sepsis in einem Jahr bei einem Krankenhaus zufallsbedingt auftreten, bei einem anderen Krankenhaus jedoch nicht. Der Einfluss des Zufalls lässt sich durch statistische Verfahren wie die Verwendung von Vertrauensbereichen kontrollieren. Allerdings nimmt die Verwendung dieser statistischen Verfahren selbst systematischen Einfluss auf eine spätere Qualitätsbewertung, zum Beispiel indem die Chance zur Auszeichnung mit besonders guter Qualität von der Leistungsmenge abhängig wird.

Zum anderen sind die Faktoren, die beispielsweise die Häufigkeit des Versterbens im Krankenhaus bei akutem Myokardinfarkt beeinflussen, nicht alle bekannt. Auch sind derartige Ereignisse nicht alle messbar; sie sind vielfältig und in komplexer Beziehung miteinander verwoben. Die Ergebnisse von Qualitätsindikatoren unterliegen daher systematischen Unsicherheiten, unabhängig von der Qualität der Leistungserbringung.

## Qualitätsindikatoren

Im Gesundheitswesen wird aufgrund der beschriebenen Beschränkungen aus methodisch-wissenschaftlicher Perspektive die Anwendung von Qualitätsindikatoren in erster Linie in der Unterstützung des internen Qualitätsmanagements gesehen. Das Einsatzgebiet von Qualitätsindikatoren wurde durch Fachgesellschaften wie folgt definiert.<sup>2)</sup>

„Ein Indikator ist ein quantitatives Maß, welches zum Monitoring und zur Bewertung der Qualität wichtiger Leitungs-, Management-, klinischer und unterstützender Funktionen genutzt werden kann, die sich auf das Behandlungsergebnis beim Patienten auswirken. Ein Indikator ist kein direktes Maß der Qualität. Er ist mehr ein Werkzeug, das zur Leistungsbeurteilung benutzt werden kann, das Aufmerksamkeit auf potenzielle Problembereiche lenken kann, die einer intensiven Überprüfung innerhalb einer Organisation bedürfen könnten.“

Diese Einschätzung wurde bereits von der Joint Commission on Accreditation of Healthcare Organizations (JCAHO) im Jahre 1990 formuliert und **festgestellt**: „Indicators are not direct measures of quality“.<sup>3)</sup> Ob die Verwendung von Qualitätsindikatoren jenseits des etablierten Einsatzgebietes möglich und nützlich ist, ist national und international Gegenstand des wissenschaftlichen Diskurses<sup>4), 5), 6), 7), 8), 9), 10)</sup>. Dies betrifft u. a. deren Verwendung in der Krankenhausplanung, bei einer qualitätsabhängigen Vergütung (Pay-for-Performance) oder zur öffentlichen Information, zum Beispiel über Qualitätsberichte oder Patientenportale. Die Auswirkungen einer unzutreffenden Qualitätsbewertung können dann dramatisch sein: ein Ausschluss guter Krankenhäuser aus der Versorgung, die nicht kostendeckende Refinanzierung der Leistungserbringung durch Abschläge oder die irrtümliche Entscheidung

eines Patienten für ein Krankenhaus entgegen seiner Präferenzen. Das Risiko einer unzutreffenden Bewertung muss daher auf ein zu definierendes Maß reduziert und die Anforderungen an eine faire Qualitätsbewertung explizit definiert werden. Angelehnt an Donabedian<sup>11)</sup> ist zu fordern, dass die Beziehung zwischen dem Ergebnis eines Qualitätsindikators und der tatsächlichen Qualität beschreibbar ist. Das bedeutet, dass absichtsvolles Handeln, das auf eine Verbesserung der Versorgungsqualität zielt und durch den Qualitätsindikator abgebildet wird, auch eine tatsächliche und messbare Verbesserung der Behandlungsqualität erreicht. Oder umgekehrt: Eine Vernachlässigung der Ziele des Qualitätsindikators führt zu einer nachweisbaren Verschlechterung der Qualitätsergebnisse. Fairness ist in diesem Sinne die zentrale Eigenschaft eines Indikators, die sich aus dieser Beziehung ableitet. Fair ist ein Vergleich von Krankenhäusern über Qualitätsindikatoren, wenn die ermittelten Unterschiede, beispielsweise einer Sterblichkeit im Krankenhaus bei akutem Myokardinfarkt, weitgehend durch die Krankenhäuser beeinflusst werden können und zudem die fehlende Beeinflussung in die gewünschte Richtung ein Versäumnis darstellt.

### **Einflussgrößen auf das Ergebnis eines Qualitätsindikators**

Die mit den Qualitätsindikatoren gemessenen Sachverhalte werden durch eine große Vielfalt von Faktoren beeinflusst:

Zu den Einflussgrößen für das interessierende Ereignis gehören unbekannte Einflussgrößen, die demzufolge auch nicht messbar sind.

Bekannte Einflussgrößen können durch das Krankenhaus nicht beeinflussbar, aber messbar sein (zum Beispiel das Alter der Patienten, das Stadium der Erkrankung, die regionale Lage des Krankenhauses), oder aber beeinflussbar, aber nicht messbar (zum Beispiel Umweltrisiken, „schicksalhafter“ Verlauf der Erkrankung, Qualität der ambulanten Versorgung).

Durch das Krankenhaus beeinflussbar und messbar wäre etwa das Vergessen eines Fremdkörpers im Körper des Patienten bei Operation oder eine falsche Wahl des therapeutischen Vorgehens. Beeinflussbar, aber nicht eindeutig messbar ist beispielsweise die Motivation der Beschäftigten.

In eine faire Qualitätsbewertung gehen ausschließlich bekannte Einflussgrößen und Unterschiede ein, die vom Krankenhaus beeinflussbar sind. Die Qualitätsbewertung wird jedoch durch Einflussgrößen verzerrt, die durch das Krankenhaus nicht zu beeinflussen sind oder bei denen eine Beeinflussung nicht erwartet werden kann – wie etwa das regionale sozioökonomische Umfeld der zu versorgenden Patienten. Überschreitet die Verzerrung ein bestimmtes Ausmaß, wird die Bewertung unfair. Aber auch bei einem bestmöglichen Ausgleich für andere Faktoren verbleibt eine Unsicherheit. Diese Unsicherheit findet ihren Ausdruck in vom Zufall bestimmten Schwankungen der Messwerte. Mittels statistischer Methoden kann der Einfluss des Zufalls korrigiert werden. Hierbei spielt allerdings die Verteilung der Fallzahlen auf die einzelnen Auswertungseinheiten eine entscheidende Rolle



Hohe Bildqualität.  
Mobilität.  
Konnektivität.

### **SonoSite iViz – Ihr persönliches medizinisches Ultraschallgerät**

Scannen, Zugriff auf Lernmittel und Patientendaten, Untersuchungen speichern, Berichte übermitteln, Austausch mit Remote-Providern für unmittelbare Beurteilungen und mehr...

Erfahren Sie mehr unter  
**[www.sonosite.com/de/sonosite-iviz](http://www.sonosite.com/de/sonosite-iviz)**  
oder kontaktieren Sie SonoSite unter  
**069 / 80 88 40 30** oder  
**[eraf-sales@sonosite.com](mailto:eraf-sales@sonosite.com)**  
für mehr Informationen.

**P21v**  
5-1 MHz  
sektorförmig



**L38v**  
10-5 MHz linear

**Tabelle 1: Qualify – Gütekriterien und Kategorien**

| Kategorie   | Methodisches Gütekriterium   |
|---|--|
| Relevanz  | Bedeutung des mit dem Qualitätsindikator erfassten Qualitätsmerkmals für das Versorgungssystem |
|   | Nutzen   |
|   | Berücksichtigung potenzieller Risiken / Nebenwirkungen   |
| Wissenschaftlichkeit                                | Indikatorevidenz   |
|   | Klarheit der Definitionen (des Indikators und seiner Anwendung)                                |
|   | Reliabilität   |
|   | Statistische Unterscheidungsfähigkeit  |
|   | Risikoadjustierung   |
|   | Sensitivität   |
|   | Spezifität   |
| Praktikabilität                                     | Validität  |
|   | Verständlichkeit und Interpretierbarkeit für Patienten und interessierte Öffentlichkeit        |
|   | Verständlichkeit für Ärzte und Pflegende   |
|   | Beeinflussbarkeit der Indikatorausprägung  |
|   | Datenverfügbarkeit   |
|   | Erhebungsaufwand   |
|   | Implementationsbarrieren berücksichtigt  |
|   | Die Richtigkeit der Daten kann überprüft werden  |
|   | Die Vollständigkeit der Daten kann überprüft werden  |
| Die Vollständigkeit der Daten kann überprüft werden |  |

„Eignung für die einrichtungsbezogene öffentliche Berichterstattung“ beschreibt und als Kriterium fordert, das Ergebnis des Indikators sollte durch einen einzelnen Leistungserbringer beeinflussbar sein.<sup>12)</sup> Die Prüfung erfolgte jedoch nicht quantitativ, sondern durch Befragung von Experten mit einer modifizierten RAND Appropriateness Method (mRAM), wobei diese eine Bewertung in ganzzahligen Stufen zwischen 1 (überhaupt nicht geeignet) und 9 (sehr geeignet) vorsah. Angesichts des komplexen Zusammenwirkens verschiedener Einflussgrößen ist diese Form der Bewertung nicht zufriedenstellend. Das Institut für Qualität und Patientensicherheit BQS hatte ein sehr feines Instrument mit 20 methodischen Gütekriterien unter dem Namen Qualify veröffentlicht.<sup>13)</sup> Die Fairness verteilt sich hier auf verschiedene Kriterien wie eine statistische Unterscheidungsfähigkeit, die Validität und die Beeinflussbarkeit der Indikatorausprägung. Ein quantitatives Maß fehlt allerdings ebenfalls. Hingegen finden sich einige Vorschläge zur Quantifizierung

und führt dazu, dass bei kleinen Fallzahlen gegebenenfalls keine Qualitätsbewertungen mit angemessener Sicherheit vorgenommen werden können.

### Fairness

Für jede einzelne Kennzahl, die als Indikator für eine Qualitätsbewertung in Betracht kommt, ist daher deren Grad an Fairness zu bestimmen. Falls ein festzulegendes Maß an Fairness nicht erreicht wird, kann über eine Risikoadjustierung eine Reduktion der Verzerrung der Ergebnisse versucht werden. Hierbei sollten die Ergebnisse für bekannte, nicht durch das Krankenhaus zu verantwortende Unterschiede ausgeglichen werden. Diese Faktoren lassen sich auch als Störgrößen bezeichnen. Auch hier bleibt selbst bei einer optimalen Risikoadjustierung eine Unsicherheit aufgrund unbekannter und nicht messbarer Größen zurück. Tatsächlich kann eine Risikoadjustierung nur die messbaren und nicht durch das Krankenhaus beeinflussbaren Störgrößen berücksichtigen. Der Nachweis einer sachgerechten Risikoadjustierung ist bei den auch dann verbleibenden Unsicherheiten für die Einschätzung der Fairness eines Indikators damit bei weitem nicht ausreichend. Die Fairness einer Qualitätsbewertung selbst muss in der Güteprüfung eines Indikators nachgewiesen werden.

In den in Deutschland propagierten Verfahren zur Güteprüfung von Qualitätsindikatoren fehlt ein quantitativer Parameter, der das Ausmaß der Fairness erfasst. Das AQUA-Institut nähert sich diesem Aspekt an, wenn es eine Prüfung auf

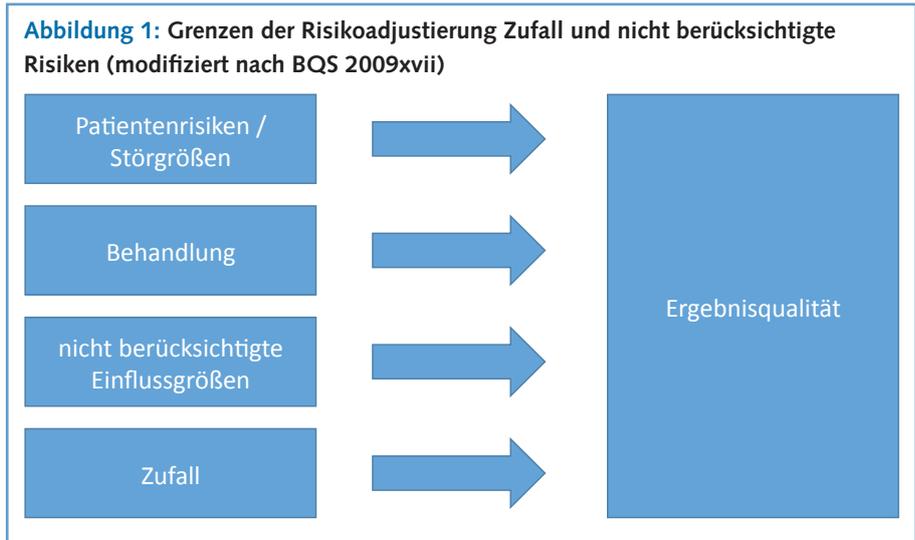
die Fairness eines Indikators bei den wissenschaftlichen Vorarbeiten zu den Patient Safety Indicators (PSI) der Agency for Healthcare Research and Quality (AHRQ)<sup>14)</sup>. Unter anderem wird als Maß die sogenannte „Signal Share“ eingeführt: „Signal share is a measure of how much of the total variation (patient and provider) is potentially subject to hospital control.“ Zusätzlich wird auf die Beschränktheit einer Risikoadjustierung und auf die Bedeutung von Zufallseffekten und unbekanntem Störgrößen hingewiesen.<sup>15)</sup>

### Gütekriterien für Qualitätsindikatoren

Qualitätsindikatoren sind Steuerungsinstrumente mit möglicherweise gravierenden Konsequenzen für Leistungserbringer und für die Gesundheitsversorgung insgesamt. Um Fehlsteuerungen zu verhindern, sollen Indikatoren selbst bestimmte Qualitätsanforderungen erfüllen. Mit Qualify hat das BQS-Institut 2007 hierfür ein durchgängig operationalisiertes Instrument zur methodischen Bewertung von Indikatoren veröffentlicht. Qualify wurde mit wissenschaftlicher Begleitung auf Basis einer systematischen Literaturrecherche entwickelt. Es umfasst 20 methodische Gütekriterien, die den drei Kategorien Relevanz, Wissenschaftlichkeit und Praktikabilität zugeordnet sind (► **Tabelle 1**). Es wurde vor seiner Veröffentlichung an 55 Indikatoren erprobt und diente in seiner ersten Anwendung als methodische Grundlage für die Bestimmung der Qualitätsindikatoren, die von Krankenhäusern im Qualitätsbericht zu veröffentlichen sind. Eine neuere Arbeit<sup>16)</sup> bestätigt im We-

sentlichen die in Qualify enthaltenen Kriterien aber gleichzeitig, wie breit gefächert die methodische Diskussion weiterhin ist – und dass sich noch kein methodischer Standard durchgesetzt hat.

In der Kategorie Wissenschaftlichkeit enthält Qualify auch das Gütekriterium „Risikoadjustierung“. Die Autoren von Qualify sehen es als besonders bedeutsam für die öffentliche Berichterstattung an. Von ihm hänge ab, ob die Ergebnisse eines Qualitätsindikators tatsächlich der Behandlungsqualität zugerechnet werden können. Bei Nichtberücksichtigung seien Indikatorwerte eher Ausdruck des Casemix des bewerteten Leistungserbringers als der Behandlungsqualität. Hierzu gelte es, die Einflussfaktoren, welche zumeist Patientenmerkmale wie Krankheitsschwere und Komorbidität darstellten, angemessen zu adjustieren. Aber auch Patientenpräferenzen seien unter Umständen zu berücksichtigen. Die Risikoadjustierung soll helfen, die richtigen Unterschiede bei Qualitätsvergleichen von Krankenhäusern sichtbar zu machen: Unterschiede, die auf den Einfluss des Krankenhauses zurückgehen und bei denen eine Beeinflussung des Krankenhauses erwartet werden kann.



Umrechnung auf eine Referenzverteilung der als relevant eingeschätzten Störgrößen, zum Beispiel Alter und Geschlecht. In den Qualitätsvergleich gehen dann nicht die gemessenen, sondern die transformierten Ergebnisse ein. Bei der indirekten Standardisierung wird die Häufigkeit des interessierenden Ereignisses in einer Referenzpopulation zur Ableitung einer erwarteten Häufigkeit in einem Krankenhaus verwendet und der beobachteten Häufigkeit in diesem Krankenhaus gegenübergestellt. Das Ergebnis wird als Quotient, als Verhältniszahl (Ra-

## Risikoadjustierung

Aufgabe einer Risikoadjustierung ist es, die Fairness zu steigern, indem eine unterschiedliche Verteilung von Störgrößen in den Krankenhäusern ausgeglichen wird. Dies kann methodisch auf verschiedenen Wegen erreicht werden (siehe auch Erläuterungen in den Qualitätsberichten der BQS, zum Beispiel Veit et. al.)<sup>17)</sup>

Die spezifische Festlegung von Einschlusskriterien versucht, patientenseitige Unterschiede auszugleichen. Bei ratenbasierten Indikatoren beeinflusst dies die Population des Nenners, für die dann die Patienten mit einem interessierenden Ereignis im Zähler ausgewiesen werden. Ähnlich verhält es sich bei der Auftrennung der Population nach dem Schweregrad der Erkrankung (beispielsweise ASA-Klassifikation, Tumorstadien), dem Alter, dem Geschlecht oder nach anderen Störgrößen und dem Ausweis getrennter Ergebnisse (Stratifizierung). Für diese Stratifizierung können auch komplexere Modelle herangezogen werden, über die zum Beispiel ein Scorewert berechnet wird. Auch können weitere Störgrößen mitbetrachtet werden, wenn etwa Qualitätsvergleiche regional differenziert vorgenommen werden. Differenzierung und Stratifizierung können auch zur Risikoadjustierung bei Messungen, etwa von Zeiten und Mengen, eingesetzt werden (siehe ► **Abbildung 1**).

Bei einer Standardisierung werden die berechneten Ergebnisse des einzelnen Krankenhauses für den Qualitätsvergleich transformiert. Bei der direkten Standardisierung erfolgt eine

Gesundheitsmanagement ist kein Zufall!

**Medizin-Controlling**

- Kodierung & Abrechnung
- MDK-Fallbearbeitung
- Erlösoptimierung

**Krankenhausberatung**

**Qualitätsorientierte Dokumentation**

elbamed GmbH: Tel.: 0800-elbamed (0800-35 22 6 33)  
www.elbamed.de ■ info@elbamed.de

tio), ausgewiesen. Bei dieser Ratio ist im Idealfall auch eine unmittelbare Qualitätsbewertung möglich, da das Ergebnis „1“ der „durchschnittlichen“ Qualität entspricht. In einer besonderen Form der indirekten Standardisierung werden nicht Subgruppen von Patienten betrachtet, sondern über ein Risikomodelle wird eine Eintrittswahrscheinlichkeit des interessierenden Ereignisses für jeden Patienten berechnet. Die Summe der Eintrittswahrscheinlichkeiten ergibt dann die erwartete Häufigkeit, zum Beispiel des Versterbens von Patienten bei akutem Myokardinfarkt, die der beobachteten Häufigkeit gegenübergestellt wird, etwa als Standardized Mortality Ratio (SMR). Bei der Standardisierung können nur bekannte und messbare Störgrößen Berücksichtigung finden.

In der Diskussion einer Risikoadjustierung gilt: Die Verwendung standardisierter Verhältniszahlen ist keine Garantie für einen fairen Vergleich. Über die üblichen Gütekriterien hinaus geht es bei der Risikoadjustierung nicht um die bestmögliche Erklärung eines Ergebnisses. Die Risikoadjustierung soll helfen, die relevanten Unterschiede bei Qualitätsvergleichen von Krankenhäusern sichtbar zu machen, Unterschiede, die auf den Einfluss des Krankenhauses zurückgehen und bei denen eine Beeinflussung des Krankenhauses erwartet werden kann.

Bis hierher kann daher festgestellt werden: Ergebnisqualität wird von verschiedenen Größen beeinflusst. Diese Einflussgrößen müssen bei der Qualitätsmessung und Qualitätsberichterstattung, bei deren Nutzung als planungsrelevante Qualitätsindikatoren oder in einem qualitätsabhängigen (P4P-) Vergütungssystem, berücksichtigt werden. Da die Patientenstruktur in verschiedenen Krankenhäusern sehr unterschiedlich sein kann, ist für einen fairen Klinikvergleich eine risikoadjustierte Betrachtungsweise unabdingbar.

## Gütekriterien zur Beurteilung eines Verfahrens zur Risikoadjustierung

Eine vollständige Risikoadjustierung nach allen möglichen Einflussgrößen ist nicht durchführbar. Daher ermöglicht auch die Risikoadjustierung keinen Vergleich von Einrichtungen im Sinne einer „exakten wissenschaftlichen Methodik“. Auch hochwertige Risikoadjustierungsmodelle können nicht alle

Einflussgrößen berücksichtigen. Dies gilt insbesondere für solche, die auf administrativen Routinedaten basieren.

Selbst hochwertige Risikomodelle ermöglichen nur eine mehr oder weniger genaue Schätzung der Zielvariablen. Risikomodelle sind daher nicht als ein Instrument zur Diagnostik der Versorgungsqualität zu verstehen, sie dienen vielmehr als Screeningverfahren, um die Aufmerksamkeit auf Bereiche zu lenken, die ggf. einer genaueren fachlichen Analyse bedürfen. Sie stellen ggf. statistische Auffälligkeiten dar, die näher untersucht werden sollen.

Keine Methode zur Risikoadjustierung ist perfekt und es wird niemals möglich sein, eine Adjustierung für alle Patientencharakteristika durchzuführen. In Anbetracht dieser grundsätzlichen Limitierung ist es umso wichtiger, das verbleibende Potenzial der Risikoadjustierung nicht durch Fehler bei der Modellplanung und Modellentwicklung weiter einzuschränken.

Auf Basis dieser Überlegungen sind für die Modellentwicklung und die anzuwendenden statistischen Verfahren – insbesondere die multiple logistische Regression – besondere Punkte zu berücksichtigen. Die Entwicklung der Modelle muss nachvollziehbar und transparent erfolgen. Die entwickelten Modelle sind zu validieren und anhand vorab zu definierender Kriterien zu bewerten.<sup>18)</sup>

## Die Kriterien im Einzelnen

Ein Verfahren zur Risikoadjustierung im Qualitätsvergleich von Kliniken sollte die folgenden Kriterien berücksichtigen:

### Transparenz

- Akzeptanz und Motivation zur Anwendung eines Risikoadjustierungsmodells und seiner Ergebnisse hängen in besonderem Maße von der Transparenz des Modells ab.
- Es wird empfohlen, zu jedem Modell Transparenz hinsichtlich der in dieser Schrift formulierten Anforderungen und Empfehlungen herzustellen.

### Modellentwicklung

- Bei der Entwicklung eines Risikoadjustierungsmodells sollen in der Literatur formulierte Konzepte und Gütekriterien berücksichtigt, Art und Umfang nachvollziehbar beschrieben werden – auch, wenn etablierte Konzepte und Gütekriterien nicht berücksichtigt werden.
- Die Spezifizierung eines Risikoadjustierungsmodells soll vorhandenes Wissen zur Fragestellung (Literatur, Leitlinien etc.) sowie klinische und statistische Expertise berücksichtigen.
- Schon vor einer ersten Datenauswertung sollen klinische Hypothesen zu Einflussvariablen inklusive ihrer Ausprägung und möglichen Wechselwirkungen sowie der möglichen unterschiedlichen Wirkungen in bestimmten Subgruppen (zum Beispiel Geschlecht, Alter) formuliert werden.

Anzeige



**RECHTSANWÄLTE WIGGE**  
BERATUNG IM MEDIZINRECHT

**Krankenhausplanungs- und -finanzierungsrecht**  
**Sektorenübergreifende Kooperationen**  
**Arbeits- und Personalrecht**

|  |   |  |
|--|---|--|
| <p><b>Beratungskompetenz für Krankenhäuser</b><br/>www.ra-wigge.de</p> | <p><b>MÜNSTER</b><br/>Scharnhorststraße 40<br/>48151 Münster<br/>TEL (0251) 53 59 5 - 0</p> | <p><b>HAMBURG</b><br/>Neuer Wall 44<br/>20354 Hamburg<br/>TEL (040) 339 87 05 - 90</p> |
|--|---|--|

- Wenn ausgewählte Einflussvariablen auf jeden Fall in einem Risikoadjustierungsmodell enthalten sein sollen, so soll die zugrunde liegende Rationale detailliert beschrieben werden.
- Das Outcome, auf welches sich ein Risikoadjustierungsmodell beziehen soll, soll exakt definiert werden und es soll genau geprüft werden, ob und in welchem Umfang es mittels Routinedaten sicher abgebildet werden kann.
- Bei der Modellentwicklung soll geprüft werden, ob die zu untersuchende Patientengruppe eindeutig definiert werden kann. Hierbei ist auch zu berücksichtigen, dass relevante Faktoren nur eingeschränkt oder nicht abgebildet werden können (zum Beispiel Krankheitsschwere, Status Reanimation ja/nein, Therapiebeschränkung).
- Systematische Fehler in den Daten können die Validität von Risikoadjustierungsmodellen erheblich verfälschen. Diese Fehler können zum Beispiel auf Unterschieden in der Beobachtung bzw. Kodierung beruhen und sich von Einrichtung zu Einrichtung stark unterscheiden. Bei einem Risikoadjustierungsmodell soll auch thematisiert werden, ob die zugrunde liegenden Risikoadjustierungsvariablen valide erhoben wurden. Ansonsten kann das Risikoadjustierungsmodell durchaus mit einer Verschlechterung der Risikoadjustierung einhergehen. Insofern muss eine sinnvolle Weiterentwicklung der Risikoadjustierung auch bedeuten, dafür Sorge zu tragen, dass Risikoadjustierungsvariablen einheitlich verstanden und dokumentiert sind.
- Die Datenqualität und auch die Häufigkeiten (absolut und relativ) der Ziel- und Einflussvariablen in den Entwicklungs- und Validierungsdaten sollen detailliert beschrieben werden.
- Zur Abbildung bestimmter Komorbiditäten und Komplikationen (zum Beispiel Sepsis) sind komplexe Definitionen erforderlich. Diese Definitionen und auch die Rationale für ihre Auswahl sollen detailliert beschrieben werden.
- Zu diesen Limitierungen gehört auch das Fehlen eines Kennzeichens, welches eine eindeutige Zuordnung von Komorbiditäten und Komplikationen im Sinne von „bei der Aufnahme vorhanden ja/nein“ ermöglicht. Aus diesem Grunde wird die Prüfung der Einführung eines solchen Kennzeichens (Present-On-Admission [POA]) im Rahmen entsprechender Forschungsprojekte empfohlen. Weitere Limitierungen sind fehlende ICD-Kodes zur eindeutigen Abbildung definierter Komplikationen sowie des „Do-Not-Resuscitate-(DNR)“-Status.
- Weiterhin sollen die möglichen Auswirkungen von Definitionen und Vorgaben aus den Deutschen Kodierrichtlinien und den Klassifikationen (ICD und OPS) auf die Datenqualität untersucht und detailliert beschrieben werden.
- Im Rahmen der Modellentwicklung sollte definiert werden, wie mit der Zuordnung des Zielereignisses (zum Beispiel Tod) im Falle einer Verlegung zum verlegenden oder aufnehmenden Krankenhaus umgegangen wird. Oder mit der Frage einer plausiblen bzw. erforderlichen Mindestverweildauer bei bestimmten Qualitätsindikatoren.
- Bei der Modellentwicklung soll paradoxen Effekten<sup>19)</sup> (auch „paradoxe Variablen“ genannt) besondere Aufmerksamkeit gewidmet werden, da sie Hinweise auf bedeutsame Probleme der Datenqualität (Kodierung, Probleme bei der Abgrenzung von Komorbiditäten und Komplikationen), Interaktionen oder Multikollinearität sein können. Die möglichen Ursachen und der Umgang mit solchen Einflussvariablen im Rahmen der Modellentwicklung sollen detailliert beschrieben werden. Verbleiben paradoxe Effekte im Risikoadjustierungsmodell, so sollen die möglichen (patho-)physiologischen Grundlagen oder andere Erklärungen des Effektes detailliert beschrieben werden.
- Bei der Modellentwicklung soll der Erfassungszeitpunkt oder Erfassungszeitraum, auf den sich das Risikoadjustie-

## Recruiting ist Vertrauenssache

**HealthCare**  
Personalmanagement GmbH

because we care

Als Experten im Gesundheitswesen bringen wir die richtigen Partner zusammen. Wir sind darauf spezialisiert, diskret und effektiv die passenden neuen Führungskräfte für Kliniken, Pharmakonzerne und Praxen zu gewinnen - seit über 12 Jahren mit großem Engagement und Erfolg. Damit Sie und Ihr Team wieder voll durchstarten können.



HealthCare Personalmanagement GmbH  
Geschäftsführung: Dorothea Rickert  
Merowingerplatz 1, 40225 Düsseldorf, Telefon: 0211 220 58 90  
info@healthcare-personal.de, www.healthcare-personal.de

Dorothea Rickert  
Geschäftsführerin

rungsmodell beziehen soll (zum Beispiel Krankenhausausschuss, 30 oder 90 Tage danach), genau beschrieben werden. Die zugrunde liegende Rationale sowie die möglichen Störeinflüsse sollen detailliert beschrieben werden.

- Ebenso wie die Auswahl soll auch die Anzahl der Einflussvariablen (Ereignisse pro Variable) und ihre möglichen Auswirkungen auf die Stabilität der Parameterschätzungen detailliert begründet und beschrieben werden.
- An die Analyse fehlender Werte und sogenannter Extremwerte sollten höchste Anforderungen gestellt werden. Die Suche nach Ursachen und ihre Ergebnisse sowie der Umgang mit den Ursachen und Ergebnissen soll detailliert beschrieben werden.
- Die Kodierung und eventuelle Transformation von Variablen sowie die zugrunde liegende Rationale muss detailliert beschrieben werden. Nach einer Umkodierung bzw. Transformation von Variablen soll eine erneute Prüfung auf Multikollinearität durchgeführt werden.
- Erfolgt die Auswahl möglicher Einflussvariablen mittels statistischer Verfahren (zum Beispiel univariate logistische Regression) unter Anwendung eines Signifikanzwertes, soll die Rationale für den gewählten Signifikanzwert detailliert beschrieben werden. (5.14.2)
- Erfolgt die Auswahl möglicher Einflussvariablen mittels univariater logistischer Regression, soll die Anwendung von Bootstrapping-Verfahren in Betracht gezogen werden. (5.14.3)

### Nutzung von administrativen und Abrechnungsdaten (Routinedaten)

- Routinedaten unterliegen Limitierungen, die man verstehen und berücksichtigen sollte. Ihre Validität soll für jede Fragestellung individuell geprüft werden.
- Selbst hochwertige Risikoadjustierungsmodelle können nicht alle Einflussgrößen berücksichtigen. Dies gilt insbesondere für Modelle, die auf administrativen Routinedaten basieren. Die Entwicklung eines Risikoadjustierungsmodells auf Basis von Routinedaten soll unter Kenntnis der Limitierungen dieser Daten erfolgen.
- Ein Modell zur Risikoadjustierung, welches auf Routinedaten basiert, ist wegen der Limitierungen der Routinedaten oftmals nur ein Surrogat für ein Modell, welches zusätzlich oder ausschließlich klinische Daten berücksichtigt. Dies gilt vor allem für Daten, die den Schweregrad einer Erkrankung beschreiben.
- Nutzung besonderer statistischer Verfahren: Multiple logistische Regression (MLR)
- Bei der Planung und Durchführung der MLR sowie der Prüfung ihrer Ergebnisse und etwaigen Korrekturen soll vorhandenes Wissen zur Fragestellung (Literatur, Leitlinien etc.), klinische und statistische Expertise sowie Fragestellung und Zielsetzung des Modells berücksichtigt werden.
- Erfolgt die MLR mittels automatisierter Variablenselektion, so soll die Auswahl des Verfahrens detailliert begründet und beschrieben werden.

- Erfolgt die MLR mittels automatisierter Variablenselektion, so wird empfohlen, diese sowohl mittels Vorwärts- als auch Rückwärtsselektion durchzuführen und die Ergebnisse dann zu vergleichen.
- Es wird empfohlen, ein „Referenzmodell“ ohne automatisierte Verfahren zu entwickeln, welches auf vorhandenem Wissen zur Fragestellung (Literatur, Leitlinien etc.) sowie klinischer und statistischer Expertise beruht. Dieses Modell kann dann einem mittels automatisierter Variablenselektion entwickeltem Modell gegenübergestellt werden.
- Wenn ausgewählte Einflussvariablen unabhängig von statistischen Selektionskriterien in einem Risikoadjustierungsmodell enthalten sein sollen, so soll die zugrunde liegende Rationale detailliert beschrieben werden.
- Erfolgt die MLR unter Anwendung eines Signifikanzwertes, soll die Rationale für den gewählten Signifikanzwert detailliert beschrieben werden.
- Es wird empfohlen, die MLR nicht nur auf Signifikanzwerte zu stützen, sondern auch andere Selektionskriterien zu verwenden, wie zum Beispiel das Bayes Information Criterion (BIC) oder Akaike Information Criterion (AIC).
- Es wird empfohlen, die Anwendung von sogenannten „Bootstrapping-Verfahren“ in Betracht zu ziehen.
- Geprüft werden sollte
  - auf einflussreiche Beobachtungen (Leverage, Diskrepanz)
  - auf wenig wahrscheinliche Fälle
  - die Anzahl der Kovariablenklassen
  - auf Interaktionen
  - auf Kollinearität und Multikollinearität. Hierbei wird die Anwendung von statistischen Kriterien (zum Beispiel Variance Inflation Factor, VIF) empfohlen.
- Die Prüfung und auch der Umgang mit ihren Ergebnissen sollte jeweils detailliert beschrieben werden. Dies gilt auch für die Nichtdurchführung.

### Modellvalidierung

- Eine Modellvalidierung soll durchgeführt werden. Die Auswahl der Methodik, die Prüfung und auch der Umgang mit ihren Ergebnissen soll detailliert beschrieben werden. Dies gilt auch für die Nichtdurchführung.
- Es wird davon abgeraten, die Modellvalidierung anhand der Daten durchzuführen, mit denen das Modell entwickelt wurde.
- Ziel der Modellvalidierung ist die Prüfung der Prognosegüte, die ein Modell ausserhalb der zur Entwicklung herangezogenen Daten erreicht.
- Bei der Planung und Durchführung der Modellvalidierung sowie den Prüfungen ihrer Ergebnisse und etwaigen Korrekturen soll vorhandenes Wissen zur Fragestellung (Literatur, Leitlinien etc.), klinische und statistische Expertise sowie Fragestellung und Zielsetzung des Modells berücksichtigt werden.
- Es wird empfohlen, die Anwendung von Bootstrapping-Verfahren in Betracht zu ziehen. Diese Empfehlung gilt gleichermaßen für die interne und externe Validierung. ▶

- Es wird empfohlen, die Anwendung von Shrinkage-Verfahren in Betracht zu ziehen.
- Es wird empfohlen, bei der Validierung von ICD-basierten Modellen die Anwendung von ICD-Daten aus anderen Ländern in Betracht zu ziehen.

## Bewertung der Prognosegüte

- In der Literatur wird eine Vielzahl von Gütemaßen beschrieben und kritisch diskutiert. Als allgemeine Erkenntnis und Anforderung lässt sich aus dieser Diskussion ableiten, dass die Prognosegüte eines Risikoadjustierungsmodells analysiert, bewertet und publiziert werden muss.
- Ein Standard im Sinne grundsätzlich oder zwingend anzuwendender Gütemaße kann aufgrund der Literatur nicht formuliert werden. Die Diskussion zeigt jedoch, dass die Entscheidung für oder gegen ein bestimmtes Gütemaß im Kontext der Zielsetzung eines Risikoadjustierungsmodells getroffen werden soll.
- Eine Bewertung der Prognosegüte soll durchgeführt werden. Die Auswahl der Gütemaße, die Festlegung der akzeptierten Mindestwerte jedes Gütemaßes, die Prüfung und auch der Umgang mit ihren Ergebnissen soll detailliert beschrieben werden. Dies gilt auch für die Nichtdurchführung.
- Bei der Planung und Durchführung der Bewertung der Prognosegüte sowie der Prüfung ihrer Ergebnisse und etwaigen Korrekturen soll berücksichtigt werden: vorhandenes Wissen zur Fragestellung (Literatur, Leitlinien etc.), klinische und statistische Expertise, Fragestellung und Zielsetzung des Modells.
- Da ein allgemein akzeptiertes, „universelles“ Prognosegütemaß nicht existiert, wird zur bestmöglichen Erkennung aller für die Prognosegüte relevanten Effekte die Anwendung verschiedener Gütemaße empfohlen.
- Die Methodenkombination soll globale Gütemaße (z. B. die erklärte Variation  $R^2$ , Brier Score) und solche zur Diskrimination (zum Beispiel Sensitivität, Spezifität, Youden-Index, Likelihood Ratios, prädiktive Werte, Separationsindex [PSEP], c-Statistik, Somer's D, Net Reclassification Improvement [NRI], Integrated Discrimination Improvement [IDI])

- und Kalibrierung (zum Beispiel Analyse von intercept und slope, Kalibrierungsgrafik, Goodness-of-Fit-Tests) umfassen.
- Werden Ergebnisse „cut-off point“ abhängiger Gütemaße berichtet, so soll der „cut-off point“, auf den sich die Ergebnisse beziehen, genannt und die Rationale für seine Auswahl detailliert erläutert werden.
- Sollen „cut-off point“ abhängige Gütemaße angewendet werden, so wird empfohlen, diese für mehrere „cut-off points“ zu berechnen und übergreifend zu bewerten.
- Es wird empfohlen, die Prüfung der Prognosegüte auch getrennt für relevante Subgruppen durchzuführen, falls diese identifiziert werden. Die Prüfung auf relevante Subgruppen und ihr Ergebnis soll detailliert beschrieben werden.
- Es wird empfohlen, die Ergebnisse der Prüfungen der Prognosegüte mittels geeigneter grafischer Verfahren zu visualisieren. ROC-Kurve und Kalibrierungsgrafik werden hierbei als Mindestanforderung betrachtet.

## Übertragbarkeit

- Es wird empfohlen, die Übertragbarkeit eines Risikoadjustierungsmodells unter besonderer Berücksichtigung etwaiger Unterschiede im Casemix detailliert zu prüfen und die Ergebnisse der Prüfung detailliert zu berichten. Dies gilt auch für die Nichtdurchführung.
- Im Sinne einer Mindestanforderung gehören dazu (soweit zutreffend) für den Entwicklungs- und Validierungsdatensatz Angaben zu den Einflussvariablen und weiteren Daten, die die verschiedenen Populationen beschreiben: absolute und relative Häufigkeiten, statistische Maßzahlen mit der Angabe von Konfidenzintervallen, Prüfung auf Unterschiede zwischen den Populationen.
- Diese Empfehlungen gelten insbesondere, wenn ein Modell für eine spezifische Modellgruppe entwickelt wurde und dann auf eine andere Population übertragen werden soll zum Zwecke externer Validierung und einer öffentlichen Berichterstattung, bei der die Ergebnisse einzelner Krankenhäuser veröffentlicht werden sollen.

## Anmerkungen

- 1) DIN EN ISO 9000:2015-11: Qualitätsmanagementsysteme – Grundlagen und Begriffe. Berlin: Beuth-Verlag 2015
- 2) Sens, B., Fischer, B., Bastek, A., Eckardt, J., Kaczmarek, D., Paschen, U., Pietsch, B., Rath, S., Ruprecht, T., Thomeczek, C., Veit, C., Wenzlaff, P., Begriffe und Konzepte des Qualitätsmanagements – 3. Auflage. GMS Med Inform Biom Epidemiol. 2007;3 (1)
- 3) Joint Commission on Accreditation of Healthcare Organisations. Primer on indicator development and application. Measuring quality in health care. Oakbrook Terrace, Illinois: Joint Commission on Accreditation of Healthcare Organisations, 1990.
- 4) Blumenstock, G., Zur Qualität von Qualitätsindikatoren. Bundesgesundheitsblatt 2011; 54: 154–159
- 5) Glattacker, M., Jäckel, WH., Evaluation der Qualitätssicherung – aktuelle Datenlage und Konsequenzen für die Forschung. Gesundheitswesen 2007; 69: 277–283
- 6) Scholle, SH., Roski, J., Adams, JL., Dunn, DL., Kerr, EA., Dugan, DP., Jensen, RE., Benchmarking physician performance: reliability of individual and composite measures. Am J Manag Care 2008; 14: 833–838

Anzeige



Ihr Partner für telemedizinische  
Produkte & Lösungen

Zentrum für Telemedizin Bad Kissingen | [www.ztm-badkissingen.de](http://www.ztm-badkissingen.de)  
E-Mail: [info@ztm-badkissingen.de](mailto:info@ztm-badkissingen.de) | Telefon: 78 55 29-0

- 7) Safford, MM., Improving the quality of quality measurement. The tinkerer, the tailor and the candlestick maker. *Medical Care* 2007; 47: 375–377
- 8) Stausberg, J., Halim, A., Färber, R., Concordance and robustness of quality indicator sets for German hospitals. *BMC Health Services Research* 2011; 11: 106
- 9) Swinglehurst, D., Emmerich, N., Maybin, J., Park, S., Quilligan, S., Confronting the quality paradox: towards new characterisations of 'quality' in contemporary healthcare. *BMC Health Services Research* 2015; 15: 240
- 10) Thielscher, C., Antoni, B., Driedger, J., Jacobi, S., Krol, B., Geringe Korrelation von Krankenhausführern kann zu verwirrenden Ergebnissen führen. *Gesundheitsökonomie & Qualitätsmanagement*. 2014; 19: 65–69
- 11) Donabedian A. The quality of medical care. In: Graham NO, ed. *Quality assurance in hospitals*. Rockville, Maryland: Aspen, 1982: 15–34
- 12) AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH. *Allgemeine Methoden im Rahmen der sektorenübergreifenden Qualitätssicherung im Gesundheitswesen nach § 137a SGB V. Version 3.0*. Göttingen, Stand: 15. Februar 2013
- 13) Reiter, A., Fischer, B., Kötting, J., Geraedts, M., Jäckel, WH., Barlag, H., Döbler, K., QUALIFY: Ein Instrument zur Bewertung von Qualitätsindikatoren. Düsseldorf: BQS, 2007
- 14) McDonald, K., Romano, P., Geppert, J., Davies, SM., Duncan, BW., Shojania, KG., Hansen, A., Measures of patient safety based on hospital administrative data. The Patient Safety Indicators. Technical Review 5 (Prepared by the University of California San Francisco-Stanford Evidence-based Practice Center under Contract No. 290–97–0013). AHRQ Publication No. 02-0038. Rockville, MD: Agency for Healthcare Research and Quality. August 2002
- 15) Davies, S., Geppert, J., McClellan, M., McDonald, KM., Romano, PS., Shojania, KG., Refinement of the HCUP quality indicators. Technical Review Number 4. Rockville, MD: (Prepared by UCSF-Stanford Evidence-based Practice Center under Contract No. 290–97–0013) Agency for Healthcare Research and Quality; 2001. Report No.: 01–0035
- 16) Rode, S., Ries, V., Petzold, T., Buch, U., Untersweg, F., Fischer, B.: Anforderungen und Anwendungshinweise für den Einsatz von Qualitätsindikatoren in der medizinischen Versorgung: Ergebnisse eines systematischen Reviews. *Das Gesundheitswesen – May 2016*. DOI: 10.1055/s-0042-102347
- 17) Veit, C., Bauer, J., Döbler, K., Fischer, B., Hrsg. *Qualität sichtbar machen. BQS-Qualitätsreport 2008*. Düsseldorf: BQS Bundesgeschäftsstelle Qualitätssicherung gGmbH, 2009
- 18) Becker, A., *Qualitätskriterien der Risikoadjustierung am Beispiel der logistischen Regression und ihre Anwendung bei der öffentlichen Berichterstattung*. Ein Diskussionsbeitrag. Hrsg.: Becker, A. 1. Auflage (Eigenverlag), Dezember 2014
- 19) Hierunter versteht man eine Einflussgröße, die entgegen bekannter (patho-)physiologischer Grundlagen oder klinischer Erfahrung oder wissenschaftlicher Erkenntnisse (beispielsweise aus prospektiven Registern) einen nicht erwarteten oder nicht plausiblen protektiven Effekt aufweist. In der Fachliteratur aufgeführte Beispiele sind u.a. Hypertonus oder Diabetes, die in Risikoadjustierungsmodellen zur Krankenhaussterblichkeit von Herzinfarktpatienten zu einer niedrigeren geschätzten Krankenhaussterblichkeit führen. Mit anderen Worten: bei Patienten mit dokumentiertem Hypertonus oder Diabetes ist die durch das Risikoadjustierungsmodell geschätzte Sterbewahrscheinlichkeit (paradoxe Weise) niedriger als bei Patienten, bei denen diese Komplikationen nicht vorliegen. Eine mögliche Erklärung ist, dass bei Patienten, die bei der Erstbehandlung versterben, keine vollständige Anamnese erhoben wird und dementsprechend Grund- oder Begleiterkrankungen unvollständig dokumentiert werden.

### Anschrift der Verfasser

*Prof. Dr. Andreas Becker, Institut Prof. Dr. Becker, Nonnenweg 120a, 51503 Rösrath/Prof. Dr. med. Jürgen Stausberg, Arzt für Medizinische Informatik und Ärztliches Qualitätsmanagement, Kordulastr. 13, 45131 Essen/Burkhard Fischer, Referatsleiter Qualitätsmanagement, IT und Datenanalyse, KGNW, Humboldtstraße 31, 40237 Düsseldorf/Dr. Dirk Carstanjen, DKG, Wegelystraße 3, 10623 Berlin/Prof. Dr. Maria Weyermann, FB Gesundheitswesen, Hochschule Niederrhein, Reinartzstraße 49, 47805 Krefeld*

## HEUKING KÜHN LÜER WOJTEK



Herz- und Diabeteszentrum NRW Arzneimittel-Versandhandel

Universitätsklinikum Aachen AöR

Beratung bei der Wahl ärztlicher Kooperations- und Rechtsformen

AKH-Gruppe

Kreiskrankenhaus Herdecke

Transaktionsberatung Privater Klinikträger

Vitos GmbH

## Rechtssicherheit für das Gesundheitswesen

Unsere Rechtsanwälte und Steuerberater vertreten Krankenhäuser und Universitätskliniken, Arztpraxen, Pharmaunternehmen, Medizinproduktehersteller, Krankenkassen und wirtschaftliche Einrichtungen. Wir beraten Sie in verschiedenen Rechtsgebieten in Bezug auf alle Fragestellungen und Neuerungen im Gesundheitswesen sowie im Bereich Life Science. Von der öffentlichen Ausschreibung und Kooperation über Ausgliederungen und Transaktionen bis hin zu Umstrukturierungen: In allen Belangen stehen Ihnen unsere Experten zur Seite. Unser Health Care Team ist auch in diesem Jahr auf der Medica in Düsseldorf vertreten. Wir freuen uns auf Ihre Kontaktaufnahme. | [www.heuking.de](http://www.heuking.de)

[www.heuking.de](http://www.heuking.de)

|            |           |         |
|------------|-----------|---------|
| Berlin     | Hamburg   |         |
| Chemnitz   | Köln      |         |
| Düsseldorf | München   | Brüssel |
| Frankfurt  | Stuttgart | Zürich  |

RECHTSANWÄLTE UND STEUERBERATER