



Von Daten zu Informationen zu Entscheidungen – wie können statistische Grafiken die klinische Qualitätsbewertung im Krankenhaus unterstützen?

Data to Information to Decisions – how can Statistical Graphics support Clinical Quality Evaluation in Hospitals?

Eine Übersicht und praktische Anwendung am Beispiel der
Krankenhaussterblichkeit bei ambulant erworbener Pneumonie

An overview and practical application, taking the Mortality Rate
in Hospitals for Community-Acquired Pneumonia as an example

Becker A¹

¹ CLINOTEL Krankenhausverbund
gemeinnützige GmbH, Köln

Vorwort

Deutschland hat zwar eine lange Tradition von Qualitätssicherung und Qualitätsmanagement mit dem Sammeln, Auswerten und Darstellen qualitätsrelevanter Kenngrößen. Die Akzentuierung des unerlässlichen Prozesses der Analyse und Bewertung in der eigenen Fachabteilung, der eigenen Versorgungseinrichtung mit der Konsequenz der faktenbasierten Initiierung von Veränderungsprozessen trat aber demgegenüber bisher in den Hintergrund.

Aufgrund aktueller gesundheitspolitischer Rahmenbedingungen hat sich aber die Relevanz der Qualitätsfrage deutlich verändert. Qualität und Transparenz über die Versorgungsqualität haben eine hohe Bedeutung erlangt, die perspektivisch noch zunehmen wird. Dies hat zur Konsequenz, dass einerseits die qualitätsorientierte Unternehmensführung mit entsprechenden Strategien, Maßnahmen und Kennzahlen zur Steuerung der Leistungsprozesse eine höhere Bedeutung erlangt. Andererseits wird es umso wichtiger, adäquate Instrumente zu nutzen, die die Qualitätsmessung und -beurteilung unterstützen. Wenn die Führungsebene einer Gesundheitseinrichtung in zunehmendem Maß aktiv in die Qualitätsentwicklung eingebunden ist, müssen die involvierten Personen sicher sein können, dass »das Richtige« gemessen und interpretiert wird.

Korrespondierender Autor

✉ Prof. Dr. med. Andreas Becker

Zitierung

Becker A. Von Daten zu Informationen zu Entscheidungen – wie können statistische Grafiken die klinische Qualitätsbewertung im Krankenhaus unterstützen? Eine Übersicht und praktische Anwendung am Beispiel der Krankenhaussterblichkeit bei ambulant erworbener Pneumonie.

Interdisciplinary Contributions to Hospital Management: Medicine, Patient Safety and Economics. 18.12.2013 #017.

<http://www.clinotel-journal.de/article-id-017.html>

Die vorliegende Arbeit setzt genau auf diesem Szenario moderner, qualitätsorientierter, kennzahlengestützter Unternehmensführung auf. Auf der Basis von Routinedaten der Krankenhausbehandlung werden sehr fundiert Methoden der biometrischen Analyse und darauf basierenden grafischen Darstellung am Beispiel einer Qualitätskennzahl dargestellt. Darunter ist die bislang viel zu wenig beachtete Qualitätsregelkarte, die kaum verbreitete Cumulative-Sum-Technik (CUSUM) und – erstmals in einer deutschsprachigen Publikation – das Variable Life Adjusted Display (VLAD). Diese Methode wird zudem weiterentwickelt, um komplizierende Nebendiagnosen erweitert und somit für die eigene Organisation maßgeschneidert (»tailor-made« im besten Sinne).

Der Fokus liegt jedoch auf der Handlungsrelevanz: »Statistische Signifikanz« soll mit entsprechender Sensitivität und Spezifität in »klinische Bedeutung« übersetzt werden. Dabei wird der Blick aber nicht nur – wie besonders in den gesundheitspolitisch oktroyierten Qualitätssicherungsmaßnahmen zu Recht kritisiert – auf negative Auffälligkeiten gelenkt, sondern auch die besonders positiv hervorzuhebenden Ergebnisse, der »run of good performance«, das Lernen von den Besten, werden in den Fokus gerückt.

Basierend hierauf werden klare Forderungen formuliert, wie aktuell eine an qualitätsorientierter Unternehmensführung ausgerichtete Analyse, Bewertung und Umsetzung von Maßnahmen zur Qualitätsverbesserung aussehen sollte, die wir in dieser Konsequenz (zu) selten lesen oder hören:

Eine qualitätsorientierte Unternehmenspolitik mit klaren qualitätsbezogenen Zielen und einem routinemäßig implementierten Qualitätsmonitoring ist Grundlage von Qualitätsdarlegung und Transparenz nach innen und außen und dient der ständigen Qualitätsverbesserung und dem langfristigen Unternehmenserfolg.

Damit stellt der Autor nicht nur Methoden und Instrumente der statistischen Darstellung von Qualitätsindikatoren vor, sondern setzt Akzente für eine neue Qualität von Krankenhausführung. Prädikat: Besonders wertvoll.

DR. PHIL. BRIGITTE SENS

Vorsitzende der Gesellschaft für Qualitätsmanagement
in der Gesundheitsversorgung e.V. (GQMG)

Zusammenfassung

Die Qualitätsbewertung klinischer Prozesse gehört zu den herausragenden Anforderungen an (klinische) Leitungskräfte. Sollen in diesem Kontext wirksame Entscheidungen getroffen werden, müssen hierzu Daten analysiert und zu Informationen transformiert werden.

Statistische Grafiken wie Zeitreihen, Regelkarten, Funnel Plots, Cumulative Sum (CUSUM) und das Variable Life Adjusted Display (VLAD) können den Transformationsprozess »Datum – Information – Entscheidung« wirksam unterstützen, wie wir am Beispiel der Daten zweier Krankenhäuser zur ambulant erworbenen Pneumonie zeigen konnten.

Falls eine weiterführende Analyse der Daten erforderlich ist, sollte sie – ebenso wie die klinische Diagnostik – geplant und zielgerichtet durchgeführt werden. Hierzu empfehlen wir die Analysepyramide (Mohammed et al. 2004).

Neben der Erhebung und Analyse prozessrelevanter Daten empfehlen wir die Morbiditäts- und Mortalitätskonferenz und das Prozessaudit als weitere Maßnahmen, die wichtige Erkenntnisse zur Qualitätsbewertung, Prävention und Qualitätsplanung liefern können.

Das von uns geschilderte Vorgehen unterstützt (klinische) Leitungskräfte bei der Bewertung klinischer Prozesse hinsichtlich der Prozess- und Ergebnisqualität, deren Kenntnis wiederum eine Voraussetzung für die Bewertung der Effizienz darstellt.

Schlüsselwörter

Qualität, Qualitätsbewertung, Qualitätsverbesserung, klinische Leitung, ambulant erworbene Pneumonie, Risikoadjustierung, Zeitreihe, Regelkarte, Funnel Plot, Cumulative Sum (CUSUM), Variable Life Adjusted Display (VLAD), DIN EN 15224

Lesehinweis

Alle [Abbildungen](#) dieses Artikels finden Sie ab Seite [34](#).

Die [Tabellen](#) dieses Artikels finden Sie ab Seite [57](#).

Summary

The evaluation of the quality of clinical processes is one of the major performance criteria required of (clinical) leadership. If effective decisions are to be made within this context, the data collected must be analysed and transformed to information.

Statistical graphics such as timelines, control charts, funnel plots, cumulative cum (CUSUM) and variable life adjusted display (VLAD) charts can effectively support »data – information – decision« transformation, as we were able to show from the data collected from two hospitals on community-acquired pneumonia.

Should further analysis of the data be required, then this should be planned and carried out with purpose and precision – as is the case with clinical diagnostics. To this end we recommend applying the analysis pyramid method (Mohammed et al. 2004).

Besides collecting and analysing process relevant data, we recommend conferring morbidity and mortality rates and executing a process audit as further measures for delivering important information relevant to quality evaluation, prevention and quality planning.

The procedure described supports (clinical) leadership in the evaluation of clinical processes with respect to process and outcome quality, the knowledge of which is a prerequisite for the evaluation of efficiency.

Keywords

Quality, Quality Evaluation, Quality Improvement, Clinical Leadership, Community-Acquired Pneumonia, Risk Adjustment, Timeline, Control Chart, Funnel Plot, Cumulative Sum (CUSUM), Variable Life Adjusted Display (VLAD), EN 15224

Einleitung

Zielsetzung

Es wird ein Überblick zu statistischen Grafiken gegeben und deren Anwendungsmöglichkeiten zur Unterstützung der klinischen Qualitätsbewertung und Entscheidungsfindung im Krankenhaus erläutert.

Vorbemerkung

Eine angemessene Patientenversorgung in Übereinstimmung mit dem aktuellen medizinischen Wissen und der besten verfügbaren Praxis ist elementarer Bestandteil der Qualitätspolitik des CLINOTEL-Krankenhausverbundes. Dieses Verständnis wird gelebt durch kontinuierliches Streben nach Verbesserung. Hierzu gehen wir den Ursachen auf den Grund, um die Fakten herauszufinden, die uns zu korrekten Entscheidungen befähigen. Wir pflegen eine umfassende interne und externe Transparenz durch offenen Umgang mit Qualitätsdaten.

Seit 1999 arbeiten wir an Qualität durch Einsatz von Instrumenten des Qualitätsmanagements. Auditierungsverfahren geben Hinweise zur Standortbestimmung sowie Empfehlungen zur Verbesserung und überprüfen die Umsetzung. Die Ergebnisqualität hat einen besonderen Stellenwert, daher messen wir sie und stellen die Informationen so zur Verfügung, dass sie bewertet und Maßnahmen abgeleitet werden können. Hierbei legen wir großen Wert darauf, dass sich die Ergebnisgrößen in die Unternehmenssteuerung integrieren lassen und einen kontinuierlichen Benchmarkingprozess unterstützen.

Einen besonderen Stellenwert hat hierbei die Qualitätssicherung mit Routinedaten, die wir intensiv seit dem Jahr 2003 betreiben. Im Laufe der Jahre haben wir systematisch Darstellungsformen aus der Literatur recherchiert und auf ihre Anwendbarkeit in der datenbasierten Steuerung von klinischen Prozessen überprüft. Dabei ist ein in sich geschlossener Werkzeugkasten entstanden, der sowohl das zeitliche Monitoring als auch das »Sich-mit-anderen-Vergleichen« unterstützt. Das alles nicht als Selbstzweck, sondern um in Steuerungs- und Verbesserungsprozessen kontinuierlich an der Qualität und Sicherheit der klinischen Behandlung zu arbeiten. Herz des Ganzen sind geprüfte, zeitnah verfügbare Routinedaten. In diesem Beitrag wird aber nicht der Datenmanagementprozess, sondern die Übersetzung in klinisch relevante Informationen und damit einer datenbasierten Entscheidungsunterstützung betrachtet.

Die Arbeit soll einerseits als Rationale für die vielen in CLINOTEL-Mitgliedshäusern für klinische Prozesse Verantwortlichen dienen, die regelmäßig mit diesen Werkzeugen arbeiten und so die Gelegenheit bekommen, ihr Verständnis für die Methoden zu vertiefen. Andererseits soll sie eine über den Verbund hinausgehende Diskussion über die besten Methoden für das Monitoring und den Vergleich klinischer Behandlungsprozesse anhand von Kennzahlen auslösen.

»Das Vergleichen ist das Ende des Glücks und der Anfang der Unzufriedenheit.«

Søren Kierkegaard

Struktur der Arbeit

Im nächsten Abschnitt (Hintergrund) wird zunächst eine Einführung zur Qualität klinischer Prozesse unter besonderer Berücksichtigung der Effizienz gegeben. Darauf aufbauend werden die verschiedenen Schritte einer sachbezogenen Entscheidungsfindung erläutert.

Bei den »Grundlagen« wird der Begriff der »Routinedaten« kurz dargestellt. Im Anschluss wird die Datengrundlage der Auswertungen und Grafiken erläutert, gefolgt von der Definition des Begriffs »Letalität« und Ausführungen zur Bedeutung und Anwendung der »Risikoadjustierung«.

Alle dann nachfolgend aufgeführten Grafiken bilden Daten zur Krankenhausletalität vollstationär behandelter Patienten mit ambulant erworbener Pneumonie ab. Dieses Vorgehen hat den Vorteil, dass die unterschiedlichen Potenziale der einzelnen Grafiken besonders gut zur Geltung kommen.

Im Abschnitt »Nicht risikoadjustierte Krankenhausletalität« werden Grafiken mit ihren Interpretationsmöglichkeiten erläutert und Empfehlungen abgeleitet. Weiterhin wird die sogenannte Analysepyramide als Konzept zur Bewertung klinischer Prozesse vorgestellt:

- Balkendiagramm
- Exkurs: Analyse klinischer Prozesse
- Zeitreihe
- Regelkarte
- Cumulative Sum (CUSUM)

In einem nächsten Abschnitt werden auch wieder Grafiken mit ihren Interpretationsmöglichkeiten erläutert und Empfehlungen abgeleitet. Diesmal geht es um die Darstellung »Risikoadjustierter Krankenhausletalität«:

- Balkendiagramm (beobachtet versus erwartet)
- Balkendiagramm (risikoadjustierte Krankenhausletalität)
- Zeitreihe
- Funnel Plot
- Variable Life Adjusted Display (VLAD)

Im letzten Abschnitt werden dann CUSUM und VLAD gemeinsam betrachtet. Abgeschlossen wird die Arbeit mit einer Gesamtsicht der Ergebnisse und den Schlussfolgerungen.

Hintergrund

Klinische Prozesse sind die Grundlage eines Krankenhauses. Sie bilden den Haupttyp der Dienstleistungen, denn hier erfolgen Diagnostik und Therapie mit der Zielsetzung, die Anforderungen der Patienten in angemessenem Umfang zu erfüllen. Der Grad, zu dem diese Anforderungen erfüllt werden, wird im Allgemeinen mit dem Begriff »Qualität« beschrieben. Um die Qualität klinischer Prozesse messbar und nachprüfbar zu gestalten, bedarf es definierter Qualitätsmerkmale, so wie sie die noch recht neue DIN EN 15224:2012-12 (DIN EN 15224 2012) beschreibt.

Eines dieser Merkmale ist die »Wirksamkeit«, die in der DIN EN 15224 wie folgt erläutert wird: »Tätigkeiten der Gesundheitsversorgung verbessern im Vergleich zu nicht oder anderweitig durchgeführten Untersuchungen oder Behandlungen in einem sinnvollen Ausmaß die Wahrscheinlichkeit eines erwarteten positiven Ergebnisses.« »Wirksamkeit« fokussiert also auf Ergebnisse und steht damit in Einklang mit dem bekannten Konzept der »Ergebnisqualität«.

Wirksamkeit beziehungsweise Ergebnisqualität werden in besonderem Maße durch eine angemessene und richtige Versorgung, die wissenschaftlich abgesichert ist und sich auf das beste Wissen und die beste Praxis stützt, beeinflusst. Diese Betrachtung erklärt auch die Notwendigkeit nach klinischen Prozessen, deren Risiken unter Kontrolle sind und vermeidbare Schäden auf ein Restrisiko reduziert sind.

Ein weiteres Qualitätsmerkmal aus der DIN EN 15224 ist »Effizienz«, hier definiert als »das bestmögliche Verhältnis zwischen den erreichten Ergebnissen und den aufgewendeten Ressourcen.« Die Effizienz verbindet also die medizinische und ökonomische Sicht auf die klinischen Prozesse, indem sie die erreichten Ergebnisse in ein Verhältnis zu den aufgewendeten Ressourcen setzt.

Über den Zusammenhang von Ergebnisqualität und Kosten sowie den (vermeidbaren) Kosten (vermeidbarer) unerwünschter Ereignisse wurde bereits an anderer Stelle umfassend ausgeführt (Becker et al. 2006).

Das Lenken und Leiten eines Krankenhauses beziehungsweise seiner medizinischen Organisationseinheiten sollte also in erster Linie auf die Qualität unter besonderer Berücksichtigung der Wirksamkeit beziehungsweise Ergebnisqualität ausgerichtet sein.

Eine gewisse Operationalisierung der Handlungen »Lenken« und »Leiten« kann wiederum aus der DIN EN 15224 abgeleitet werden. Diese finden sich in der Norm unter der Überschrift »Grundsätze des Qualitätsmanagements«.

Im Kontext der bisherigen Ausführungen hat der Normgrundsatz »sachbezogene Herangehensweise an das Treffen von Entscheidungen« eine besondere Bedeutung, denn »wirksame Entscheidungen beruhen auf der Analyse von Daten und Informationen«. Hieraus lässt sich die im Titel der Arbeit enthaltene Sequenz »Daten – Informationen – Entscheidungen« ableiten.

Grundlage sachbezogener Entscheidungsfindung sind also geeignete Daten, die die Organisation ermittelt. Solche Daten zu klinischen Prozessen und deren Ergebnissen können beispielsweise aus Routinedaten gewonnen werden (weitere Ausführungen dazu im nächsten Abschnitt).

Daten an sich haben keinen Wert. Sie erhalten einen Wert erst durch den ersten Transformationsschritt (**Abb. 1**), in dem einem Datum durch qualifizierte Mitarbeiter eine Bedeutung gegeben wird. Somit wird das Datum zur »Information«. Bateson versteht unter der »elementaren Informationseinheit« einen »Unterschied, der einen Unterschied ausmacht.« (Bateson 1972, Seite 582) und lenkt unsere Aufmerksamkeit damit auf eine notwendigerweise zu stellende Frage, deren Beantwortung wiederum Ausgangspunkt für die Entscheidungsfindung ist. Als Beispiele aus der klinischen Leitungspraxis können hier etwa die folgenden Fragen gestellt werden:

- »Ist bei Patienten mit ambulant erworbener Pneumonie eine 1%ige Abweichung der Krankenhausletalität von entsprechenden Vergleichsdaten ein Unterschied, der einen Unterschied ausmacht, der also Handeln erfordert?«
- »Wenn die monatliche Krankenhausletalität von Patienten mit ambulant erworbener Pneumonie in meiner Fachabteilung in einem Bereich von 8% bis 12%

schwankt, ist das eine zufällige Variation oder handelt es sich um einen Unterschied, der einen Unterschied macht, der also Handeln erfordert?»

Im nächsten Transformationsschritt wird der Information dann wiederum durch qualifizierte Mitarbeiter eine Bedeutung gegeben und eine Entscheidung kann getroffen werden.

Der hier beschriebene Transformationsprozess mit seinem Fokus auf dem bedeutsamen Unterschied soll ermöglichen, sachbezogene Entscheidungen zu treffen, die nur dann zu Interventionen führen, wenn diese auch wirklich erforderlich sind. Es geht also – wie auch in den klinischen Prozessen selbst – um die Indikationsstellung – das Für und Wider von Aktivitäten mit dem Ziel einer Qualitätsverbesserung.

Besteht in Wirklichkeit gar keine Notwendigkeit für derartige Aktivitäten (zum Beispiel, weil es sich bei den oben erwähnten Schwankungen nur um zufällige monatliche Variation der Daten handelt), werden möglicherweise finanzielle und personelle Ressourcen gebunden, die dann an anderer Stelle fehlen.

Hinzu kämen natürlich auch negative Effekte auf die Motivation der beteiligten Mitarbeiter, wenn sie erkennen, dass die durchgeführten Aktivitäten nicht zu den gewünschten Veränderungen der Daten (hier also der Krankenhausletalität) führen.

Wenn einem Datum eine Bedeutung beigemessen werden soll, kann eine angemessene grafische Darstellung äußerst hilfreich sein.

Grundlagen

An dieser Stelle werden einige Punkte erläutert, die für das Verständnis der weiteren Ausführungen von Bedeutung sind:

Routinedaten

Im Krankenhausbereich können unter dem Begriff »Routinedaten« sowohl klinische (zum Beispiel anästhesiologische Daten im Rahmen einer Risikoeinschätzung/Narkoseführung oder die Basisdokumentation in der Psychiatrie) als auch administrative Daten (zum Beispiel nach § 21 Krankenhausentgeltgesetz – KHEntgG) verstanden werden. Die vorliegende Arbeit beschäftigt sich nachfolgend mit administrativen Routinedaten in Krankenhäusern, hier kurz »Routinedaten« genannt. Weitere Ausführungen zur Begriffsklärung, allgemeine Grundlagen und Beispiele zur Anwendung sind beschrieben bei Benchimol et al. (2011), Becker (2012) sowie Becker et al. (2005; 2012a; 2012b; 2013).

Daten

Datengrundlage der in dieser Arbeit abgebildeten Grafiken sind die von den Mitgliedskrankenhäusern an die CLINOTEL-Geschäftsstelle gelieferten Falldaten der in den Jahren 2010 bis 2012 aus vollstationärer Behandlung entlassenen 1.809.643 Patienten gemäß Definition nach § 21 Krankenhausentgeltgesetz (KHEntgG).

Die nicht personenbezogenen Daten werden im laufenden Jahr monatlich kumulativ vom 01. Januar an übermittelt. Dies hat den Vorteil, dass unterjährige Korrekturen der Daten einfließen, die zum Beispiel infolge der Ergebnisse umfangreicher Kodierungsprüfungen durchgeführt wurden (Becker et al. 2003). Die Speicherung und Verarbeitung der Daten erfolgt in einer ORACLE®-Datenbank (Oracle Corporation, California, USA).

Die nachfolgend in den Grafiken abgebildeten Daten zur Krankenhausletalität stammen von 24.689 vollstationär behandelten Patienten (Alter ≥ 18 Jahre) mit ambulant erworbener Pneumonie in den Jahren 2010 bis 2012.

Die Grafiken sind Teil der monatlichen Auswertungen für unsere Mitgliedshäuser und da die Daten im laufenden Jahr monatlich kumulativ vom 01. Januar an übermittelt werden, wachsen die zugrunde liegenden Fallzahlen von Monat zu Monat.

Letalität

Anzahl der durch eine bestimmte Erkrankung verursachten Sterbefälle bezogen auf die Anzahl der Erkrankten. Die sogenannte »In-Hospital-Letalität« oder »Krankenhausletalität« erfasst alle Patienten einer bestimmten Erkrankungs- beziehungsweise Therapiegruppe, die während ihres stationären Aufenthalts versterben.

Allgemein: prozentualer Anteil der Patienten mit einer Krankheit, die daran sterben: »Fallsterblichkeit« (in Anlehnung an Kreienbrock et al. 2012, Seite 32).

Risikoadjustierung

Ein Ziel der Analyse von Daten ist unter anderem, Behandlungsergebnisse medizinischer Versorgungseinrichtungen einander gegenüberzustellen, um Verbesserungspotenziale für die untersuchten Einrichtungen aufzuzeigen. Bei solchen Vergleichen können jedoch patientenbezogene, von der Einrichtung nicht beeinflussbare Risikofaktoren eine entscheidende Rolle spielen und bewirken, dass trotz gleicher Qualität der erbrachten Leistung unterschiedliche Behandlungsergebnisse erzielt werden.

Es können sowohl demografische Merkmale wie Alter oder Geschlecht der Patienten, als auch krankheitsspezifische Faktoren wie Schweregrad der Erkrankung oder Komorbiditäten das Behandlungsergebnis einer Einrichtung beeinflussen. So wird etwa ein Krankenhaus, das viele ältere und multimorbide Patienten behandelt, im Einrichtungsvergleich vermeintlich schlechter abschneiden als ein Krankenhaus mit einer jüngeren, gesünderen Patienten Klientel – unabhängig von der Versorgungsqualität.

Werden solche relevanten Risikofaktoren beim Gegenüberstellen der Ergebnisse vernachlässigt, kann dies zu unfairen Vergleichen und falschen Schlussfolgerungen führen. Um entscheiden zu können, ob tatsächlich Unterschiede in der Versorgungsqualität vorliegen, müssen

daher zwischen den Einrichtungen gleiche Ausgangsbedingungen im Hinblick auf die zu vergleichenden Ergebnisse geschaffen werden. Für einen fairen Klinikvergleich ist eine risikoadjustierte Betrachtungsweise also unabdingbar.

Hierzu dient die Risikoadjustierung, die mit geeigneten statistischen Methoden Unterschiede im Patientenspektrum ausgleichen soll. Ihr Ziel ist es, Einrichtungsvergleiche insofern gerecht zu gestalten, dass Unterschiede in den Ergebnissen auch tatsächlich Unterschiede in der Versorgungsqualität abbilden und nicht auf verschiedene Patientenstrukturen zurückzuführen sind. Die Risikoadjustierung ist somit vor allem für Daten von Bedeutung, die Behandlungsergebnisse abbilden sollen.

Die in dieser Arbeit verwendeten risikoadjustierten Daten basieren auf einem von uns entwickelten, noch nicht veröffentlichten Modell, welches mit der sogenannten multiplen logistischen Regression berechnet wurde. Die multiple logistische Regression ist ein statistisches Standardverfahren für die Analyse des Einflusses verschiedener Größen auf eine binäre, das heißt nur eine von zwei Möglichkeiten zulassende Zielvariable (zum Beispiel »Patient verstorben: ja/nein«). So kann unter Anwendung des Modells für jeden Patienten unter Berücksichtigung seines individuellen Risikoprofils die erwartete Wahrscheinlichkeit des vorhergesagten Ereignisses (zum Beispiel: Versterben im Krankenhaus) berechnet werden.

Nicht risikoadjustierte Krankenhausletalität

Balkendiagramm

Die auf Routinedaten basierende nicht risikoadjustierte (sogenannte rohe) Krankenhausletalität kann auf verschiedene Art und Weise grafisch dargestellt werden. Eine erste Möglichkeit ist das Balkendiagramm, welches in **Abb. 2** zu sehen ist und zu den monatlichen Standardauswertungen in unserem Krankenhausverbund gehört.

Abgebildet wird hier die nicht risikoadjustierte Krankenhausletalität der einzelnen Krankenhäuser (schwarze Balken) und auch der Gesamtwert über alle Häuser (grauer Balken). In den Auswertungen sind die einzelnen Balken mit den Namen der Krankenhäuser beschriftet, diese werden zur Anonymisierung hier nicht genannt.

In **Abb. 2** sind die Ergebnisse des Gesamtjahres 2012 zu sehen, es handelt sich also um die Auswertung zum Datenstand 31.12.2012.

Für ein einzelnes Krankenhaus, wie zum Beispiel die markierten Krankenhäuser A und B, kann die Frage gestellt werden, ob die Werte 12,0% beziehungsweise 20,7% den Erwartungen der Häuser (also einem definierten Sollwert) entsprechen. Dies setzt natürlich voraus, dass ein solcher Wert prospektiv formuliert wurde, was bei einer qualitätsorientierten klinischen Leitung zu erwarten wäre. Bei der Formulierung des Erwartungswertes könnte man Daten aus der Fachliteratur, aus Leitlinien, der externen Qualitätssicherung oder anderen Krankenhäusern heranziehen.

Diese Werte können dann natürlich auch bei der Formulierung eines Zielwertes für einen Verbesserungsprozess herangezogen werden. Zu beachten ist hierbei, dass die berücksichtigten Quellen mit hoher Wahrscheinlichkeit unterschiedliche Datengrundlagen und Definitionen verwenden. Schon hier wird klar, dass eine vergleichende Qualitätsbewertung eine sehr komplexe Aufgabe sein kann, die nicht als exakte Wissenschaft betrachtet werden sollte, da es hierbei einige Unwägbarkeiten zu berücksichtigen gilt.

Aus den bisherigen Ausführungen ergibt sich die erste Empfehlung:

Die Unternehmensleitung eines Krankenhauses und/oder die klinischen Leitungskräfte sollten Zielwerte für die Prozess- und Ergebnisqualität definierter Patientengruppen prospektiv definieren.

Der Gesamtwert über alle Krankenhäuser (grauer Balken/13,6%) bietet eine erste Hilfe, die nach unseren Erfahrungen von den Anwendern intuitiv angewendet wird, indem sie die Grafik in zwei Bereiche »links und rechts vom Verbundwert« unterteilt.

So ist es nachvollziehbar, wenn (klinische) Leitungskräfte als Ziel ausgeben »rechts vom CLINOTEL-Wert«. Auf den ersten Blick hat dieses Ziel natürlich einen gewissen Charme, da es auf der robusten Einteilung in »links und rechts vom Verbundwert« basiert. Auf den zweiten Blick verliert es jedoch an Attraktivität, da die Krankenhäuser innerhalb dieser Gruppe Werte im Bereich von 5,8% bis 13,3% aufweisen. Sinnvoll könnte die Zielformulierung »rechts vom CLINOTEL-Wert« sein, wenn ein Krankenhaus bei der ersten Teilnahme an dieser Auswertung (zum Beispiel nach Eintritt in den Verbund) einen Wert zeigt, der über dem Gruppenwert 13,6% liegt. Ein Wert $\leq 13,6\%$ könnte als Richtung für einen ersten Verbesserungsprozess formuliert werden, sollte dann aber in der Folge durch einen konkreten Zahlenwert ersetzt werden.

Auch in der anderen Gruppe finden wir unterschiedliche Werte, die sich hier in einem Bereich von 13,6% bis 20,7% ansiedeln. Eines der ersten Argumente, welches bei der Diskussion dieser Daten oftmals angebracht wird, ist der Schweregrad der Erkrankung (in diesem Fall der Pneumonie) und auch die Komorbidität der Patienten. Hier wird also ein Erklärungsmodell mit einem stetigen Zusammenhang zwischen Krankheitsschwere/Komorbidität einerseits und Krankenhausletalität andererseits formuliert. Höhere Letalitätswerte werden also mit einem höheren Anteil von Patienten mit »schwererer« Pneumonie und erhöhter Komorbidität begründet. Diesem klinisch nachvollziehbaren Erklärungsmodell kann erst einmal gefolgt werden, allerdings sollte dabei nicht verges-

sen werden, zu fragen, ob man in der betreffenden Klinik über Daten zur Krankheitsschwere beziehungsweise Komorbidität verfügt. Ohne solche Daten ist das obige Erklärungsmodell nur eine Hypothese, die nicht verifiziert und somit nicht als Grundlage der eingangs geforderten sachbezogenen Entscheidungsfindung herangezogen werden kann.

Denkbar wäre auch, dass alle Kliniken »links des Gruppenwertes« für sich das schwerste und komplexeste Patientengut in Anspruch nehmen, was wiederum zu der Frage führt, ob die in dieser Gruppe unterschiedlichen Werte (13,6% bis 20,7%) nur zufällige Variationen darstellen. Problematisch wird das Erklärungsmodell zusätzlich dadurch, dass nach unserer Erfahrung auch Kliniken »rechts des Gruppenwertes« ein schweres beziehungsweise komplexes Patientengut unterstellen.

Ein weiteres Ziel der Qualitätsbemühungen kann natürlich eine Verschiebung innerhalb der Bereiche »links und rechts vom Verbundwert« sein, die genaue Festlegung ist der Eigenverantwortung der Unternehmens- beziehungsweise Organisationsleitung zuzuordnen.

Der Gruppenwert 13,6% wird mittels der Verbunddaten berechnet, er ist also ein »selbstreferenzieller« Verbundwert, der über dem Wert der externen Qualitätssicherung (EQS) liegt, der im Auswertungsjahr 2011 mit 12,7% über alle Patienten (mit und ohne dokumentierter Therapieeinstellung) angegeben wird (EQS 2011). Es kann wegen der unterschiedlichen Datengrundlagen und Verfahren an dieser Stelle nicht entschieden werden, ob der Unterschied zwischen 12,7% und 13,6% ein Unterschied ist, der einen Unterschied ausmacht, und wir sehen daher auch den Wert der EQS nur als ein weiteres Hilfsmittel zur Qualitätsbewertung.

Der EQS-Wert ist mit der horizontalen Linie in der Grafik abgebildet und teilt die Verbundergebnisse nun in die Gruppen »oberhalb und unterhalb des EQS-Wertes« ein. Die möglichen Ausführungen bei der Diskussion der Daten unterscheiden sich nicht von den oben erwähnten.

Das Balkendiagramm ist folglich eine gute Möglichkeit zur Abbildung der Letalitätsrate sowie zum Vergleich mit Orientierungswerten (andere Krankenhäuser, Gruppenwert, Daten aus Literatur und von externen Verfahren).

Trotz der bereits erwähnten unterschiedlichen Datengrundlagen können wir aus dem hier vorgestellten wenig komplexen Balkendiagramm die folgenden Empfehlungen ableiten:

Krankenhäuser, deren Letalitätsrate über der von Orientierungswerten und/oder der eigenen Vorgabe liegt, sollten weitere Analysen zur Ursachenidentifizierung durchführen und diese verifizieren. Ziel sollte die Entscheidung sein, ob eine Verbesserung der Letalitätsrate als möglich erachtet und somit durch geeignete Maßnahmen angestrebt wird.

Liegt die Letalitätsrate unter den Orientierungswerten und/oder der eigenen Vorgabe, so sollte geprüft werden, ob eine weitere Verbesserung der Letalitätsrate als möglich erachtet und somit durch geeignete Maßnahmen angestrebt wird.

Exkurs: Analyse klinischer Prozesse

Ein in der Praxis gut anwendbares Konzept zur Analyse der Struktur-, Prozess- und Ergebnisqualität klinischer Prozesse ist die Analysepyramide (Mohammed et al. 2004; Duckett et al. 2007), die in **Abb. 3** vorgestellt wird.

Es ist gut nachvollziehbar, dass die unterste Ebene das Fundament der Qualitätsbewertung und somit der sachbezogenen Entscheidungsfindung bildet, da qualitativ oder quantitativ inadäquate Daten schnell zu Fehlinterpretationen führen können (**Abb. 4**). Daher werden die vorab formulierten Empfehlungen wie folgt ergänzt:

Unabhängig von der Letalitätsrate selbst und ihrer Lage zu einem Orientierungswert und/oder der eigenen Vorgabe sollte die Datenqualität immer sicher bekannt sein beziehungsweise bei Unsicherheit geprüft werden.

Die ebenfalls in **Abb. 4** abgebildete zweite Ebene fordert zur Analyse der Patienteneigenschaften auf, die ja wiederum in den Daten korrekt abgebildet werden müssen.

In unserem Beispiel der Pneumonie wäre bei den Daten und den Patienteneigenschaften besonderes Augenmerk auf die Parameter des in der EQS verwendeten Systems zur Einteilung der Patienten in Risikoklassen (sogenannter CRB-65-Score) zu lenken, die mittels Routinedaten sicher abbildbar sind (Alter und Beatmung). So weisen beide Krankenhäuser A und B den Anteil von Patienten mit einem Alter ≥ 65 Jahre mit rund 77% aus. Mindestens 25 Stunden beatmet wurden 4,4% (Krankenhaus A) beziehungsweise 6,3% (Krankenhaus B) der Patienten.

Auch die Fachliteratur bietet vielfältige Informationen hinsichtlich klinischer Risikofaktoren, die in Form von Nebendiagnosen kodierbar sind (zum Beispiel Fine et al. 1997; Bratzler et al. 2011).

Auf Ebene 3 geht es um Strukturen und Ressourcen, die sich auf die Ergebnisqualität auswirken können. Dies ist natürlich ein wichtiger Punkt und auch die DIN EN 15224 fordert, dass die Organisation die erforderlichen Ressourcen ermittelt und bereitstellt, die für die Erfüllung

der Anforderungen notwendig sind. Zu diesen Ressourcen zählen unter anderem die Infrastruktur und die Arbeitsumgebung.

Ebene 4 (Versorgungsprozess) und 5 (Personal) werden in **Abb. 5** aufgeführt. Der Versorgungsprozess läuft im Umfeld der Strukturen und Ressourcen ab, er wird von ihnen beeinflusst und ist daher immer in diesem Zusammenhang zu betrachten. Besondere Bedeutung erhält hierbei natürlich die Frage, ob der Versorgungsprozess wissenschaftlich abgesichert ist und sich auf das beste Wissen und die beste Praxis stützt, wie die DIN EN 15224 fordert. Wurde das beste Wissen – hier also die Leitlinie zur Pneumonie (Leitlinie Pneumonie 2009) – in die Organisation übernommen und kommt es wirksam am Patienten an? Funktioniert hier der »Transfer Research Into Practice (TRIP)«?

In den klinischen Prozessen sind Menschen tätig und daher sind die auf Ebene 5 formulierten Fragen für die Qualität eines Prozesses von elementarer Bedeutung und so fordert die DIN EN 15224 auch folgerichtig: »Personal, dessen Tätigkeit die Erfüllung der Anforderungen an die Dienstleistung in der Gesundheitsversorgung beeinflusst, muss aufgrund der angemessenen Ausbildung, Schulung und Fertigkeiten und Erfahrung kompetent sein.« Auch hier ist zu beachten, dass das Personal im Umfeld der Strukturen und Ressourcen sowie anderer organisationaler Umgebungsfaktoren in den klinischen Prozessen tätig ist.

Die bisherigen Ausführungen gelten auch für alle Grafiken, die nachfolgend besprochen werden, und natürlich nicht nur für die Letalitätsrate, sondern auch für andere Parameter (zum Beispiel nosokomiale Infektionen) und Patientengruppen.

Wir kommen an dieser Stelle zu zwei weiteren Empfehlungen:

Eine Analyse der Letalitätsrate sollte die Wechselwirkungen des klinischen Prozesses mit anderen (klinischen) Prozessen und auch alle möglichen Einflussfaktoren, die auf diese Prozesse wirken können, angemessen berücksichtigen.

Es wird ausdrücklich empfohlen, die Analyse der Letalitätsrate standardisiert mittels eines definierten Verfahrens durchzuführen, welches mit der klinischen Realität der Versorgungsprozesse kompatibel ist. Die Analysepyramide wird hierzu ebenfalls ausdrücklich empfohlen.

Zeitreihe

Eine weitere Darstellung der nicht risikoadjustierten Krankenhausletalität erfolgt über die sogenannte Zeitreihe.

Eine Zeitreihe ist eine zeitlich geordnete Folge von Daten (Beobachtungen). Auch wenn der Wortteil »Zeit« impliziert, dass eine zeitliche Folge abgebildet wird, so kann eine Folge von Daten oder Beobachtungen auch nach anderen Kriterien erfolgen, wie zum Beispiel eine Fallnummer von 1 bis x, die nach dem Entlassungszeitpunkt vergeben wird. Somit wäre diese Fallnummer dann auch wieder als zeitliches Kriterium anzusehen.

Da die Zeitreihe in unserem Beispiel monatlich ergänzt wird, dient sie der Bewertung des aktuellen Monats oder auch eines Zeitraumes auf Basis historischer Daten, also der Daten aus dem Zeitraum davor.

In **Abb. 6** werden die Daten zur nicht risikoadjustierten Krankenhausletalität des Krankenhauses A für die Jahre 2010 bis 2012 mit insgesamt 1.295 Fällen abgebildet.

Neben den monatlichen Krankenhausletalitätsraten wird eine sogenannte Korrelationsgerade eingezeichnet. Hierunter versteht man eine Gerade, die sich den einzelnen Punkten möglichst gut »anpasst« beziehungsweise zwischen den Punkten möglichst gut ausgleicht, also die beste Angleichung an die Datenpunkte darstellt. Die Korrelationsgerade zeigt an, ob die Daten im Laufe der Zeit einen bestimmten Trend annehmen, daher nennen wir sie auch »Trendlinie«.

Im vorliegenden Beispiel ist zu sehen, dass die Trendlinie im Januar 2010 bei rund 16% startet und im Dezember 2012 bei einem Wert von rund 12% endet. Hieraus ist also zunächst abzulesen, dass die nicht risikoadjustierte Krankenhausletalität im Laufe der Zeit abgenommen hat. Dies entspricht auch den einzelnen hier nicht gezeigten Jahresgesamtwerten, die für das Krankenhaus A 15,2%; 15,9% und 12,0% betragen (Jahre 2010/2011/2012). An diesen drei Werten ist sehr gut zu erkennen, dass die Trendlinie den Trend über den gesamten Zeitraum abbildet und insgesamt sinkt, während die Jahresgesamtwerte von 2010 auf 2011 zunächst sogar ansteigen, um dann von 2011 auf 2012 deutlich abzufallen.

Wir erkennen weiterhin, dass die einzelnen Datenpunkte mitunter deutlich von der Trendlinie abweichen. Um interpretieren zu können, wie gut die Trendlinie an die einzelnen Datenpunkte angepasst ist, oder mit anderen Worten: wie verlässlich die Tendaussage ist, ist der angegebene Korrelationsfaktor (r) hilfreich. Für den Korrelationsfaktor r gilt (in Anlehnung an Hüsler & Zimmermann 2006, Seite 192; Muche et al. 2005, Seite 55):

- Er ist ein Wert zwischen -1 und $+1$.
- Ein positiver Wert zeigt eine steigende und ein negativer Wert eine abfallende Trendlinie an.
- Ein Wert gleich Null bedeutet, dass die Trendlinie keine verwertbare Anpassung aufweist.
- Je stärker die Anpassung, desto näher liegt der r -Wert bei $+1$ (steigende Trendlinie) oder -1 (abfallende Trendlinie).
- Eine gute Anpassung der Trendlinie liegt bei einem r -Wert $\geq 0,7$ vor.

Die vorliegende Trendlinie zeigt also mit einem Wert für $r = -0,2$ keine gute Anpassung, was bei den stark schwankenden Werten nicht verwundert.

Es sind natürlich die schwankenden Werte, die eine Qualitätsbewertung so schwierig gestalten. Würde die Trendlinie eindeutig sinken oder ansteigen und eine gute Anpassung der Trendlinie vorliegen (also: $r \geq 0,7$), so wäre die Interpretation viel einfacher. In der Realität haben wir es jedoch meist mit nicht so eindeutigen Datenverläufen zu tun.

Die grundsätzliche Frage, die bei der Betrachtung solcher Verläufe gestellt wird, lautet: »Handelt es sich bei den Schwankungen um zufällige Variationen oder liegen nicht zufallsbedingte Variationen vor, die aus medizinischer Qualitätssicht positiv oder negativ zu bewerten sind?«

Zunächst mag der Betrachter geneigt sein, den Durchschnittswert der EQS (12,7%, siehe voriger Abschnitt) aus dem Jahr 2012 für den entsprechenden Abschnitt in **Abb. 6** heranzuziehen. Besonders hilfreich wäre dies nicht, denn auch hier lägen die Datenpunkte recht variabel um den Orientierungswert. Zu beachten ist hierbei auch, dass es sich bei dem Orientierungswert um den Ganzjahreswert über alle in der EQS unter dieser Diagnose ausgewerteten Fälle und Krankenhäuser handelt. Wesentlich hilfreicher wären daher die uns nicht zur Verfügung stehenden monatlichen Werte, die eventuell ebenfalls variabel ausfallen.

Die zu erkennende Variation der Letalitätsraten kann auch durch abweichende Fallzahlen beeinflusst werden. In **Abb. 7** werden für die markierten Extremwerte die jeweiligen Fallzahlen (verstorben/gesamt) angegeben. Bei im Median 5 von 35 verstorbenen Patienten im Zeitraum 2010 bis 2012 stellt man fest: Die Extremwerte unterhalb der Trendlinie zeichnen sich durch eine im Vergleich zum Median geringe Anzahl verstorbenen Patienten aus, im Monat 2011-01 liegt die Fallzahl zusätzlich deutlich über dem Median von 35.

Die Extremwerte über der Trendlinie zeigen ein umgekehrtes Bild mit geringeren Fallzahlen und in 2 Monaten (2011-05 und 2011-07) ist auch die Anzahl der Verstorbenen über dem Median von 5.

Natürlich ist es sinnvoll, die Frage nach den Fallzahlen zu stellen. Es soll aber davor gewarnt werden, schwankende Fallzahlen alleine schon als ein ausreichendes Erklärungsmodell zu sehen. Hier gilt die Warnung, dass Zahlen hochgradige Komplexität in unzulässiger Weise reduzieren können!

Eine weitere Hilfe bei der Beantwortung der Frage »Handelt es sich bei den Schwankungen um Unterschiede, die einen Unterschied ausmachen?« bietet die **Abb. 8**. Hier werden zusätzlich die monatlichen Werte über alle Fälle des Verbundes abgebildet (insgesamt 24.689). Zusätzlich wird für die Verbundwerte auch das 95%-Konfidenzintervall (95%-KI, auch genannt: Vertrauensintervall, Altman et al. 2000; Bender & Lange 2007; du Prel et al. 2009) berechnet. Das 95%-KI unterstützt die Beantwortung der Frage insofern, als dass man folgende Faustregel anwenden kann: Liegt ein Wert meines Krankenhauses außerhalb des 95%-KI der Verbundwerte, so handelt es sich in der Regel um einen statistisch bedeutsamen Unterschied.

Hierbei ist zu beachten, dass dieser Unterschied durch den Vergleich der Krankenhauswerte mit den Verbundwerten gebildet wird. Eine andere Bezugsgröße kann zu einem anderen Ergebnis führen, wie in den nächsten Abschnitten zu sehen sein wird.

Es sollte auch Sensibilisierung für Serien bestehen, die wir hier als mindestens drei Datenpunkte oberhalb beziehungsweise unterhalb des 95%-KI definieren.

Für das Krankenhaus A wird wie folgt formuliert: Die Trendlinie sinkt im Verlauf von 2010 bis 2012 ab und liegt dabei fast vollständig innerhalb des 95%-KI der Verbunddaten. Einzelne Datenpunkte beziehungsweise mehrere Datenpunkte liegen außerhalb des 95%-KI, sie sollten näher analysiert werden.

Die Daten des Krankenhauses B zeigt **Abb. 9**, hier fallen die Extremwerte in den Monaten 2012-06 und 2012-11 sofort ins Auge. Die Fallzahlen in diesen Monaten weichen deutlich von den Medianwerten der Verstorbenen (5) beziehungsweise der Fallzahl (35) der Verbundwerte ab: Sie betragen im Juni 11/20 und im November 9/15.

Besonders auffällig ist eine Serie von Datenpunkten, die an der Obergrenze beziehungsweise oberhalb des 95%-KI liegen (2011-09 bis 2012-04, in der Abbildung farbig markiert). Die Trendlinie zeigt bei geringer Anpassung einen Anstieg von rund 13% auf rund 21% ($r = 0,2$).

Wir sehen hier also eine steigende Trendlinie, die in weiten Bereichen über dem 95%-KI der Verbunddaten liegt. In Verbindung mit den beschriebenen Extremwerten und

der Serie solcher Werte ist eine weitere Analyse dringend angeraten. Eine solche Analyse geht natürlich über eine Betrachtung der Fallzahlen hinaus, wenn sie der Komplexität der hinter den Daten stehenden klinischen Prozesse gerecht werden soll.

Folgerichtig kann auch hier die Anwendung der Analysepyramide weiterhelfen.

Wir kommen an dieser Stelle zu weiteren Empfehlungen:

Ein regelmäßiges Monitoring der Letalitätsraten mithilfe von Zeitreihen sollte durchgeführt werden.

Die in der Zeitreihe abgebildeten Daten eines Krankenhauses sollten um eine Trendlinie ergänzt werden, die den gesamten Beobachtungszeitraum und/oder einzelne Zeitabschnitte (zum Beispiel Jahre) umfasst.

Die zu den einzelnen Datenpunkten gehörenden Fallzahlen sollten verfügbar sein, um deren Einfluss auf die Datenpunkte einschätzen zu können. Daten einer Vergleichsgruppe inklusive der dazu gehörenden Fallzahlen und des Konfidenzintervalls sollten in der Zeitreihe ebenfalls abgebildet werden.

Die Vergleichsdaten sollten in der Auflösung der Krankenhausdaten, möglichst pro Monat, abgebildet werden.

Krankenhausdaten, die außerhalb des Konfidenzintervalls der Vergleichsgruppe liegen, sollten standardisiert einer näheren Analyse unter Anwendung eines definierten Verfahrens unterzogen werden. Dies gilt natürlich auch dann, wenn die Krankenhausdaten unterhalb des Konfidenzintervalls liegen – Darstellung guter Praxis.

Die Empfehlungen für eine nähere Datenanalyse gelten insbesondere dann, wenn abweichende Daten in Serie vorkommen.

Regelkarte

Wie bereits bei der Zeitreihe zu sehen war, hilft die kontinuierliche Darstellung und Interpretation von Daten (hier: nicht adjustierte Krankenhausletalitätsraten) bei der Qualitätsbewertung, die ja zu angemessenen Managemententscheidungen führen soll.

Zu diesem Zweck können auch Regelkarten eingesetzt werden, deren Charme darin besteht, dass sie auch ohne Statistikprogramme berechnet und dargestellt werden können. Für die Regelkarten gilt, wie für die Zeitreihen, dass sie auch risikoadjustierte Daten abbilden können.

In der Literatur findet man eine Vielzahl von Arbeiten und Büchern, die Grundlagen und Anwendungsbeispiele im klinischen Umfeld beschreiben, so zum Beispiel Cook et al. (2008), Duclos et al. (2009), Hart et al. (2003), Henderson et al. (2008), Kottner & Hauss (2013), Mohammed et al. (2008), Mohammed & Worthington (2013), Noyez (2009), Perla et al. (2011), Poelaert et al. (2007), Tennant et al. 2007, Winkel & Zhang (2007) und Woodall (2006). Besondere Erwähnung soll hier die Arbeit von Mohammed et al. aus dem Jahr 2013 finden, in der die Autoren Probleme und Lösungen aufzeigen, die bei der Anwendung von Regelkarten bei sehr großen Datenmengen relevant sind. Da solche Probleme gerade bei der Anwendung von Routinedaten entstehen können, wird diese Arbeit besonders hervorgehoben.

Zu den Grundlagen von Regelkarten führen Kottner & Hauss (2013) in einer Übersichtsarbeit aus: Alle empirisch ermittelten Daten werden von systematischen und zufälligen Fehlern beeinflusst. Eine Möglichkeit, mit zufälliger Variation bei Datenvergleichen adäquat umzugehen, bietet die Theorie der Statistischen Prozesskontrolle (SPC). Innerhalb dieser Theorie wird angenommen, dass es in nahezu allen Prozessen, Produkten und Ergebnissen Variationen gibt. Keine zwei erbrachten Leistungen und keine zwei Produkte sind hundertprozentig identisch. Die Fragen, die die SPC versucht, zu beantworten, lauten: »Wie viel Variation ist natürlicherweise innerhalb stabiler Prozesse zu erwarten (sogenannte common cause variation)? Wann weisen Prozesse und Ergebnisse höhere Variationen auf, die, basierend auf den vorliegenden Daten, nicht mehr natürlicherweise erklärt werden können (special cause variation)?

Winkel & Zhang (2007, Seite 12) formulieren sehr treffend: »Statistical control is a concept fundamental to the theory of control charts. It is based on a distinction between two types of variation: one resulting from unavoidable causes, which one cannot identify (random variation), and one resulting from causes, which may be identified (assignable causes of variation). A process in which sample values vary due to random causes alone is said to be in a state of statistical control. Additional variation caused by assignable causes may occur. If this is the case, the process is said to be out of statistical control. Since the causes may be identified, it is often possible to regulate and control them so that the process may be brought back into a state of statistical control.«

Eine auf Basis der Daten nicht mehr natürlicherweise erklärbare Variation (auch genannt: »Signal«) könnte beispielsweise durch Abweichungen von den Empfehlungen der Leitlinie zur Pneumonie (Leitlinie Pneumonie 2009) bedingt sein.

Der englische Begriff »control chart« wurde in den 1920er-Jahren von Walter A. Shewart geprägt und wird im Deutschen mit »Regelkarte« übersetzt. Es handelt sich um den Vergleich von Daten über eine Zeitspanne und so

mit um einen »Selbstvergleich« (in Anlehnung an Kottner & Hauss 2013). Es findet also kein Vergleich mit Daten anderer Krankenhäuser statt.

In Abhängigkeit von der Art der betrachteten Merkmale kommen unterschiedliche Regelkarten zur Anwendung, in dieser Arbeit wird die sogenannte p-Regelkarte näher erläutert, die in **Abb. 10** zu sehen ist.

Aufgetragen sind die bereits bekannten monatlichen Letalitätsraten des Krankenhauses A und der arithmetische Mittelwert über alle Datenpunkte. Zusätzlich werden sogenannte obere und untere Warngrenzen sowie die oberen und unteren Kontrollgrenzen eingezeichnet. Per Konvention befinden sich die Warngrenzen ± 2 Standardabweichungen und die Kontrollgrenzen ± 3 Standardabweichungen vom arithmetischen Mittelwert, eine genaue Beschreibung der Berechnungen ist bei Kottner & Hauss (2013) zu finden.

Bei einem stabilen Prozess befinden sich nahezu alle Datenpunkte innerhalb der Kontrollgrenzen, die sich als sinnvoll erwiesen haben, da bei einem stabilen Prozess aus statistischer Sicht nur 0,27% der Datenpunkte fälschlicherweise außerhalb der Kontrollgrenzen liegen. Dies bedeutet, dass nur einer von 370 Datenpunkten fälschlicherweise als »Signal« klassifiziert würde.

Die Daten des Krankenhauses A sprechen also für einen aus statistischer Sicht stabilen Prozess, im zeitlichen Verlauf tangiert kein Datenpunkt die Warn- oder Kontrollgrenze. Selbst die im vorherigen Abschnitt als Extremwerte bezeichneten Datenpunkte (farbig markiert) gelten nun als nicht auffällig.

An dieser Stelle wird die Bedeutung der Vergleichsdaten deutlich: Bei der Zeitreihe haben wir die Verbundwerte benutzt, hier nun den arithmetischen Mittelwert (relative Häufigkeit) des Krankenhauses selbst (also: selbstreferenziell) und um genau diesen Mittelwert werden die Warn- und Kontrollgrenzen ja berechnet.

Es gibt für die Interpretation noch weitere Regeln, die nach Mohammed et al. (2008) ebenfalls als Signale betrachtet werden, die auf nicht zufällige Variation im Prozess hinweisen:

- acht (manche Autoren favorisieren sieben) zusammenhängende Datenpunkte auf einer Seite der Mittellinie (Verschiebung)
- zwei von drei aufeinanderfolgenden Datenpunkten, die sich außerhalb einer Warngrenze befinden
- einen »run« von acht (manche Autoren favorisieren sieben) aufeinanderfolgenden Datenpunkten, die kontinuierlich an- oder absteigen (Trend)

Keine dieser Bedingungen wird von Krankenhaus A in negativer Ausprägung erfüllt, dies ist ein anzustrebendes Ergebnis.

Der Prozess mag zwar »unter Kontrolle« sein, aber es sollte in jedem Fall auch geprüft werden, wie die Daten im Vergleich zu einer repräsentativen Vergleichsgruppe liegen. Aus unserer Sicht ist die kombinierte Anwendung der von uns vorgestellten Zeitreihe mit der Regelkarte anzustreben.

Weiterhin gilt: Auch ein aus statistischer Sicht als »kontrolliert« einzustufender Prozess sollte zusätzlich immer aus klinischer Sicht hinterfragt werden. Dies gewährleistet, dass wir deutliche Datenschwankungen auch dann als Hinweis auf gute oder weniger gute Praxis erkennen können, wenn sie sich innerhalb der Warn- oder Kontrollgrenzen befinden.

Die Daten von Krankenhaus B werden in **Abb. 11** gezeigt: Die bereits bei der Zeitreihe markierten Extremwerte sind eindeutige Signale, da sie die obere Kontrollgrenze überschreiten. Die bei der Zeitreihe auffällige Serie von Datenpunkten (2011-09 bis 2012-04) stellt in der Regelkarte jedoch kein Signal dar. Auch hier zeigt sich die Überlegenheit der Kombination beider Verfahren, da wir die Zeitreihe zum Anlass nehmen würden, auch diesen Bereich weiter zu analysieren.

Zwei weitere Signale, die bereits bei der Zeitreihe aufgefallen waren, sind bei den Monaten 2010-04 und 2011-08 zu sehen – hier wäre interessant herauszufinden, welche Faktoren aus der Analysepyramide diese Signale im Sinne bester Praxis erklären können. Was war in diesen beiden Monaten und auch in den Monaten 2012-06 sowie 2012-11 anders? Was hat sich geändert? Was wurde in den anderen Monaten anders gemacht? Was davon sollten wir vermehrt tun, was sollten wir möglichst nicht mehr tun?

Die Regelkarte bietet eine gut handhabbare und zugängliche Methode für das regelmäßige Datenmonitoring. Da sie keine externen Vergleichsdaten benötigt, eignet sie sich sehr gut auch für Krankenhäuser, die nicht in einem Verbund organisiert sind. Berechnung und grafische Darstellung erfordern kein Spezialwissen und auch keine Statistiksoftware. Da die Grundlagen in der Literatur gut beschrieben sind (Kottner & Hauss 2013), können Regelkarten in der Praxis mit vertretbarem Aufwand erstellt werden.

Empfehlungen:

Ein regelmäßiges Monitoring der einrichtungsbezogenen Letalitätsraten mithilfe von Regelkarten sollte durchgeführt werden.

Die Festlegung, was ein Signal ist und was nicht, sollte definierten Regeln folgen.

Signale sollten standardisiert einer näheren Analyse unter Anwendung eines definierten Verfahrens unterzogen werden.

Auch ein aus statistischer Sicht »kontrollierter Prozess« sollte aus klinischer Sicht hinterfragt werden. Dies gilt insbesondere dann, wenn Datenschwankungen erkannt werden, die klinisch nicht nachvollziehbar oder – im Hinblick auf eventuell definierte Sollwerte – nicht akzeptabel sind. Verwechsle nie statistische Signifikanz mit klinischer Bedeutung.

Die kombinierte Anwendung von Regelkarte und Zeitreihe wird empfohlen.

Die (klinische) Leitungsaufgabe sollte verstanden werden als Sicherstellung einer angemessenen Variabilität von Prozessen.

Cumulative Sum (CUSUM)

Die »Cumulative Sum (CUSUM)«-Technik wurde bereits vor nunmehr 60 Jahren in England im Rahmen industrieller Qualitätskontrolle eingeführt und erstmals 1994 zum Qualitätsmonitoring kardiochirurgischer Operationen eingesetzt (Grunkemeier et al. 2003).

Noyez (2009) gibt in seinem Review eine Übersicht zu Methoden, die zum »Performancemonitoring« im Gesundheitswesen angewendet werden können. Er gibt hierbei auch einen Überblick zu den verschiedenen CUSUM-Typen:

- Cumulative failure chart
- Standard non-risk-adjusted CUSUM chart
- Risk-adjusted CUSUM chart
(auch genannt: CRAM, VLAD)

Die »Cumulative failure chart« ist die einfachste Form des CUSUM, sie bildet die kumulative Summe von Ereignissen (zum Beispiel Tod, Nachblutung) ab.

Die »Standard non-risk-adjusted CUSUM chart« wird in Zusammenhang mit konstanten Erwartungswerten (zum Beispiel einer Komplikationsrate) eingesetzt. Dies bedeutet, dass ein individuell für jeden Patienten berechneter Erwartungswert (zum Beispiel für die Krankenhaussterblichkeit) nicht vorliegt. Bei vielen Themen, die einer Qualitätsmessung und Qualitätsbewertung unterzogen werden sollen, ist dies jedoch der Fall und daher erfüllt diese CUSUM hier eine wichtige Aufgabe.

Die »Risk adjusted CUSUM chart« bildet kumulative Summen erwarteter minus beobachteter Sterblichkeit (oder auch bestimmter Komplikationen) grafisch ab. An dieser Stelle soll erwähnt werden, dass die Literatur bezüglich der Bezeichnung dieser Grafik durchaus als verwirrend bezeichnet werden kann, da verschiedene Autoren neben »Risk adjusted CUSUM chart« unterschiedliche Bezeichnungen verwenden: Poloniecki et al. (1998) nennen die Grafik »Cumulative Risk-Adjusted Mortality (CRAM)«. Lovegrove et al. (1997) berichten zu einer Weiterentwicklung der kumulativen Summenmethode und führen die Bezeichnung »Variable Life Adjusted Display (VLAD)« ein (zitiert nach Grunkemeier et al. 2003; Noyez 2009).

Zur eindeutigen sprachlichen Abgrenzung wird in dieser Arbeit für die Abbildung kumulierter Summen risikoadjustierter Daten die Bezeichnung VLAD verwendet.

An dieser Stelle wird zunächst CUSUM erläutert, zu VLAD finden sich dann Ausführungen in einem späteren Abschnitt.

Wird ein konstanter Erwartungswert für die Anwendung von CUSUM bei einer bestimmten Fragestellung gesucht, so gibt es viele Quellen, in denen sich eine Recherche lohnen kann, so zum Beispiel:

- externe Qualitätssicherung
- Register von Fachgesellschaften
- Krankenhausverbände
- Krankenhauszweckverbände
- Fachliteratur
- Leitlinien
- Krankenhausdaten

In unserem Beispiel wird als konstanter Erwartungswert die Krankenhaussterblichkeit verwendet, die in den Ergebnissen der externen Qualitätssicherung aus dem Jahr 2011 angegeben wird (12,7%).

Die Berechnung der Daten für die in **Abb. 12** gezeigte CUSUM-Grafik wird in **Tab. 1** näher beschrieben:

Die 432 Patienten werden chronologisch nach ihrem Entlassungsdatum aufsteigend nummeriert (Spalte: Patient). Die erwartete Sterblichkeit (0,127) wird für jeden Patienten in Spalte E mit dem jeweiligen beobachteten Ergebnis (überlebt: 0/verstorben: 1, Spalte: B) aufgetragen.

Sodann wird für jeden Patienten die Differenz E – B berechnet. Hierbei wird die folgende Logik verfolgt:

Patient 1 ist nicht verstorben, daher wurden 0,127 (0,127 – 0) »Leben gerettet«, oder mit anderen Worten: »Es sind 0,127 Patienten weniger verstorben als erwartet«.

Patient 5 ist verstorben, daher wurden 0,873 (0,127 – 1) »Leben verloren«, oder mit anderen Worten: »Es sind 0,873 Patienten mehr verstorben als erwartet«.

In der Spalte CUSUM werden dann die einzelnen Differenzen (Spalte: E – B) kumulativ addiert: So ergibt die Addition der Werte von Patient 1 (0,127) + Patient 2 (0,127) den CUSUM-Wert 0,254 bis zum Ergebnis bei Patient 432 mit 2,864.

Betrachtet man die Werte, so wird deutlich, dass der konstante Erwartungswert dazu führt, dass ein »gewonnenes Leben« mathematisch mit 0,127 in den CUSUM eingeht, während ein »verlorenes Leben« zu einem Abzug von 0,873 führt.

Dies bedeutet also: CUSUM wird wegen des konstanten Erwartungswerts durch verstorbene Patienten deutlich stärker beeinflusst als durch überlebende Patienten. Auf diese Eigenschaft wird bei den Ausführungen zum VLAD noch eingegangen.

Im Gesamtergebnis sind im Krankenhaus A im Jahr 2012 rund 3 Patienten weniger verstorben, als erwartet wurde. Dies ist kein überraschendes Ergebnis, da in **Abb. 2** zu erkennen war, dass der Wert von Krankenhaus A mit 12,0% unterhalb des Orientierungswertes der externen Qualitätssicherung liegt.

Schon an dieser Stelle kann man erkennen, dass die CUSUM-Grafik einen guten Beitrag leistet, wenn man »statistische Signifikanz« in »klinische Bedeutung« übersetzen möchte. Der besondere Vorteil ist bei dieser Technik, dass jeder Datenpunkt der Linie einem Patienten entspricht und daher auffällige CUSUM-Abschnitte den betreffenden Patienten und so auch Zeiträumen zugeordnet werden können. Solche Zeiträume sind in **Abb. 12** auf der x-Achse erkennbar, sie ist mit den Monatsnamen beschriftet und zeigt, dass in den einzelnen Zeiträumen unterschiedliche Fallzahlen vorliegen.

Zur Interpretation einzelner Abschnitte des CUSUM wird auf die entsprechenden Erklärungen zum VLAD in dieser Arbeit verwiesen.

In **Abb. 12** sind nun zusätzlich zwei Limits zu sehen, ein oberes und ein unteres. Die Berechnung der Limits wird von Noyez (2009) erläutert, daher wird hier auf eine detaillierte Beschreibung verzichtet. Nur zu den Orientierungsgrößen für die Berechnung der Limits soll hier ausgeführt werden:

Die von Noyez (2009) vorgestellte Berechnung der Limits erfordert die Festlegung einer »accepted failure rate« und einer »unaccepted failure rate«. Diese Werte definieren also den Bereich, innerhalb dessen man von akzeptierter Qualität nach den Vorgaben der (klinischen) Leitung

eines Krankenhauses oder einer Fachabteilung sprechen würde. Dieser Bereich würde dann im Transformations-schritt »Datum zu Information« (**Abb. 1**) angewendet werden. An dieser Stelle wird deutlich, dass die Festlegung von Zielwerten beziehungsweise Zielbereichen als Grundlage der Qualitätsbewertung eine wesentliche Aufgabe (klinischer) Leitungskräfte ist.

Als »accepted failure rate« und »unaccepted failure rate« wurden in **Abb. 12** 12,7% beziehungsweise 19,1% festgelegt, wobei die 19,1% einer 1,5-fachen Erhöhung der »accepted failure rate« entsprechen.

Im Zusatzmaterial zu dieser Arbeit befindet sich eine Datei, in der ein CUSUM berechnet wird.

Die Limits dienen zur Beantwortung der Frage, ob der Kurvenverlauf (zum Beispiel Kurvenschwankungen/kontinuierlich steigender oder fallender Verlauf über einen bestimmten Zeitraum) einer »normalen Variation« entspricht oder Hinweis auf eine positive beziehungsweise negative Auffälligkeit (auch genannt: Signal) ist.

Für Krankenhaus A ist **Abb. 12** zu entnehmen, dass der CUSUM einen schwankenden Verlauf zeigt und die Limits überschritten werden (März und Oktober).

Der CUSUM von Krankenhaus B (**Abb. 13**) zeigt ebenfalls Phasen schwankender Verläufe. Der wichtige Unterschied zu Krankenhaus A besteht in der Tatsache, dass am Ende des Jahres der CUSUM negativ abschließt (-19) und im Jahresverlauf mehrfach das untere Limit überschreitet. Somit wäre Anfang März das erste Signal aufgetreten und spätestens mit dem Signal Ende Mai hätte die Frage nach den Ursachen und einer möglichen Intervention gestellt werden können.

Von den einzelnen Datenpunkten kann jederzeit der Bezug zu dem dahinter stehenden Fall hergestellt werden, dies ist eine der wichtigsten Eigenschaften der CUSUM-Technik. Weitere Ausführungen dazu sind später beim Thema VLAD zu finden.

An dieser Stelle kann zusammenfassend gesagt werden, dass die CUSUM-Technik die Krankenhausletalität als Zielgröße unserer Betrachtungen nicht in einem Prozentwert, sondern als zu einem bestimmten Zeitpunkt ge-

wonnene oder verlorene Leben seit dem gewählten Anfangszeitpunkt (hier: Jahresbeginn) darstellt. Somit kann CUSUM auch sehr plastisch die Auswirkungen von Interventionen aufzeigen.

Empfehlungen:

Ein regelmäßiges Monitoring der beobachteten und erwarteten Letalitätsraten mittels CUSUM sollte durchgeführt werden.

Jedes Signal sollte eine weiterführende Analyse auslösen.

Auch Trends ohne Signal sollten einer weiterführenden Analyse unterzogen werden.

Risikoadjustierte Krankenhausletalität

Balkendiagramm (beobachtet versus erwartet)

In **Abb. 2** wurde die nicht risikoadjustierte Krankenhausletalität grafisch dargestellt, man spricht hier auch von der »beobachteten« Krankenhausletalität (BKHL). Die beobachteten Raten werden in **Abb. 14** ebenfalls dargestellt (Balken) und um die für das jeweilige Krankenhaus berechnete erwartete Rate ergänzt. Die »erwartete« Rate (EKHL) ist als Raute mit dem dazugehörenden 95%-KI angegeben. Auch diese Auswertung wird unseren Krankenhäusern in monatlichen Abständen zur Verfügung gestellt.

Die markierten Krankenhäuser (x) haben für die Berechnung der erwarteten Werte zu geringe Fallzahlen und daher sollten für sie keine Qualitätsbewertungen unter Berücksichtigung der erwarteten Werte durchgeführt werden (Kriterien nach Ash et al. 2003, Seite 305/312).

Zu den beobachteten Werten der Krankenhäuser A und B wurde bereits oben ausgeführt.

Für jeden Fall wird der Erwartungswert für die Krankenhaussterblichkeit berechnet. Dies erfolgt mittels eines statistischen Modells, welches mit der eingangs erwähnten multiplen logistischen Regression berechnet wurde. Die für einen Behandlungsfall individuell berechnete erwartete Krankenhaussterblichkeit steigt mit der Anzahl der sogenannten Risikofaktoren, die bei dem Patienten vorliegen. Dies sind zum Beispiel: Hauptdiagnose Pneumonie durch Pseudomonas oder Aspirationspneumonie, höheres Alter, gleichzeitig bestehende maligne Erkrankung, Zustand nach ischämischem Hirninfarkt, Linksherzinsuffizienz mit Ruhebeschwerden, chronische Niereninsuffizienz im Stadium IV.

Es gilt: Je älter der Patient und je mehr Risikofaktoren er aufweist, umso höher ist die berechnete erwartete Krankenhaussterblichkeit.

Zu den monatlichen Auswertungen werden dann die Erwartungswerte aller abgeschlossenen vollstationären Pneumoniefälle eines Krankenhauses addiert und dieser Wert wird dann in der Grafik als Raute mit 95%-KI abgebildet.

Auch wenn in der Abbildung alle Krankenhäuser berücksichtigt sind, geht es hier nicht um einen Einrichtungsvergleich im Sinne der Frage: Wer hat die beste Qualität? Diese Darstellungsform ermöglicht den Vergleich des beobachteten und erwarteten Wertes eines Krankenhauses, es handelt sich also wieder, bezogen auf das einzelne Krankenhaus, um ein selbstreferenzielles Vorgehen.

Betrachtet man die Erwartungswerte der verschiedenen Krankenhäuser, so kann es allenfalls um die Frage gehen, warum der Erwartungswert einer Abteilung (deutlich) von dem einer anderen Abteilung abweicht, obwohl der Betrachter beide Abteilungen kennt und davon ausgeht, dass sich die jeweilige Risikostruktur der Pneumoniepatienten beider Abteilungen nicht oder kaum voneinander unterscheiden. Dies ist der Vorteil, wenn man sich in einer Verbundstruktur befindet, in der man Informationen auch zur Häufigkeit der Risikofaktoren bei den Pneumoniepatienten in den einzelnen Krankenhäusern erlangen kann. Dies ist in unserem Verbund gewährleistet, da auch diese Daten monatlich zur Verfügung gestellt werden.

Der Erwartungswert basiert auf den Risikofaktoren und diese wiederum werden, mit Ausnahme des Alters, über Nebendiagnosen kodiert. Somit besteht also die Sequenz »Dokumentation – Kodierung – Erwartungswert«. Bei Zweifeln an der Güte der Daten gilt es, die Fragen aus Ebene 1 und 2 der Analysepyramide zu stellen (**Abb. 4**).

Durch die für die Erwartungswerte angegebenen Konfidenzintervalle sind Rückschlüsse auf die statistische Signifikanz möglich. Enthält ein Vertrauensbereich den beobachteten Wert (Balken) nicht, so kann man in der Regel von einem »statistisch signifikanten« Unterschied ausgehen.

Bei Krankenhaus A ist dies der Fall ($BKHL < EKHL$) und wenn weniger Pneumoniepatienten als erwartet versterben, so wird dieses Ergebnis auch als klinisch bedeutsam einzustufen sein. Damit ist natürlich noch nicht die Frage beantwortet, was die Ursachen dieses Ergebnisses sind

(zum Beispiel Überkodierung und dadurch falsch hohe erwartete Sterblichkeit? Diagnostik und Therapie in Übereinstimmung mit der Pneumonieleitlinie?).

Die Weite des Vertrauensbereichs hängt von Stichprobengröße und Standardabweichung der untersuchten Gruppen ab. Eine große Stichprobe führt zu »mehr Vertrauen«, also zu einem engen Konfidenzintervall. Ein breites Konfidenzintervall kann von einer kleinen Stichprobe herrühren (siehe hierzu die mit dem »x« markierten Krankenhäuser). Bei großer Streuung der Werte wird die Aussage unsicherer, das heißt, das Konfidenzintervall wird breiter. Diese Streuung kann natürlich auch auf eine hohe Variabilität des Kodierungsprozesses hinweisen und eine Prüfung sollte auf Ebene 1 und 2 der Analysepyramide stattfinden.

Der Vorteil von Konfidenzintervallen ist, dass Ergebnisse direkt auf der Ebene der Datenmessung angegeben werden. Konfidenzintervalle geben Informationen sowohl über die statistische Signifikanz als auch über die Richtung und Größe des Effekts. Damit kann man auch über die klinische Relevanz der Ergebnisse entscheiden. In die Breite des Vertrauensbereichs bei vorgegebener Irrtumswahrscheinlichkeit gehen zudem die Variabilität der Daten und die Fallzahl der untersuchten Stichprobe ein. Zur weiteren Lektüre zum Thema Konfidenzintervall seien an dieser Stelle die Übersichtsarbeiten von Bender & Lange (2007) und du Prel et al. (2009) empfohlen, aus denen hier zitiert wurde.

Bei Krankenhaus B sehen wir $BKHL > EKHL$ und hier wäre ebenfalls die Frage nach der Kodierung beziehungsweise dem Behandlungsprozess zu stellen: Unterkodierung der Risikofaktoren und dadurch zu geringer erwarteter Wert? Behandlung nicht in Übereinstimmung mit der Pneumonieleitlinie?

Bei Krankenhaus C ist zu sehen: $BKHL < EKHL$, jedoch ragt $BKHL$ in das untere Konfidenzintervall hinein. Hier wäre eine weitere Analyse ausdrücklich anzuraten. Bei Krankenhaus D wird das untere Konfidenzintervall marginal tangiert und bei Krankenhaus E wurde es »knapp verfehlt«. Dem (klinischen) Management beider Krankenhäuser würden wir eine Analyse empfehlen.

Die vorgestellte Grafik bietet weitere Ansätze, die bei der Qualitätsbewertung hilfreich sein können. Aus unserer Erfahrung reicht dies alleine natürlich nicht, da es immer darauf ankommt, wer mit welchem Erfahrungs- und Erwartungshintergrund diese Grafik betrachtet. Hierzu in der Praxis oft zu hörende Aussagen:

- Gilt » $BKHL \leq EKHL$ «, so handelt es sich um ein Ergebnis, welches exakt den Erwartungen der (leitenden) Mitarbeiter entspricht und selbstverständlich einer Handlungsweise gemäß bester Evidenz und klinischer Praxis entspricht.
- Liegt als Ergebnis » $BKHL > EKHL$ « vor, so könnte dies allein an der Kodierung oder insuffizienten und realitätsfernen statistischen Verfahren liegen. Man hat zwar die schwersten Fälle, kann sie aber mit den vorhandenen Mitteln in den Daten nicht korrekt abbilden.

Wichtig ist, dass der Zielgruppe einer Auswertung die folgenden Punkte vermittelt werden, dies gilt besonders für Anwender mit limitiertem statistischem Wissen: Statistische Verfahren verfolgen nicht den Anspruch, eine exakte Wahrheit abzubilden. Das geht schon deshalb nicht, weil die Ergebnisse immer mit einer gewissen Irrtumswahrscheinlichkeit behaftet sind, von der Qualität der formulierten Fragestellung abhängen und natürlich auch der Datengüte.

Die Berechnung einer erwarteten risikoadjustierten Letalitätssrate ist ein prognostisches Verfahren, kein diagnostisches. Auffälligkeiten müssen daher eine weiterführende Analyse auslösen, dies ist der diagnostische Teil der Qualitätsbewertung.

An dieser Stelle formulieren wir weitere Empfehlungen:

Ein regelmäßiges Monitoring der beobachteten und erwarteten risikoadjustierten Letalitätsraten sollte durchgeführt werden.

Zu den erwarteten und beobachteten Werten sollte das Konfidenzintervall angegeben werden.

Jede Überschreitung des Konfidenzintervalls durch den beobachteten Wert ist als Signal zu verstehen und sollte eine weiterführende Analyse auslösen.

Werden risikoadjustierte Daten zur Qualitätsbewertung eingesetzt, so sollten auch Daten zu den Fallzahlen und den Häufigkeiten der Risikofaktoren im betreffenden Krankenhaus und der Vergleichsgruppe zur Unterstützung der Interpretation vorliegen.

Weiterhin sollten Informationen zur statistischen Güte des angewendeten Verfahrens zur Risikoadjustierung vorliegen, damit diese bei der Qualitätsbewertung berücksichtigt werden können.

Balkendiagramm (risikoadjustierte Krankenhausletalität)

Eine andere Form der Darstellung ist die risikoadjustierte Krankenhaussterblichkeit (**Abb. 15**): Zu ihrer Berechnung für die einzelnen Krankenhäuser wird das Verhältnis beobachtet/erwartet (%) mit der verbundweit beobachteten Krankenhausletalität (%) multipliziert.

Der so berechnete Wert beschreibt dann für jedes Krankenhaus das Ergebnis, das erreicht würde, wenn alle Krankenhäuser Patienten mit der gleichen Krankheitschwere behandelt hätten. Damit werden alle Krankenhausergebnisse direkt miteinander vergleichbar (Ash et al. 2003, Seite 306).

Zu erkennen ist hier, dass das Krankenhaus B jetzt nicht mehr den höchsten Wert aufweist und auch Krankenhaus A nach rechts gewandert ist. Dies liegt an der Ausprägung der beobachteten und erwarteten Werte und soll hier nicht näher vertieft werden.

Beide Kliniken haben offensichtlich von dem Verfahren profitiert. Krankenhaus B liegt jedoch auch hier im Bereich der höchsten Werte.

Zeitreihe

Auch die risikoadjustierte Krankenhausletalität kann in einer Zeitreihe grafisch dargestellt werden, wie **Abb. 16** und **Abb. 17** für die Krankenhäuser A und B zeigen.

Der Vergleich dieser Zeitreihen mit denen der nicht risikoadjustierten Daten (**Abb. 8, 9**) zeigt Unterschiede: Bei Krankenhaus A liegen nun 5 Datenpunkte über dem Konfidenzintervall des Vergleichswertes, in **Abb. 8** waren es noch 12 Datenpunkte. Auch die Anzahl der Datenpunkte unterhalb des Konfidenzintervalls des Vergleichswertes ist nun größer (14 versus 11).

Auch Krankenhaus B zeigt hier weniger Signale: Ehemals lagen 15 Datenpunkte oberhalb des Konfidenzintervalls des Vergleichswertes (**Abb. 9**), nun sind es 11. Die Anzahl der Datenpunkte unterhalb des Konfidenzintervalls des Vergleichswertes ist unverändert.

Bezüglich der möglichen Fragen und weiteren Analyse gelten die vorherigen Ausführungen und als Empfehlungen werden formuliert:

Ein regelmäßiges Monitoring der risikoadjustierten Letalitätsraten sollte durchgeführt werden, da die Krankenhäuser durch dieses Verfahren direkt verglichen und so Hinweise auf gute Praxis identifiziert werden können.

Das regelmäßige Monitoring der risikoadjustierten Letalitätsraten sollte auch mittels Zeitreihen stattfinden.

Funnel Plot

Grundlegende Informationen zu Funnel Plots sind zu finden bei Spiegelhalter (2005) sowie Lack & Gerhardinger (2009). Mayer et al. (2011) zeigen die Unterschiede zwischen konventionellen Darstellungen (zum Beispiel Balkendiagramme, Caterpillar-Diagramme) und Funnel Plots am Beispiel der radikalen Zystektomie (Daten des National Health Service aus den Jahren 2000 bis 2006). Becker & Eissler (2013) stellen mittels Funnel Plot risikoadjustierte primäre Sectorraten dar.

Bei einem Funnel Plot handelt es sich um eine fallzahlabhängige Darstellung aus der statistischen Prozesslenkung mit globalem Sollwert sowie Eingriffs- und Warngrenzen. Anstelle der in Kontroll- und Regelkarten üblichen Zeitachse wird die jeweilige Messgröße in Abhängigkeit des Stichprobenumfangs dargestellt (in Anlehnung an: Lack & Gerhardinger 2009). Mit den Funnel Plots werden Daten also in geeigneter Form dargestellt und dabei gleichzeitig die Fallzahl berücksichtigt.

Der Funnel Plot in **Abb. 18** enthält zwei sogenannte Funnel Limits, die um den Mittelwert (horizontale graue Linie) berechnet werden und Konfidenzintervallen entsprechen. Unter der Annahme, dass vom Durchschnitt abweichende Raten nur zufallsbedingt sind, besteht eine 95%ige Wahrscheinlichkeit, innerhalb des inneren Limits zu liegen und eine 99,8%ige Wahrscheinlichkeit, innerhalb des äußeren Limits zu liegen (Bragg et al. 2010).

Es ist gut zu erkennen, dass die Limits mit steigender Fallzahl enger werden.

Um eine anonymisierte Darstellung zu erreichen, verzichten wir auf die Angabe der Fallzahl auf der x-Achse, dieses Vorgehen wird in abgewandelter Form beispielsweise auch von der britischen Healthcare Commission verwendet (zitiert in Lack & Gerhardinger 2009).

Insgesamt sehen wir im Funnel Plot ein gutes Ergebnis, da nur eine Klinik ein oberes Limit (hier das 95%-Limit) überschreitet.

Da den Krankenhäusern in der Regel »zu hohe« Sterblichkeitsraten vorgeworfen werden, wird in der Abbildung zunächst der Blick auf Häuser gerichtet, die die oberen Limits tangieren (Kreismarkierungen). Eine Klinik liegt auf dem 95%-Limit und eine Klinik zwischen den oberen Limits. Beiden Kliniken wäre eine weiterführende Analyse der Ergebnisse zu empfehlen.

Doch wir sehen auch Hinweise auf möglicherweise beste Praxis, denn zwei Kliniken liegen auf dem unteren 95%-Limit (Pfeilmarkierungen), ihre Ergebnisse können also nicht zufallsbedingt niedrig sein.

In beiden Fällen gelten die in den vorigen Abschnitten formulierten Ausführungen bezüglich der Fragen aus den Ebenen 1 und 2 der Analysepyramide.

Krankenhaus A liegt in Anbetracht der bisherigen Ergebnisse und der im Vergleich hohen Fallzahl erwartungsgemäß unter dem Mittelwert und nahe dem unteren 95%-Limit.

Die fallzahlabhängige Berechnung der Limits zeigt ihre Bedeutung bei Krankenhaus B, welches innerhalb der 95%-Limits liegt. Aus unserer Sicht liegt hierin kein Widerspruch zu den Feststellungen auf Basis der beobachteten und erwarteten Krankenhausletalität sowie der risikoadjustierten Krankenhausletalität, da der Wert des Krankenhauses sehr nah bei dem oberen 95%-Limit liegt. Aus Sicht einer (klinischen) Leitung würden wir diese Erkenntnis aus dem Funnel Plot als eine Ergänzung zu den bisherigen Feststellungen werten und eine weitere Analyse des klinischen Prozesses durchführen.

Bereiche unterschiedlicher Letalitätsraten bei ähnlichen Fallzahlen (»Ergebniscluster«, schraffierte Markierungen) führen zu der Frage, welchen Einfluss Strukturen und Behandlungsprozesse der betreffenden Kliniken auf die abgebildete risikoadjustierte Krankenhausletalität haben könnten. Diese Fragen können Anlass zu einem Benchmarking der dahinterliegenden Strukturen des klinischen Prozesses geben.

Dimick et al. (2004) haben in einer grundlegenden Arbeit gezeigt, dass die Berücksichtigung der Fallzahlen bei der Beurteilung von Ergebnissen elementar ist. Auch der hier abgebildete Funnel Plot zeigt, dass Kliniken mit geringeren Fallzahlen wesentlich größere Spannweiten der Werte aufweisen können, bevor sie die Limits tangieren, also »auffällig« werden. Hierbei spielt es zunächst keine Rolle, ob es sich um nicht adjustierte oder adjustierte Werte handelt.

Ist das Ziel jedoch eine seriöse Abbildung entsprechender Daten, so sollte dieser Umstand nicht als Chance für »kleine Fallzahlen« oder gar »kleine Kliniken« betrachtet werden, sich einer internen wie externen Qualitätsbewertung zu entziehen. Es liegt in der Natur der Sache, dass eine fallzahlabhängige Streuung existiert und außerdem könnte ja auch die Sicht eingenommen werden, dass die geringeren Fallzahlen eine Datenbewertung deutlich erschweren, da sich Trends oder Auswirkungen von Veränderungsprozessen nur sehr schwer erkennen lassen und nur mit großer Unsicherheit behaftet bewertet werden können. Dies bedeutet also, dass sich Kliniken mit höheren Fallzahlen – zumindest aus statistischer Sicht – bei der Daten- und auch Qualitätsbewertung in einer vorteilhaften Situation befinden.

Empfehlungen:

Die grafische Darstellung risikoadjustierter Letalitätsraten sollte unter Berücksichtigung von Fallzahlen mittels Funnel Plot erfolgen.

Werden Ergebniscluster identifiziert, so sollten die darin enthaltenen Kliniken eine weiterführende Analyse durchführen und ein Benchmarking in Betracht ziehen.

Variable Life Adjusted Display (VLAD)

Hierbei handelt es sich um eine noch wenig verbreitete Methode, wie eine Recherche in PubMed (US National Library of Medicine, National Institutes of Health) zeigt: Die Suche nach »Variable Life Adjusted Display« oder »VLAD« in den Feldern Titel und Abstract ergibt insgesamt 68 Treffer, davon jedoch keiner aus dem deutschsprachigen Raum (Stand 30.05.2013). Daher dürfte diese Arbeit die erste deutschsprachige Publikation zur Anwendung der Methode in einer Fachzeitschrift sein.

Eine gute Übersicht zu Grundlagen und Anwendungen geben die folgenden Autoren: Andrianopoulos et al. (2012), Coory et al. (2008), Duckett et al. (2007), Flaatten (2005), Fusco et al. (2012), Grunkemeier et al. (2003), Guest et al. (2012), Noyez (2009), Pagel et al. (2012), Pagel et al. (2013), Roberts et al. (2012), Sherlaw-Johnson et al. (2000), Sherlaw-Johnson (2005), Tan et al. (2005) und Winkel & Zhang (2007).

Viele Anregungen zur Anwendung sind auf den Webseiten australischer Bundesstaaten zu finden, so zum Beispiel von Queensland und Victoria (besonders interessant, da hier VLAD und die Analysepyramide verbunden werden). Die jeweiligen Webadressen sind im Literaturverzeichnis angegeben.

VLAD basiert als grafische Darstellung der kumulierten Ergebnisqualität auf der »Cumulative Sum (CUSUM)«-Technik, die in dieser Arbeit bereits erläutert wurde.

Die Daten für die VLAD-Grafik in **Abb. 19** wurden für das Krankenhaus A wie folgt berechnet (**Tab. 2**): Die 432 Patienten werden chronologisch nach ihrem Entlassungsdatum aufsteigend nummeriert (Spalte: Patient). Die erwartete Sterblichkeit wird für jeden Patienten mittels des oben bereits erwähnten statistischen Modells berechnet (Spalte: E) und mit dem beobachteten Ergebnis (überlebt/verstorben, Spalte: B) aufgetragen.

Sodann wird für jeden Patienten die Differenz [E – B] berechnet, dies erfolgt vom Prinzip her wie bei CUSUM:

- Für Patient 1 wurde der Erwartungswert 0,146 berechnet, der Patient ist nicht verstorben und daher wurden 0,146 (0,146 - 0) »Leben gerettet«, oder mit anderen Worten: »Es sind 0,146 Patienten weniger verstorben als erwartet«.
- Bei Patient 5 ist der berechnete erwartete Wert 0,253 und da der Patient verstorben ist, lautet die Aussage nun: »Es wurden 0,747 Leben verloren« oder »Es sind 0,747 Patienten mehr verstorben als erwartet«.

In der Spalte VLAD werden dann die einzelnen Differenzen (Spalte: E - B) kumulativ addiert: So ergibt die Addition der Werte von Patient 1 (0,146) + Patient 2 (0,218) den VLAD-Wert 0,364 bis zum Ergebnis bei Patient 432 mit 20,883.

Dies bedeutet also für das Krankenhaus A, dass im Jahr 2012 rund 21 Patienten weniger verstorben sind, als erwartet wurde. Dies war zu erwarten, da in **Abb. 14** zu erkennen war, dass der beobachtete Wert unter dem 95%-KI des erwarteten Wertes lag.

Die unterschiedlichen Effekte des Überlebens beziehungsweise Versterbens von Patienten mit niedriger beziehungsweise hoher erwarteter Sterblichkeit sollen an dieser Stelle noch anhand eines Beispiels erläutert werden:

Überlebt ein Patient mit einem niedrigen Erwartungswert (0,1), so führt dies nur zu einem geringen Anstieg des VLAD um 0,1 (0,1 - 0). Verstirbt dieser Patient, so geht er mit dem Wert -0,9 in den VLAD ein (0,1 - 1).

Anders verhält es sich bei einem Patienten mit hoher erwarteter Sterblichkeit, zum Beispiel 0,8. Überlebt der Patient, steigt das VLAD um 0,8 (0,8 - 0). Stirbt der Patient, fällt die Kurve um 0,2 (0,8 - 1).

Dies bedeutet also: Die stärksten Anstiege der VLAD werden durch überlebende Patienten mit hohem Erwartungswert und die stärksten Abfälle durch verstorbene Patienten mit geringem Erwartungswert verursacht.

Die in Anlehnung an Pagel et al. (2013) erstellte **Abb. 20** zeigt diese VLAD-Eigenschaften anhand von zwei vergrößerten Abschnitten aus **Abb. 19**.

Auch die VLAD-Grafik leistet einen guten Beitrag, wenn man »statistische Signifikanz« in »klinische Bedeutung« übersetzen möchte. Auch hier entspricht jeder Datenpunkt der Linie einem Patienten und daher können auffällige VLAD-Abschnitte den betreffenden Patienten und so auch Zeiträumen zugeordnet werden.

Solche Zeiträume sind auch in **Abb. 19** auf der x-Achse erkennbar, sie ist mit den Monatsnamen beschriftet und zeigt, dass in den einzelnen Zeiträumen unterschiedliche Fallzahlen vorliegen.

Von Januar bis Mitte Februar steigt die VLAD an, um dann bis Mitte März abzufallen. Zu diesem Zeitpunkt sind kumulativ betrachtet (!) also genau so viele Patienten verstorben, wie erwartet wurde.

Ab Mitte März zeigt die VLAD einen beinahe kontinuierlich steigenden Verlauf und so stellt sich natürlich die Frage, ob dieser Anstieg zu irgendeinem Zeitpunkt im positiven Sinne »auffällig« wird. Die Erkennung eines solchen »run of good performance« (Sherlaw-Johnson 2005) ist natürlich wichtig, da er einen Hinweis auf gute Praxis darstellen kann. Hier ist also wieder zu fragen: »Was war in diesem Zeitraum anders? Was haben wir besonders gut gemacht, was wir dann in Zukunft vermehrt tun können?« Ebenso wichtig ist natürlich die Erkennung eines »run of bad performance« (Sherlaw-Johnson 2005), da hier die Möglichkeit eines qualitativ insuffizient ablaufenden Prozesses besteht.

In **Abb. 21** sind nun zusätzlich zwei Limits zu sehen, ein oberes und ein unteres. Beide Limits setzen sich aus Datenpunkten zusammen, die nach einem komplexen Verfahren für den aktuellen VLAD-Wert berechnet werden. Die Berechnung der Limits wird von verschiedenen Autoren erläutert, daher wird hier auf eine detaillierte Beschreibung verzichtet (zum Beispiel Andrianopoulos et al. 2012; Coory et al. 2008; Grunkemeier et al. 2003; Sherlaw-Johnson et al. 2000; Sherlaw-Johnson 2005; Spiegelhalter et al. 2003; Steiner et al. 2000).

Im Zusatzmaterial zu dieser Arbeit befindet sich eine Datei, in der ein VLAD exemplarisch berechnet wird. Die Berechnung der Limits erfolgt dort, wie in den Arbeiten von Andrianopoulos et al. (2012) und Coory et al. (2008) be-

schrieben. Die Limits dienen zur Beantwortung der Frage, ob es sich bei den oben zitierten »runs« um positive beziehungsweise negative Auffälligkeiten handelt, die auch aus Sicht der Statistik auffällig sind, also »Signale« darstellen.

Werden zu diesem Zweck Limits berechnet, so sollten diese möglichst wenig falsch-positive und falsch-negative Signale auslösen. Zusätzlich sollten richtig-positive Signale möglichst früh erkannt werden und falsch-positive Signale sollten möglichst selten ausgelöst werden (Coory et al. 2008).

Coory et al. (2008) berichten in ihrer Publikation zur Berechnung von drei Limits, die jeweils für eine 30-, 50- beziehungsweise 75%ige Erhöhung beziehungsweise Verringerung des relativen Risikos stehen. In unserem Beispiel ist das obere Limit auf eine 30%ige Verringerung und das untere Limit auf eine 30%ige Erhöhung des relativen Risikos ausgelegt. Durch die Anwendung dieses Limits werden positive Signale am frühesten entdeckt.

Der VLAD-Wert bei Patient 223 (Ende Mai, erster Pfeil) ist höher als der korrespondierende Wert des oberen Limits, es handelt sich also um ein positives Signal. Es besagt, dass zu diesem Zeitpunkt kumulativ mehr Patienten überlebt haben, als erwartet wurde. Diese Feststellung basiert auf den folgenden Daten: Von den 223 Patienten sind 27 verstorben, der erwartete Wert beträgt 36. Wären 30% mehr Patienten tatsächlich verstorben, so wären dies 35 ($27 + (27 \times 0,3)$). Der so berechnete Wert (also: 30% oberhalb des Beobachtungswertes) liegt immer noch unter dem Erwartungswert (36) – es handelt sich um ein positives Signal zu einem »run of good performance«.

Sobald ein Signal erkannt wurde, wird ein Limit »zurückgesetzt«, was bedeutet, dass die Limitberechnung ab diesem Punkt neu beginnt und erneut nach einer 30%igen Verringerung oder Erhöhung des relativen Risikos starten kann.

Dies ist dann der Fall bei Patient 340 (Anfang Oktober, zweiter Pfeil) – der zweite »good run« wurde also erkannt. Das Limit wird nach dem ersten Signal neu berechnet, jetzt also ab Patient 224. Bei den Patienten 224 bis 340 beträgt der Beobachtungswert 11 und der Erwartungswert 21. Erneut liegt der um 30% erhöhte Beobachtungswert ($11 +$

$(11 \times 0,3)$) mit 14,3 unter dem Erwartungswert (hier: 21). Erneut wurde ein positives Signal erkannt.

Die VLAD-Grafik für Krankenhaus B zeigt **Abb. 22**. Insgesamt sterben kumulativ rund 5 Patienten mehr, als aufgrund der Risikoadjustierung erwartet wurde. Dennoch gibt es kein Signal für einen »bad run«, was daran liegt, dass einem Abfall der VLAD immer ein Anstieg vorausgeht. Die VLAD zeigt keinen langfristig eindeutigen Trend, wir sehen so etwas wie ein Sägezahnmuster.

Spätestens bei Patient 139 (Mitte Juni) würde man die Auslösung eines negativen Signals intuitiv erwarten. Die nachfolgende Berechnung bestätigt aber, dass dies zu Recht nicht geschieht: 30 Patienten sind verstorben, der Erwartungswert beträgt 25. Da es sich jetzt um das untere Limit handelt, lautet die Frage nun also: Wurde das untere Limit um mindestens 30% unterschritten? Dies ist nicht der Fall, da der Beobachtungswert mit 30 verstorbenen Patienten immer noch kleiner ist als der um 30% erhöhte Erwartungswert ($25 + (25 \times 0,3)$).

In der Abbildung sind drei Bereiche gekennzeichnet, deren Beginn und Ende durch die gut erkennbaren Tiefpunkte der VLAD bestimmt wird. In den drei Bereichen kann man nun mithilfe eines statistischen Verfahrens untersuchen, ob sich die beobachtete (B) und erwartete (E) Letalitätsrate statistisch signifikant ($p < 0,05$) unterscheiden (hierzu verwendeten wir den Befehl »smyrb« der Statistiksoftware STATA® StataCorp LP, Texas, USA, in der Version 12.1). In keinem der drei Bereiche wurde ein signifikanter Unterschied festgestellt.

Dies bedeutet natürlich nicht, dass die in der VLAD-Grafik gut zu erkennenden abfallenden Abschnitte nicht dennoch weiter analysiert werden können und auch sollten.

Duckett et al. (2007) betonen hierzu ausdrücklich: »VLADs do not provide definitive answers about the quality of care. They are used to develop ideas about why variations in reported outcomes occur and suggest possible solutions, be they ways of improving data quality, improving casemix adjustment, or implementing system changes to improve quality of care.«

Eine Weiterentwicklung des VLAD stellen Pagel et al. (2013) vor und erhöhen damit nochmals die Anwendungsmöglichkeiten, indem einzelne Datenpunkte – die ja einzelnen Fällen entsprechen – farblich markiert werden. Diese Idee haben wir in die Entwicklung der VLAD im CLINOTEL-Krankenhausverbund übernommen, sie wird am Beispiel von Krankenhaus A in **Abb. 23** erläutert. Die Inhalte dieser Abbildung sind im Wesentlichen bereits aus **Abb. 21** bekannt, erkennbar sind jedoch zwei Unterschiede: Auf der x-Achse befinden sich nun keine Monatsangaben, sondern Angaben zu Fallnummern, die eine ungefähre Zuordnung von Fallnummern zu Datenpunktabschnitten ermöglichen.

Der besonders innovative Ansatz in dieser Abbildung liegt in der zusätzlichen farbigen Markierung von Datenpunkten, bei der wir die folgenden Farben verwenden:

- Ein grüner Datenpunkt zeigt, dass das obere Limit erreicht wurde beziehungsweise ein Patient mit einem hohen Wert der individuell berechneten risikoadjustierten Sterblichkeit überlebt hat. Die Berechnung erfolgt mittels des bereits erwähnten statistischen Modells, als Schwellenwert verwenden wir hier das 95%-Perzentil, welches in unseren Daten bei dem Wert 0,46 beginnt.
- Ein Datenpunkt ist mit rot markiert, wenn das untere Limit erreicht wurde beziehungsweise ein Patient mit einem niedrigen Erwartungswert verstorben ist. Hier beträgt der Schwellenwert 0,015 und entspricht dem 5%-Perzentil.
- Da uns die Routinedaten der Patienten vorliegen, können wir die Nebendiagnosen der Patienten auch auf unerwünschte Ereignisse hin untersuchen, die als Nebendiagnosen kodiert vorliegen können. Liegt eines oder sogar mehrere der in der Abbildung aufgeführten Ereignisse vor, so wird der Datenpunkt mit der Farbe Orange gekennzeichnet.

Welches Ereignis bei einzelnen Patienten vorliegt, kann einer Tabelle entnommen werden, die die Mitglieder des CLINOTEL-Krankenhausverbundes mit den Auswertungen zur Qualitätssicherung mit Routinedaten ab dem Jahr 2014 erhalten werden (**Tab. 5**).

Aus der Tabelle kann man zu den einzelnen Fällen verschiedene Informationen entnehmen: In der ersten Spalte (Lfd. Nr) die Nummer des Datenpunktes in der VLAD-Grafik. In den nächsten Spalten werden dann die hier nicht zu sehende Fallnummer des Patienten sowie die ebenfalls hier in der Veröffentlichung nicht abgebildeten Datums- und Zeitangaben zu Aufnahme und Entlassung gelistet.

Die Spalte »Expected« zeigt den berechneten Erwartungswert der Krankenhaussterblichkeit, gefolgt vom beobachteten Ergebnis.

In der Spalte »VLAD« wird der VLAD-Wert des Datenpunktes angegeben, gefolgt von den dazugehörigen Werten des oberen und unteren Limits.

Der letzten Spalte können dann die oben erläuterten Ereignisse entnommen werden.

Der Vollständigkeit halber soll hier noch erwähnt werden, dass die farbige Markierung von Datenpunkten natürlich auch bei der CUSUM-Technik angewendet werden kann.

Auch zum VLAD kann zusammenfassend gesagt werden, dass das VLAD die Krankenhausletalität als Zielgröße unserer Betrachtungen nicht in einem Prozentwert, sondern als zu einem bestimmten Zeitpunkt gewonnene oder verlorene Leben darstellt. Somit kann auch VLAD sehr plastisch die Auswirkungen von Interventionen aufzeigen.

Empfehlungen:

Ein regelmäßiges Monitoring der beobachteten und erwarteten risikoadjustierten Letalitätsraten mittels VLAD sollte durchgeführt werden.

Jedes Signal sollte eine weiterführende Analyse auslösen.

Auch Trends ohne Signal sollten einer weiterführenden Analyse unterzogen werden.

CUSUM und VLAD in der Gesamtschau

CUSUM und VLAD leisten einen guten Beitrag, wenn man »statistische Signifikanz« in »klinische Bedeutung« übersetzen möchte.

Neben diesen Gemeinsamkeiten gibt es auch Unterschiede, die sich im Wesentlichen aus den unterschiedlichen Erwartungswerten ergeben: Während CUSUM konstante Erwartungswerte verwendet, werden bei VLAD patientenbezogene, risikoadjustierte Erwartungswerte eingesetzt.

Die sich hieraus ergebenden Unterschiede der CUSUM- und VLAD-Werte sollen hier nochmals unter Verweis auf die **Tabellen 1** und **2** erläutert werden:

Die Berechnung der Differenz [E – B] ergibt bei Patient 2 deutlich unterschiedliche Werte, nämlich 0,127 (CUSUM) versus 0,218 (VLAD). Bei Patient 5 beträgt der Wert –0,873 (CUSUM) versus –0,747.

Hieraus folgt: Ist bei einem individuellen Patienten der berechnete Erwartungswert (VLAD) größer als der konstante (CUSUM), so ist auch der resultierende Wert [E – B] des VLAD größer.

Der umgekehrte Fall, also Erwartungswert VLAD < CUSUM, führt dazu, dass der Wert [E – B] des VLAD kleiner ist.

Diese Feststellungen wirken sich natürlich auch auf die Grafiken aus, hierzu betrachten wir zunächst CUSUM (**Abb. 12**) und VLAD (**Abb. 21**) von Krankenhaus A: CUSUM und VLAD zeigen ein positives kumulatives Endergebnis, jedoch liegt der Wert im CUSUM (+3) deutlich unter dem des VLAD (+21). Dieser Unterschied ergibt sich aus den oben dargestellten Auswirkungen der Erwartungswerte, nämlich fix versus individuell berechnet und sollte daher nicht zu falschen Schlussfolgerungen führen. Für das kumulative Endergebnis steht daher im Vordergrund, dass es ein positives ist, also mehr Patienten überlebt haben, als erwartet wurde. Und dies stellen beide Techniken dar.

In beiden Grafiken können steigende und fallende Kurvenverläufe erkannt werden. Beide Grafiken zeigen Überschreitungen der oberen Limits, das untere Limit wird nur im CUSUM erreicht.

Bei Krankenhaus B kommen wir in CUSUM (**Abb. 13**) und VLAD (**Abb. 22**) zu in der Tendenz identischen Erkenntnissen.

Das kumulative Endergebnis liegt mit –19 im CUSUM deutlich unter dem im VLAD (–5). Während im VLAD kein Signal ausgelöst wird, ist dies im CUSUM mehrfach der Fall. Auch bei diesen Feststellungen gelten die obigen Begründungen. Hinzu kommt natürlich, dass sich die höhere Krankenhaussterblichkeit von Krankenhaus B (20,7% versus CLINOTEL 13,6%/Krankenhaus A 12,0%) nun mit einer relativ hohen Zahl von Fällen mit negativen [E – B]-Werten (0,127 – 1 = –0,873) deutlich auswirkt.

Beide Techniken ermöglichen hinsichtlich der relevanten Aspekte prinzipiell übereinstimmende Schlussfolgerungen und somit verfügen wir über ein Instrumentarium, welches im gesamten Spektrum risikoadjustierter und nicht risikoadjustierter Erwartungswerte anwendbar ist.

Der besondere Vorteil dieser Techniken ist, dass jeder Datenpunkt der Linie einem Patienten entspricht und daher auffällige Abschnitte den betreffenden Patienten und so auch Zeiträumen zugeordnet werden können.

Auf der Webseite von Victorian Government wird hierzu treffend ausgeführt:

»A key strength ... is that each episode of care has an impact on the chart. That is, every patient outcome moves the line up or down according to the outcome of their stay.«

Victorian Government Website

Ergebnisse in der Gesamtsicht

Die Erkenntnisse, die bei der Analyse der Grafiken zu den Krankenhäusern A und B formuliert wurden, sind in **Tab. 3** und **Tab. 4** zusammengefasst.

Hätten wir zwischen den beiden Krankenhäusern eine Entscheidung zu treffen, welches den klinischen Prozess »Diagnostik und Therapie der ambulant erworbenen Pneumonie« einer weiterführenden Analyse unterziehen sollte, so würden wir uns für Krankenhaus B entscheiden. Im Funnel Plot liegt der Wert des Krankenhauses deutlich über dem Mittelwert nur knapp innerhalb des inneren Limits und die risikoadjustierte Krankenhausletalität befindet sich weit links vom Gruppenwert. In den Zeitreihen, in der Regelkarte, CUSUM und VLAD finden wir deutliche Auffälligkeiten. Hier geht es um die Frage, ob das Verhältnis der beobachteten zu den erwarteten Fällen verstorbener Patienten in der Zukunft positiv beeinflusst werden kann. Mit anderen Worten: Gab es in der Vergangenheit vermeidbare Todesfälle und was können wir für die Zukunft lernen?

Zusätzlich würden wir Krankenhaus A darum bitten, zu untersuchen, ob sich im Versorgungsprozess beste Praxis erkennen lässt, die dann anderen Krankenhäusern vermittelt werden kann. Ebenso wäre die Frage zu stellen, ob im Beobachtungszeitraum Veränderungen des klinischen Prozesses (also: Interventionen) durchgeführt wurden.

Schlussfolgerung

Statistische Grafiken können den Transformationsprozess »Datum – Information – Entscheidung« wirksam unterstützen, wie wir am Beispiel der beiden Krankenhäuser zeigen konnten.

Es handelt sich hierbei um einen komplexen Prozess, der wiederum die Komplexität klinischer Prozesse widerspiegelt. Versuche zur Reduktion dieser Komplexität sind nicht zielführend. Aus Sicht mancher Betrachter mag es ja sehr charmant sein, die Qualität eines Versorgungsprozesses, einer medizinischen Fachabteilung oder gar eines ganzen Krankenhauses auf Ampelfarben, Daumenzeichen, Sterne oder Lebensbäume zu reduzieren. Man sollte jedoch nicht ernsthaft glauben, dass hierdurch nachhaltige Verbesserungen der Versorgungsqualität gefördert werden.

Grundlegende Voraussetzung für die Initiierung eines Veränderungsprozesses ist die Akzeptanz der Professionalität der beteiligten Mitarbeiter. Dies beginnt bereits bei der Sequenz »Daten – Informationen – Entscheidungen« und wird erfolgsentscheidend bei der Umsetzung und der Aufrechterhaltung der erzielten Veränderungen. Nicht nur das Problem, hier zum Beispiel die Abweichung von einer evidenzbasierten Leitlinie, wird von repetitiven Handlungsmustern aufrechterhalten. Auch die Lösung muss sich in neuen, beobachtbaren und repetitiven Handlungsmustern zeigen und manifestieren.

Dies erfordert viel Energie, die Menschen nur dann dauerhaft zu investieren bereit sind, wenn die Komplexität ihrer Arbeit und der Rahmenbedingungen anerkannt wird.

Ist die Lösung kein repetitives Handlungsmuster, wird sie nicht lange anhalten. Dieser Effekt wird in der Praxis häufig beschrieben, wenn externe Berater eine Lösung mitbringen und sich nach ihrer Installation wieder zurückziehen. Da es sich in solchen Fällen um die Lösung des Beraters und nicht die der Prozessbeteiligten handelt, nimmt der Berater die Lösung – bildlich gesprochen – wieder mit und der Prozess schwingt in seinen Ausgangszustand zurück. Gute Übersichten zu den Erfolgsfak-

toren erfolgreicher Veränderungsprozesse mit großer klinischer Relevanz geben Goeschel (2011) und Dixon-Woods et al. (2011).

Die weiterführende Analyse sollte – ebenso wie die klinische Diagnostik – geplant und zielgerichtet durchgeführt werden. Hierzu empfehlen wir die dargestellte Analysepyramide.

Neben der Erhebung und Analyse prozessrelevanter Daten empfehlen wir zwei weitere Maßnahmen, die nach unserer Erfahrung wichtige Erkenntnisse zur Qualitätsbewertung, Prävention und Qualitätsplanung liefern können und im Kontext dieser Arbeit Erwähnung finden sollten. Für diese Maßnahmen spricht auch, dass sie (zum Beispiel im Gegensatz zu einem Peer Review) von einem Krankenhaus alleine durchgeführt werden können:

Die Morbiditäts- und Mortalitätskonferenz kann unterjährig zur Reflektion des Versorgungsprozesses eingesetzt werden. Anlass könnte zum Beispiel die Erkennung eines Signals im Variable Life Adjusted Display (VLAD) sein. Folgt die Morbiditäts- und Mortalitätskonferenz einem definierten Ablauf und erfüllt sie weitere Qualitätskriterien, können »Signalfälle« einer qualifizierten Diskussion durch das therapeutische Team zugeführt werden. Dies ist die beste Chance, unsichere Handlungen und Systemfaktoren herauszuarbeiten und von ihnen zu lernen (Becker 2013a).

Ein sogenanntes Prozessaudit bietet sich für definierte klinische Versorgungsprozesse ganz besonders an, da hierbei entlang dem Patientenpfad auditiert wird. Solche internen (DIN EN 15224) Audits können anlassbezogen (zum Beispiel Auffälligkeit in der Regelkarte oder im VLAD) durchgeführt werden oder Bestandteil eines prospektiv geplanten internen Auditprogrammes sein.

So eignet sich der klinische Prozess »Diagnostik und Therapie der ambulant erworbenen Pneumonie« sehr gut als Beispiel, da er der Messung mittels Routinedaten sehr gut zugänglich ist und eine evidenzbasierte Leitlinie existiert (Leitlinie Pneumonie 2009). Aus der Leitlinie können kritische Prozesspunkte herausgearbeitet und gezielt auditiert werden. Dies sind beispielsweise: frühzeitige Risikostratifizierung, Indikationsstellung zur primären Aufnahme auf die Intensivstation, Durchführung angemessener Diagnostik, Einleitung der antibiotischen Therapie etc. Im Idealfall korrespondieren die Messwerte mit den kritischen Prozesspunkten.

So durchgeführt kann das Prozessaudit Auslöser und auch Folge der Sequenz »Daten – Informationen – Entscheidungen« sein.

Das von uns geschilderte Vorgehen unterstützt (klinische) Leitungskräfte bei der Bewertung klinischer Prozesse hinsichtlich der Prozess- und Ergebnisqualität, deren Kenntnis wiederum eine Voraussetzung für die Bewertung der Effizienz darstellt.

Die Kombination verschiedener Grafiken zur Analyse zeigt dabei ihren besonderen Wert, denn sie ermöglicht auch eine unterjährige Betrachtungsweise und so kann jederzeit eine weiterführende Analyse durchgeführt werden. Entscheidungen zu Interventionen im klinischen Prozess können daher frühzeitig getroffen werden und nicht erst nach Ablauf einer längeren Beobachtungszeit. Die so gewonnene Zeit kann aus Sicht der Patienten in der Tat einen Unterschied ausmachen, der einen Unterschied bedeutet.

Wir haben gezeigt, dass eine sachbezogene Herangehensweise an das Treffen von wirksamen Entscheidungen möglich ist. Daher greifen wir das eingangs aufgeführte Zitat von Kierkegaard und verändern es mit Optimismus zu:

»Das Vergleichen ist das Ende des Unglücks und der Anfang der Zufriedenheit.«

Abbildungen

Abb. 1:
Vom Datum zur Information und zur Entscheidung

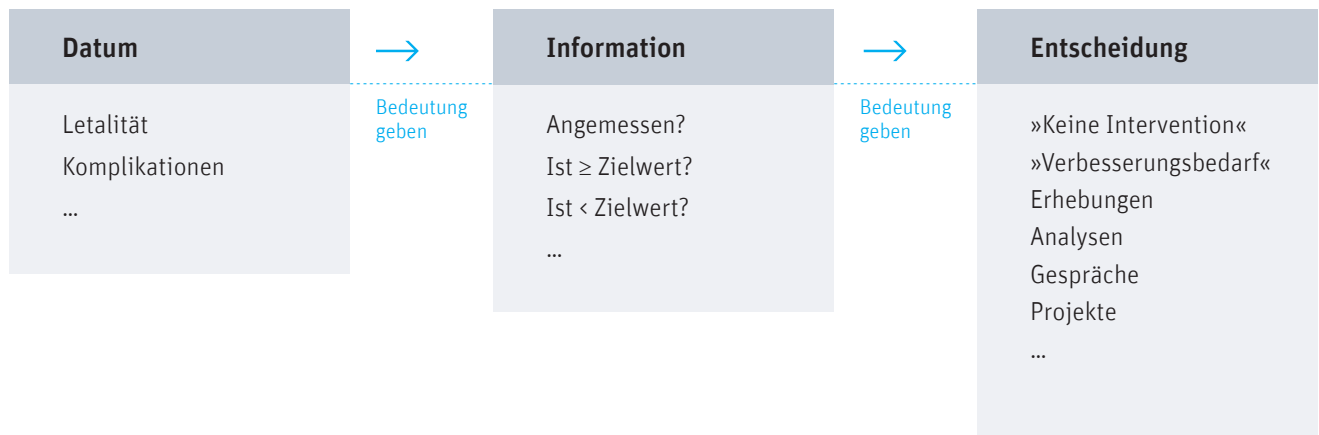


Abb. 2:

Krankenhausletalität (%) ambulant erworbene Pneumonie: Balkendiagramm
Daten 2012 | 10.379 Patienten

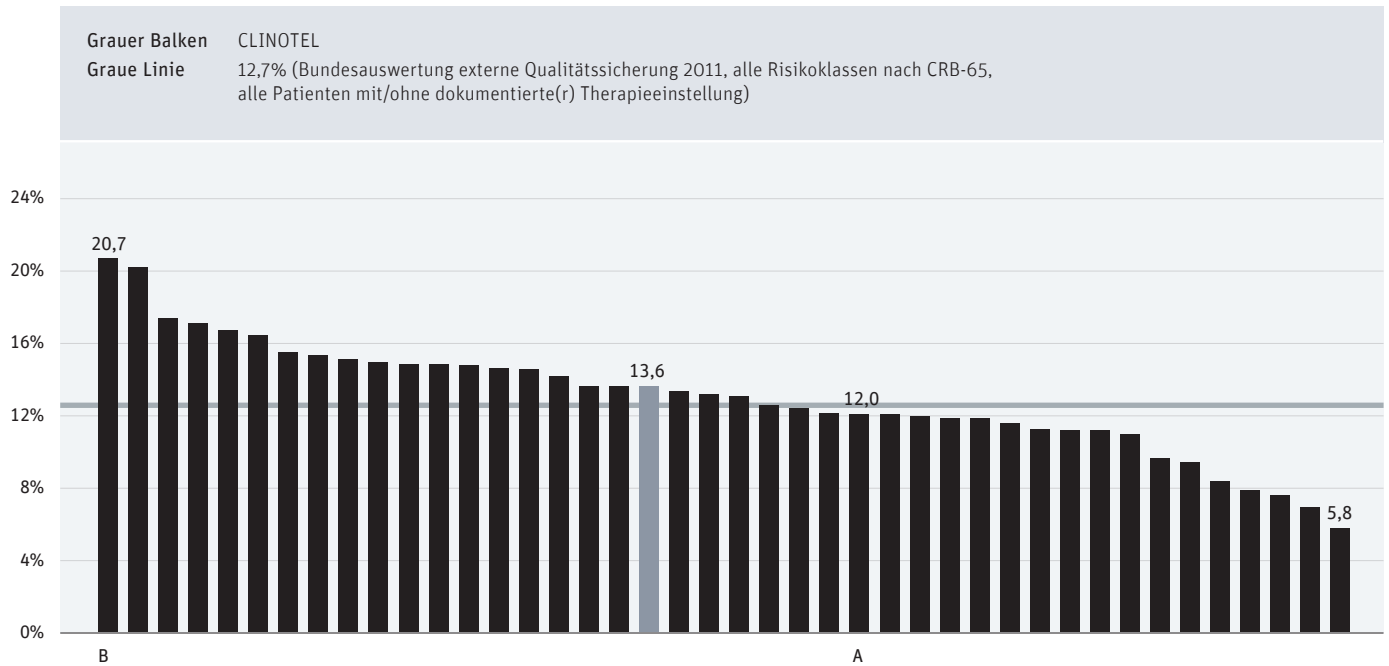


Abb. 3:
Analysepyramide («Pyramid model of investigating hospital performance») in Anlehnung an Mohammed et al. 2004



Abb. 4:

Analysepyramide (»Pyramid model of investigating hospital performance«) in Anlehnung an Mohammed et al. 2004

Ebene 1: Daten

- Wurden die Daten korrekt kodiert?
- Wurde die Hauptdiagnose korrekt zugewiesen?
- Wurden alle Nebendiagnosen korrekt kodiert (qualitativ und quantitativ; auch bei Patienten, die früh versterben)?
- Gab es eine Veränderung der Schlüssel bzw. der Kodierrichtlinien und wurden diese korrekt umgesetzt?
- Gab es eine Veränderung in der Kodierungspraxis?
- Ist die klinische Dokumentation klar, komplett und konsistent?
- Wurden Definitionen korrekt angewendet?
- Sind die Daten vollständig?

Ebene 2: Patienteneigenschaften (Case mix)

- Gibt es bei den Patienten irgendwelche Faktoren, die sich auf die Ergebnisqualität auswirken können und bei der Risikoadjustierung nicht berücksichtigt werden?
- Gab es Änderungen im Zusammenhang mit Aufnahmen bzw. Entlassungen (Indikationsstellung, Vorgehen ...)?
- Patienten mit Status »Nicht reanimieren« (beeinflusst die Ergebnisqualität, kann aber nicht kodiert werden und taucht daher auch nicht in der Risikoadjustierung auf). Bei diesen Patienten werden eventuell auch für die Risikoadjustierung relevante Nebendiagnosen nicht mehr kodiert.

Ebene 3: Strukturen & Ressourcen

- Gab es irgendwelche Veränderungen bei den zur Verfügung stehenden Strukturen und Ressourcen, die sich auf die Ergebnisqualität bzw. die Risikoadjustierung auswirken können?
- Engpässe in den Kapazitäten der OPs, ITS, Funktionsstellen etc.?
- Veränderungen in Verfahren (Leitlinien, SOPs etc.) und deren Kommunikation beziehungsweise Schulung?
- Einführung neuer Verfahren?

Abb. 5:

Analysepyramide (»Pyramid model of investigating hospital performance«)
in Anlehnung an Mohammed et al. 2004

Ebene 4: Versorgungsprozess

- Gab es irgendwelche Veränderungen im Bereich der Versorgungsprozesse, die sich auf die Ergebnisqualität bzw. die Risikoadjustierung auswirken können?
- Gab es Änderungen im Zusammenhang mit Aufnahmen, Entlassungen bzw. internen Verlegungen (Indikationsstellung, Vorgehen ...)?
- Veränderungen beziehungsweise neue Leitlinien?
- Anwendung neuer Geräte, Medikamente etc.?
- Unterschiede zwischen Schichten, Wochentagen, Urlaubszeiten?

Ebene 5: Personal

- Gab es irgendwelche Veränderungen im Bereich des Personals, die sich auf die Ergebnisqualität beziehungsweise die Risikoadjustierung auswirken können?
- Quantitativ (Arbeitsdichte)?
- Qualitativ (Kompetenzen, Fähigkeiten, Personalmix)?
- Verlust von Mitarbeitern mit besonderer Bedeutung für Prozesse und Ergebnisqualität?
- Veränderung von Teamstrukturen?

Abb. 6:
Krankenhausletalität (%) ambulant erworbene Pneumonie: Zeitreihe
Krankenhaus A | Daten 2010 bis 2012 | 1.295 Patienten

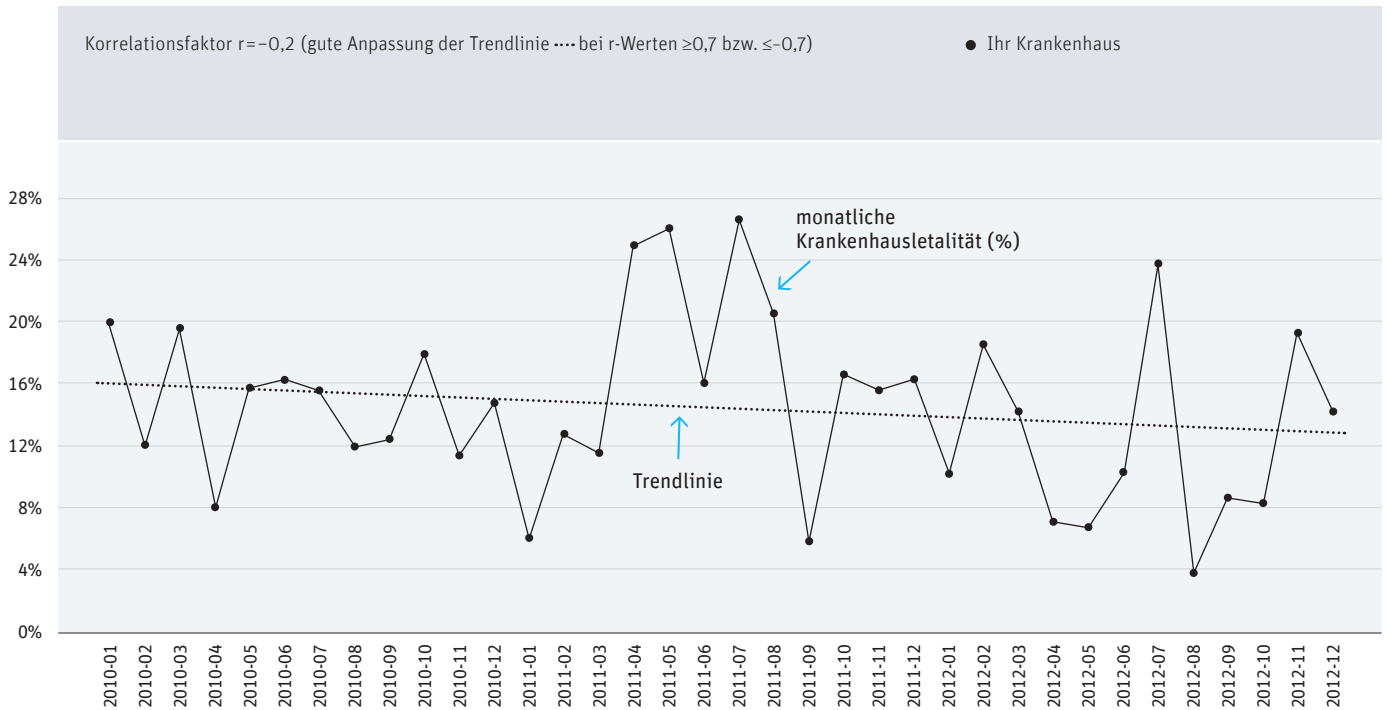


Abb. 7:
 Krankenhausletalität (%) ambulant erworbene Pneumonie: Zeitreihe
 Krankenhaus A | Daten 2010 bis 2012 | 1.295 Patienten

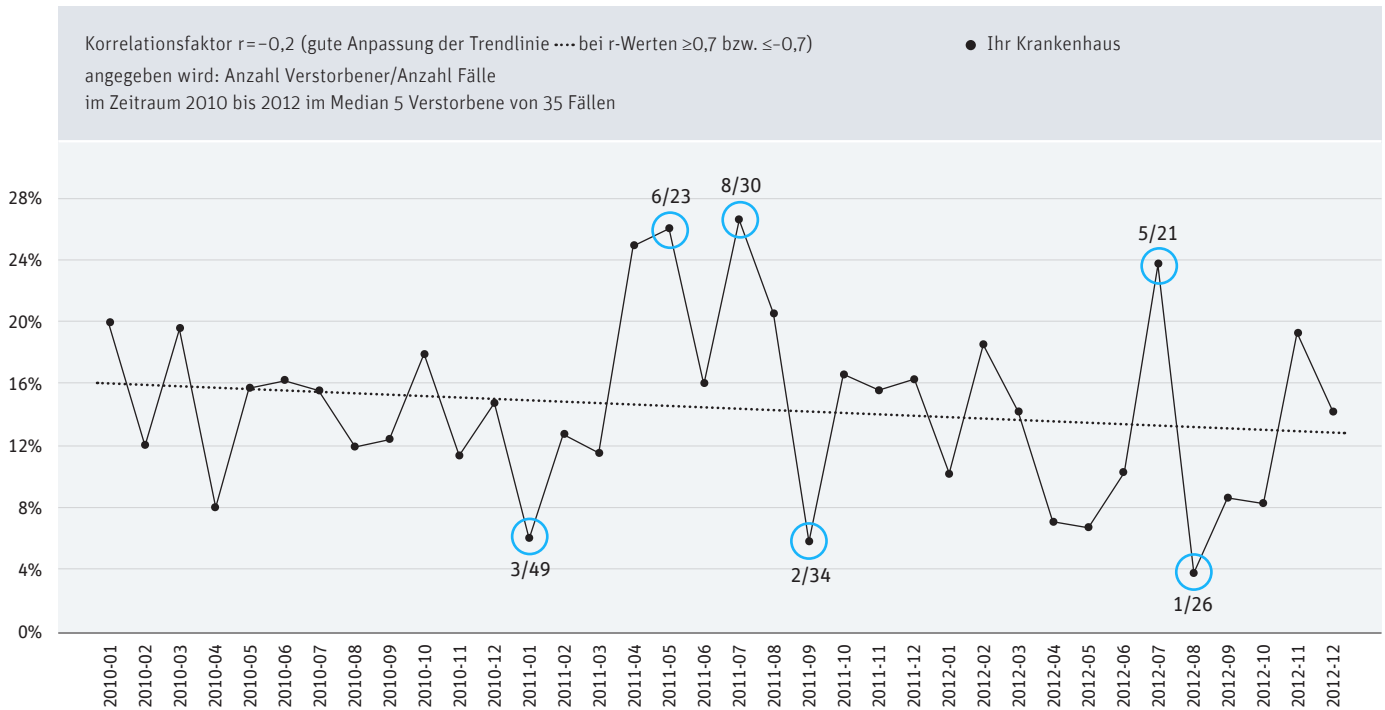


Abb. 8:
Krankenhausletalität (%) ambulant erworbene Pneumonie: Zeitreihe
Krankenhaus A | Daten 2010 bis 2012 | 1.295 Patienten

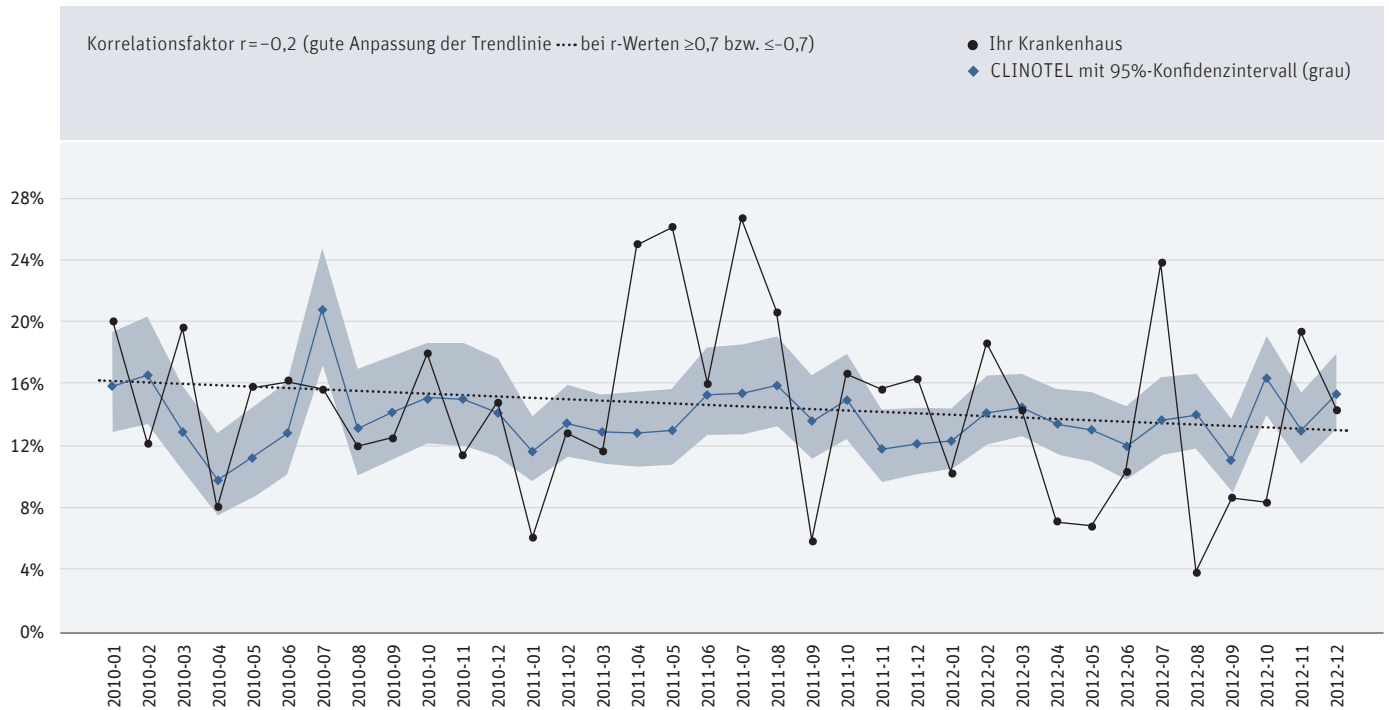


Abb. 9:
Krankenhausletalität (%) ambulant erworbene Pneumonie: Zeitreihe
Krankenhaus B | Daten 2010 bis 2012 | 786 Patienten

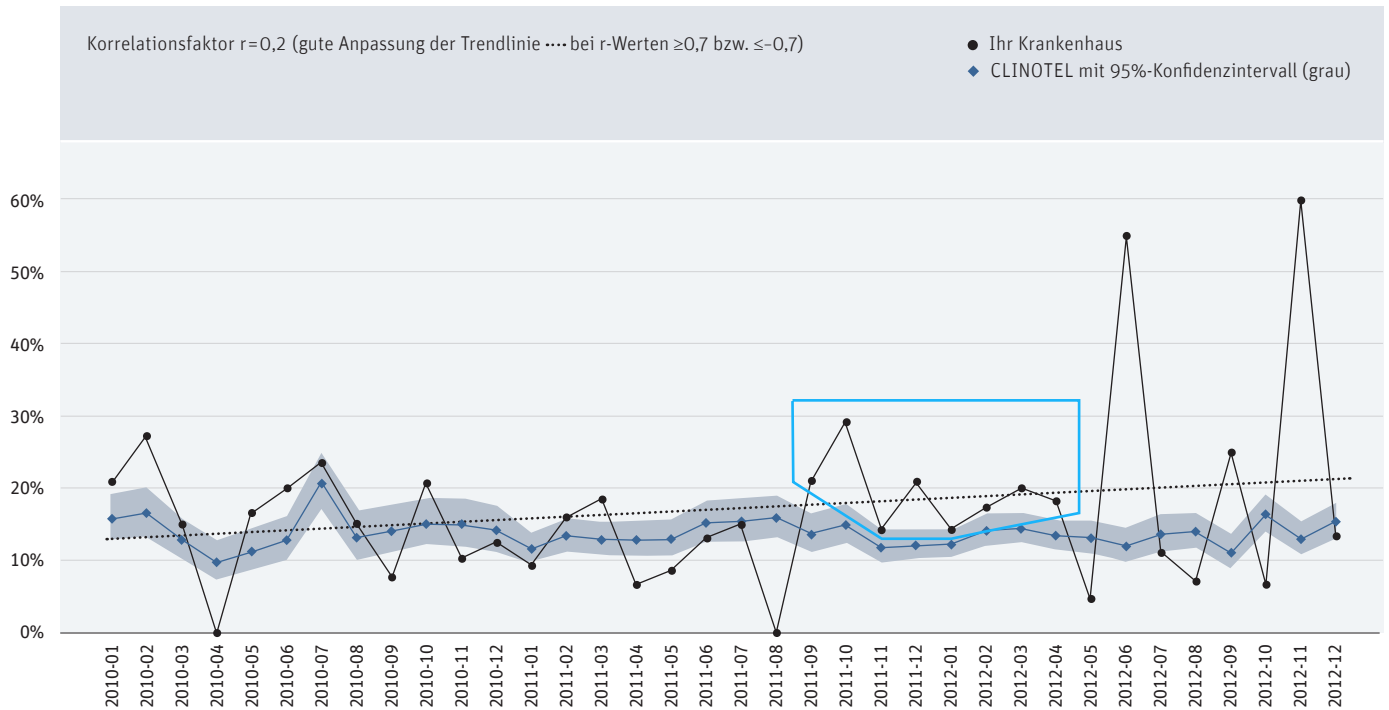


Abb. 10:
Krankenhausletalität (%) ambulant erworbene Pneumonie: p-Regelkarte
Krankenhaus A | Daten 2010 bis 2012 | 1.295 Patienten

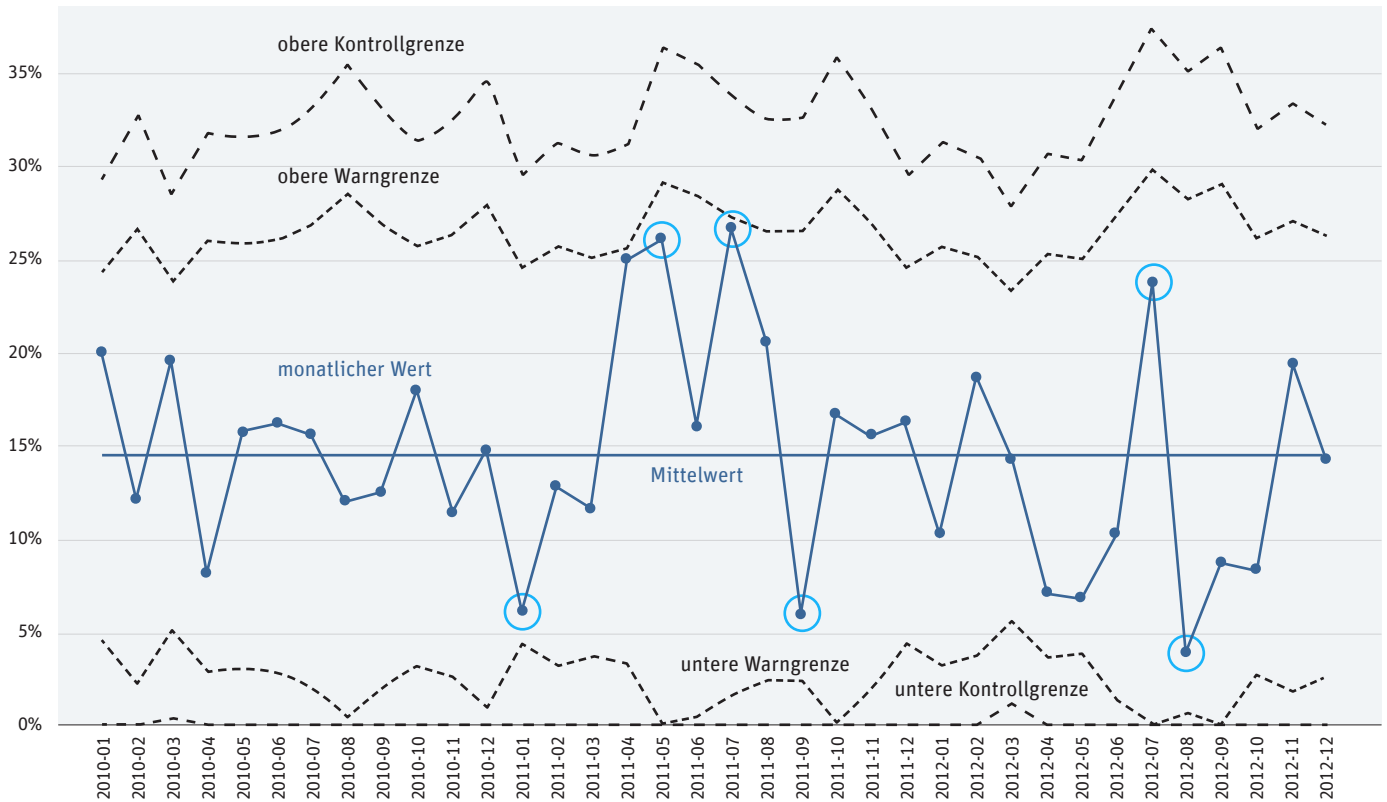


Abb. 11:
Krankenhausletalität (%) ambulant erworbene Pneumonie: p-Regelkarte
Krankenhaus B | Daten 2010 bis 2012 | 786 Patienten

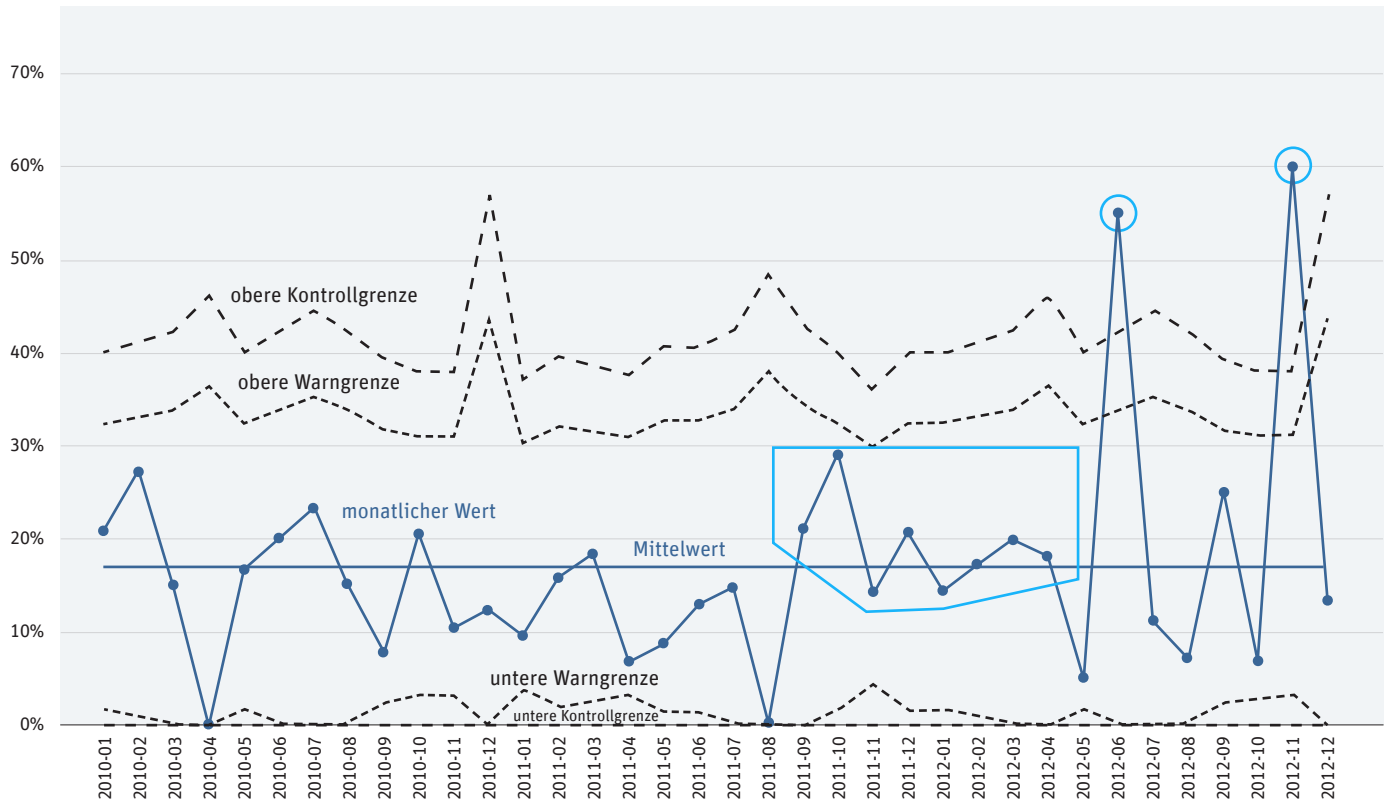


Abb. 12:
Krankenhausletalität ambulant erworbene Pneumonie: Cumulative Sum (CUSUM)
Krankenhaus A | Daten 2012 | 432 Patienten

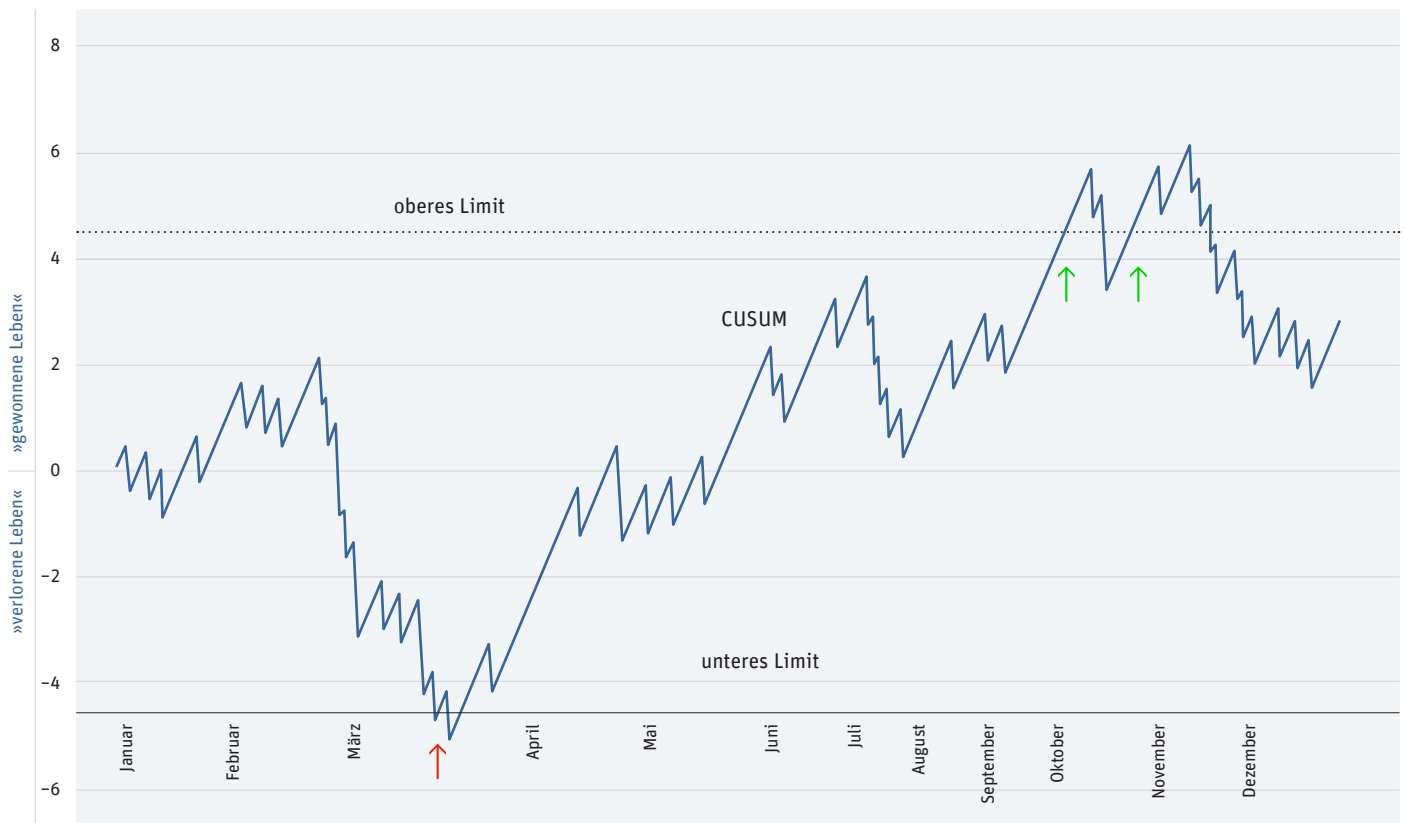


Abb. 13:
Krankenhausletalität ambulant erworbene Pneumonie: Cumulative Sum (CUSUM)
Krankenhaus B | Daten 2012 | 237 Patienten

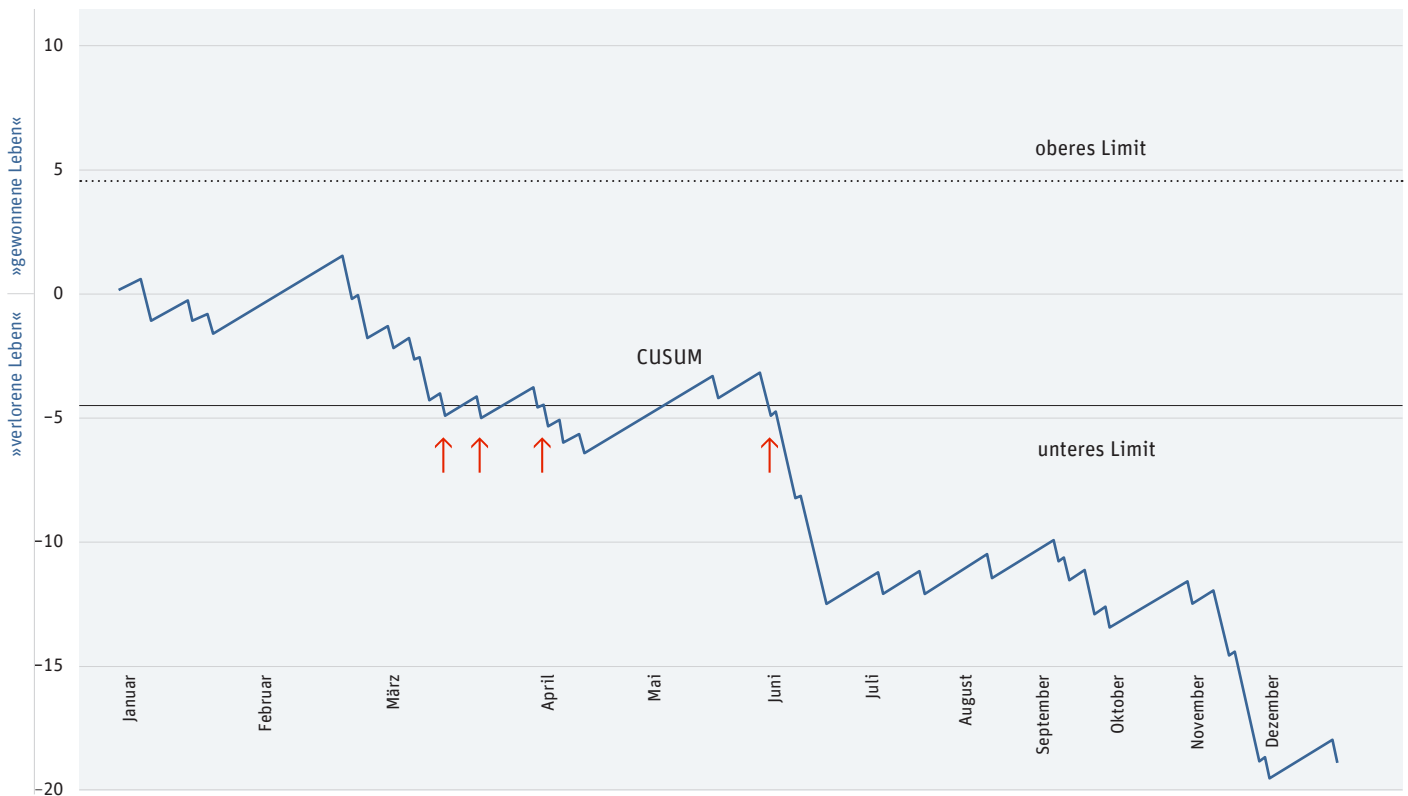


Abb. 15:
Risikoadjustierte Krankenhausletalität (%) ambulant erworbene Pneumonie
Daten 2012 | 10.379 Patienten

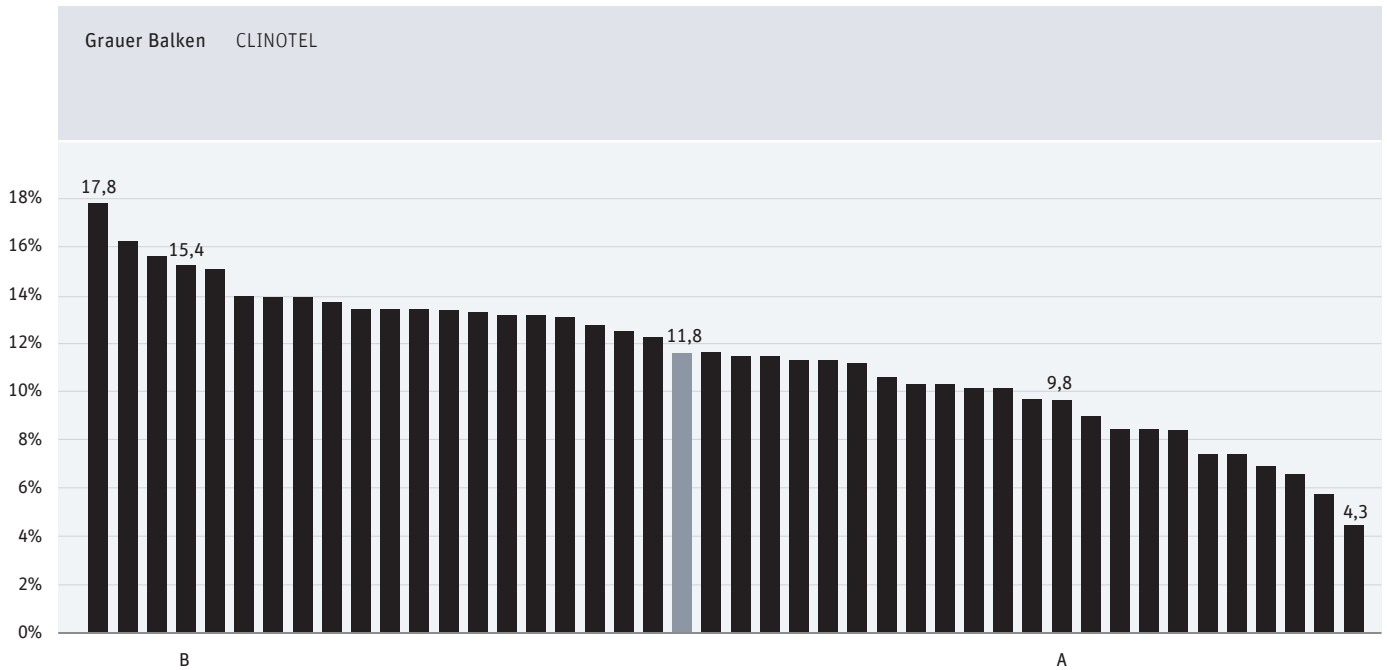


Abb. 16:
Risikoadjustierte Krankenhausletalität (%) ambulant erworbene Pneumonie: Zeitreihe
Krankenhaus A | Daten 2010 bis 2012 | 1.295 Patienten

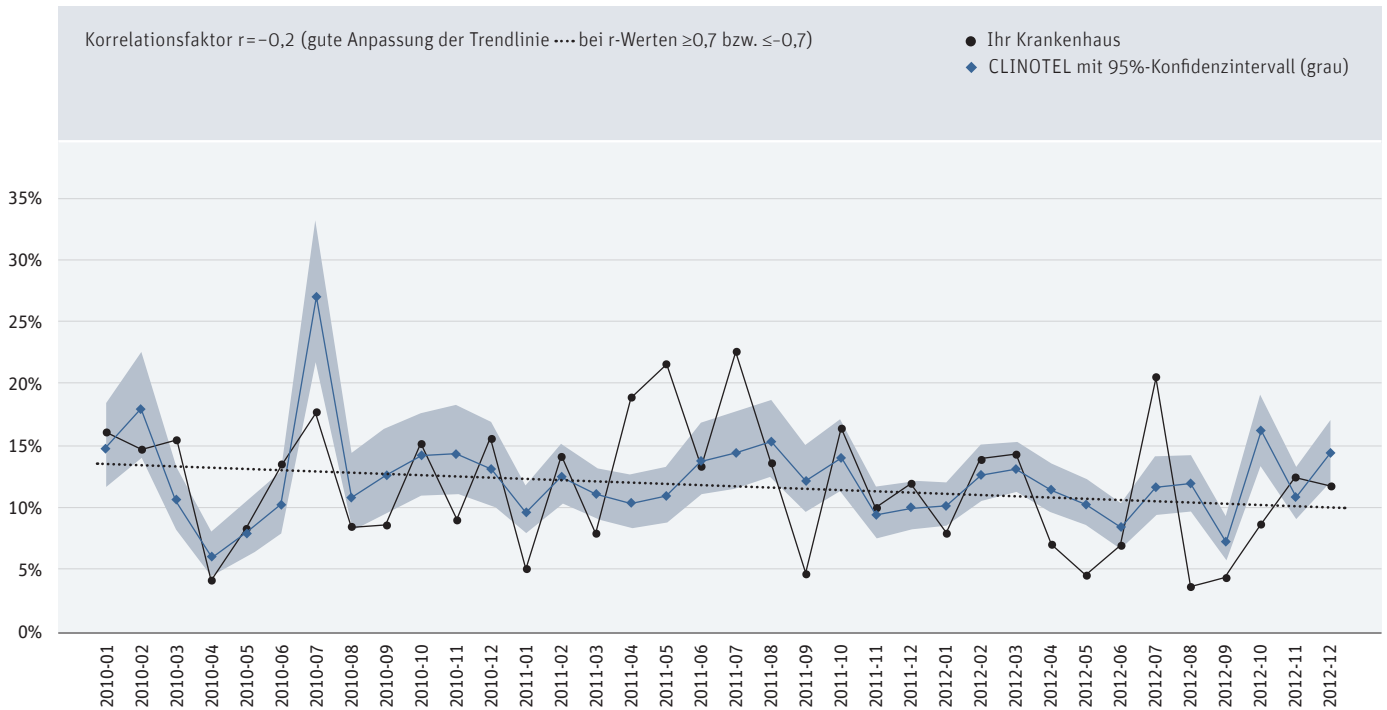


Abb. 17:
Risikoadjustierte Krankenhausletalität (%) ambulant erworbene Pneumonie: Zeitreihe
Krankenhaus B | Daten 2010 bis 2012 | 786 Patienten

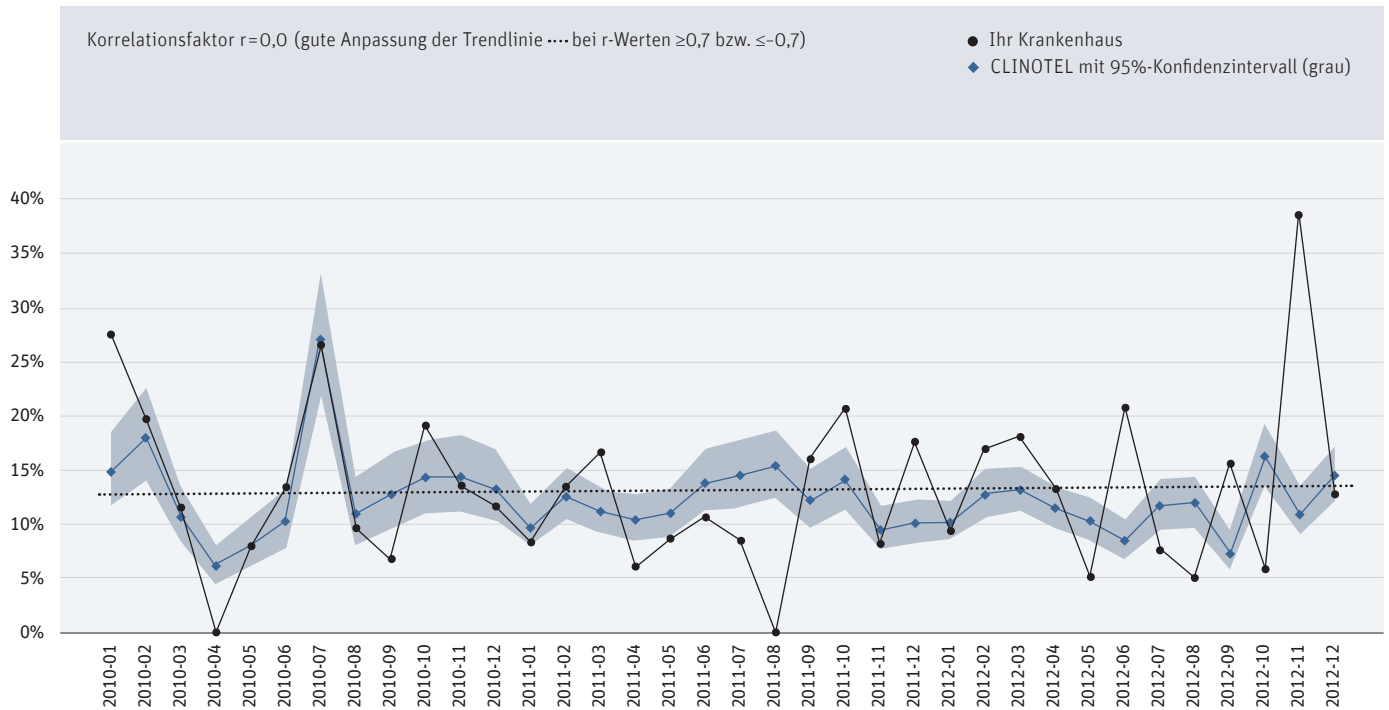


Abb. 18:
Risikoadjustierte Krankenhausletalität (%) ambulant erworbene Pneumonie: Funnel Plot
Daten 2012 | 10.379 Patienten

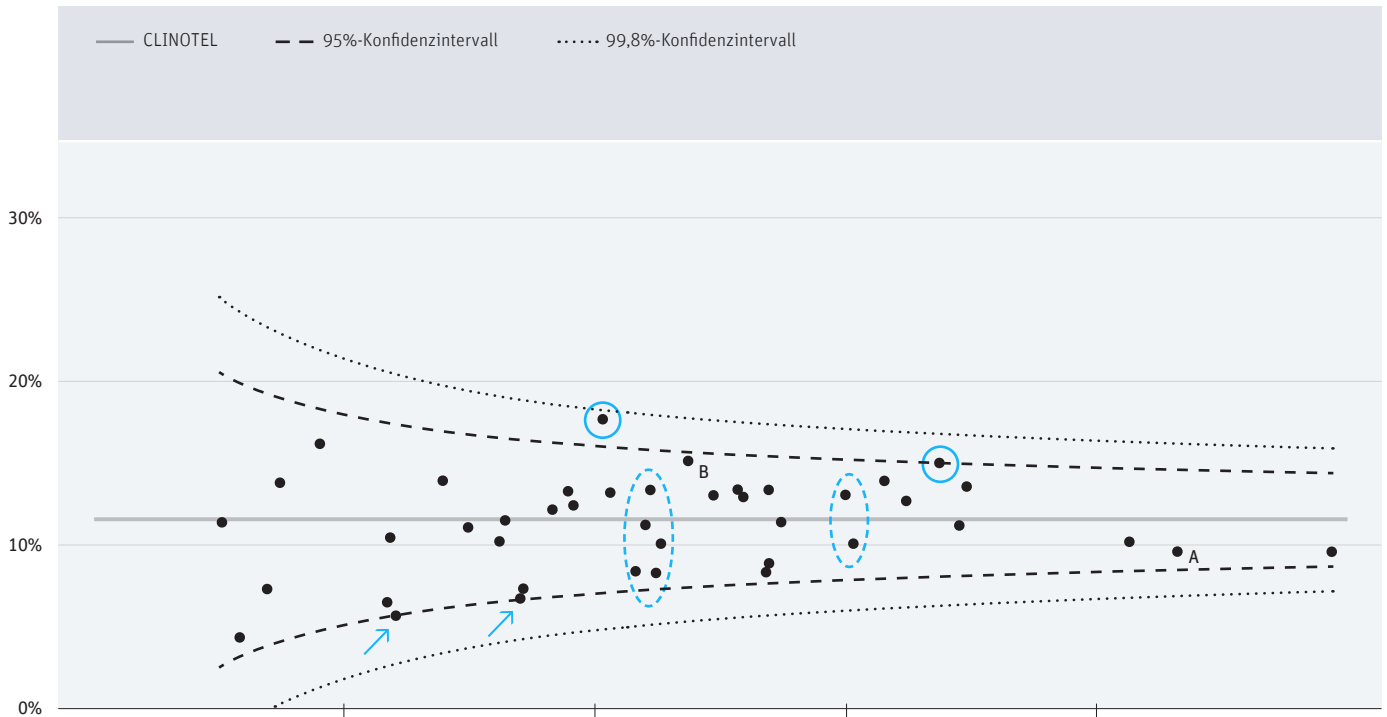


Abb. 19:

Krankenhausletalität ambulant erworbene Pneumonie: Variable Life Adjusted Display (VLAD)

Krankenhaus A | Daten 2012 | 432 Patienten

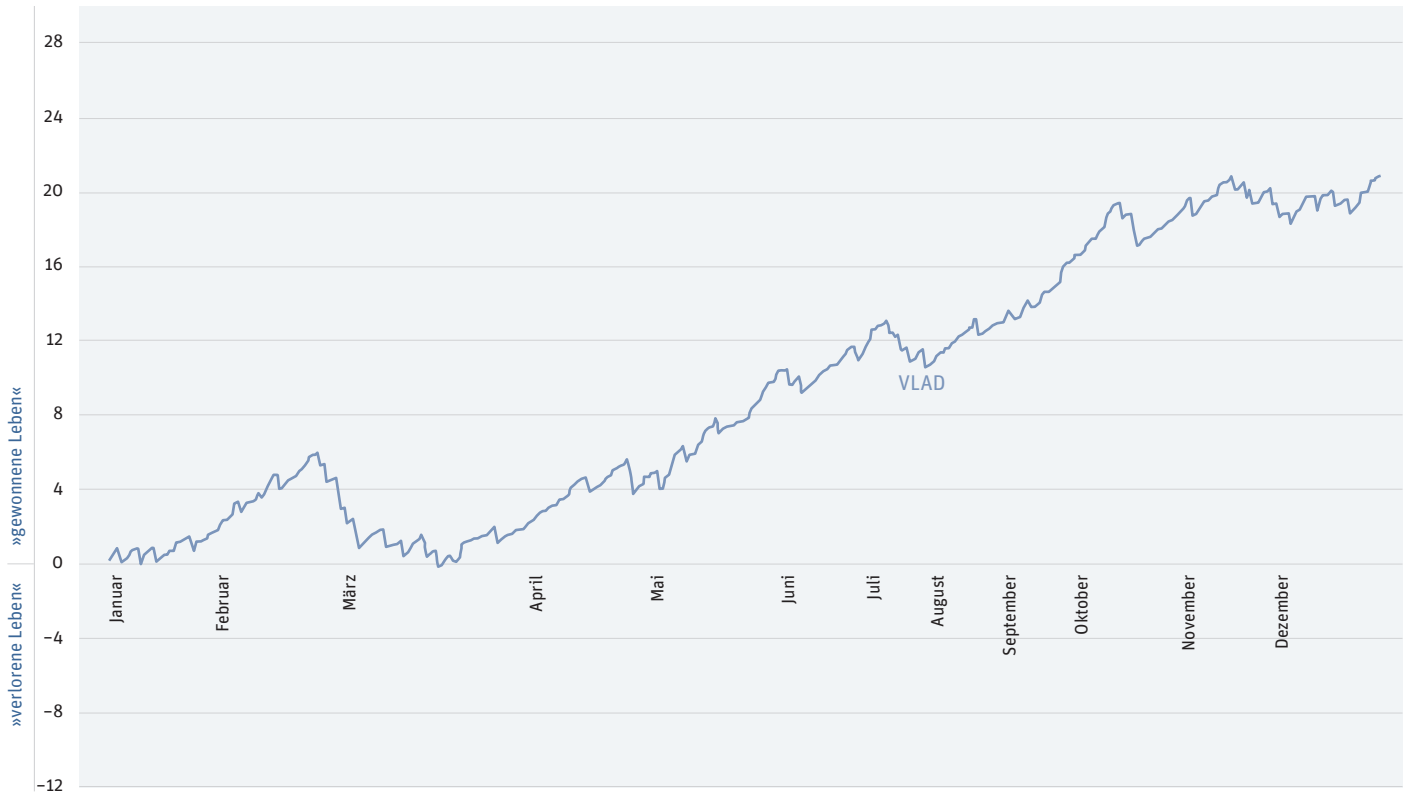


Abb. 20:
Vergrößerte Ausschnitte von zwei Abschnitten zur Erläuterung von VLAD-Eigenschaften

Krankenhausletalität ambulant erworbene Pneumonie: Variable Life Adjusted Display (VLAD)
Krankenhaus A | Daten 2012 | 432 Patienten

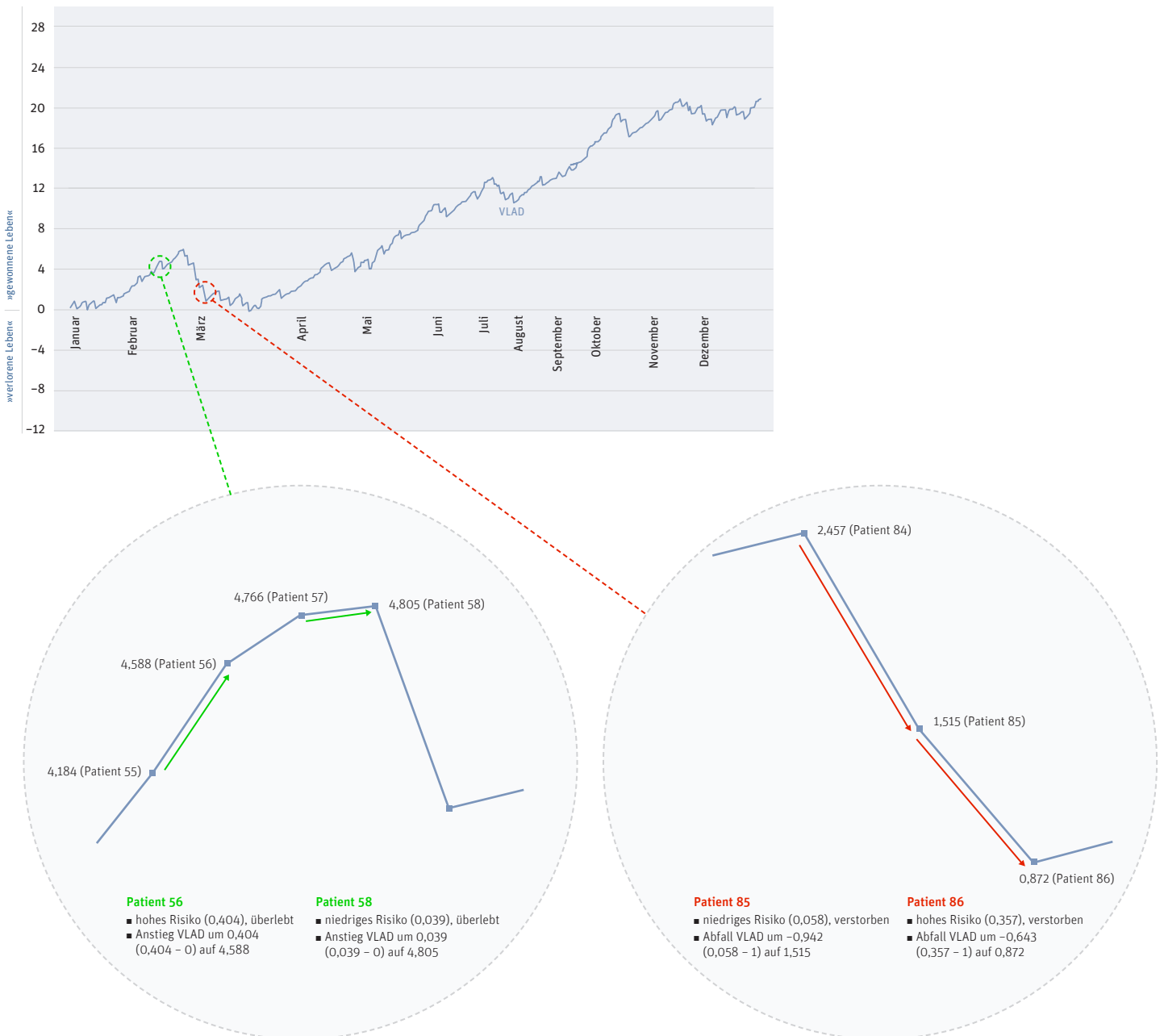


Abb. 21:

Krankenhausletalität ambulant erworbene Pneumonie: Variable Life Adjusted Display (VLAD)
Krankenhaus A | Daten 2012 | 432 Patienten

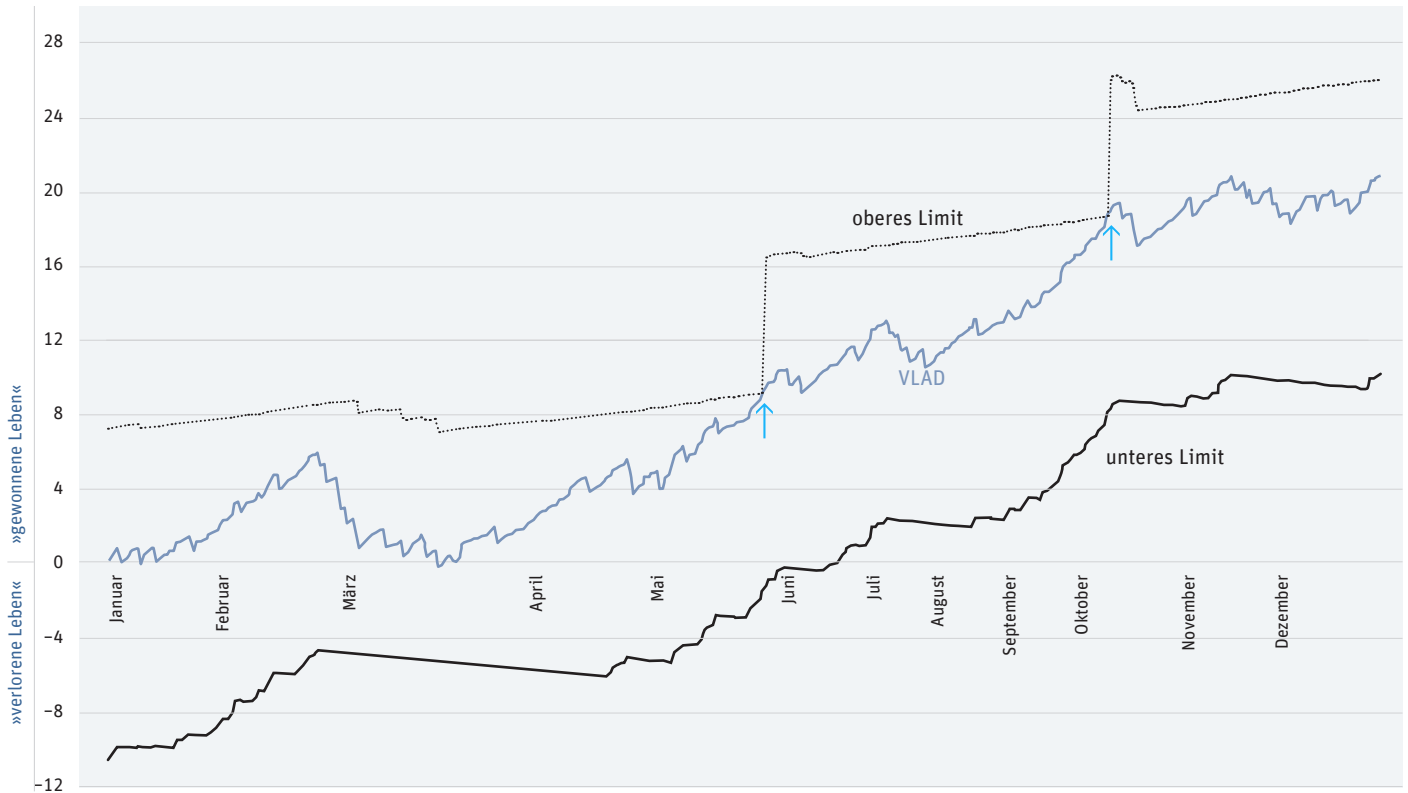


Abb. 22:

Krankenhausletalität ambulant erworbene Pneumonie: Variable Life Adjusted Display (VLAD)
Krankenhaus B | Daten 2012 | 237 Patienten

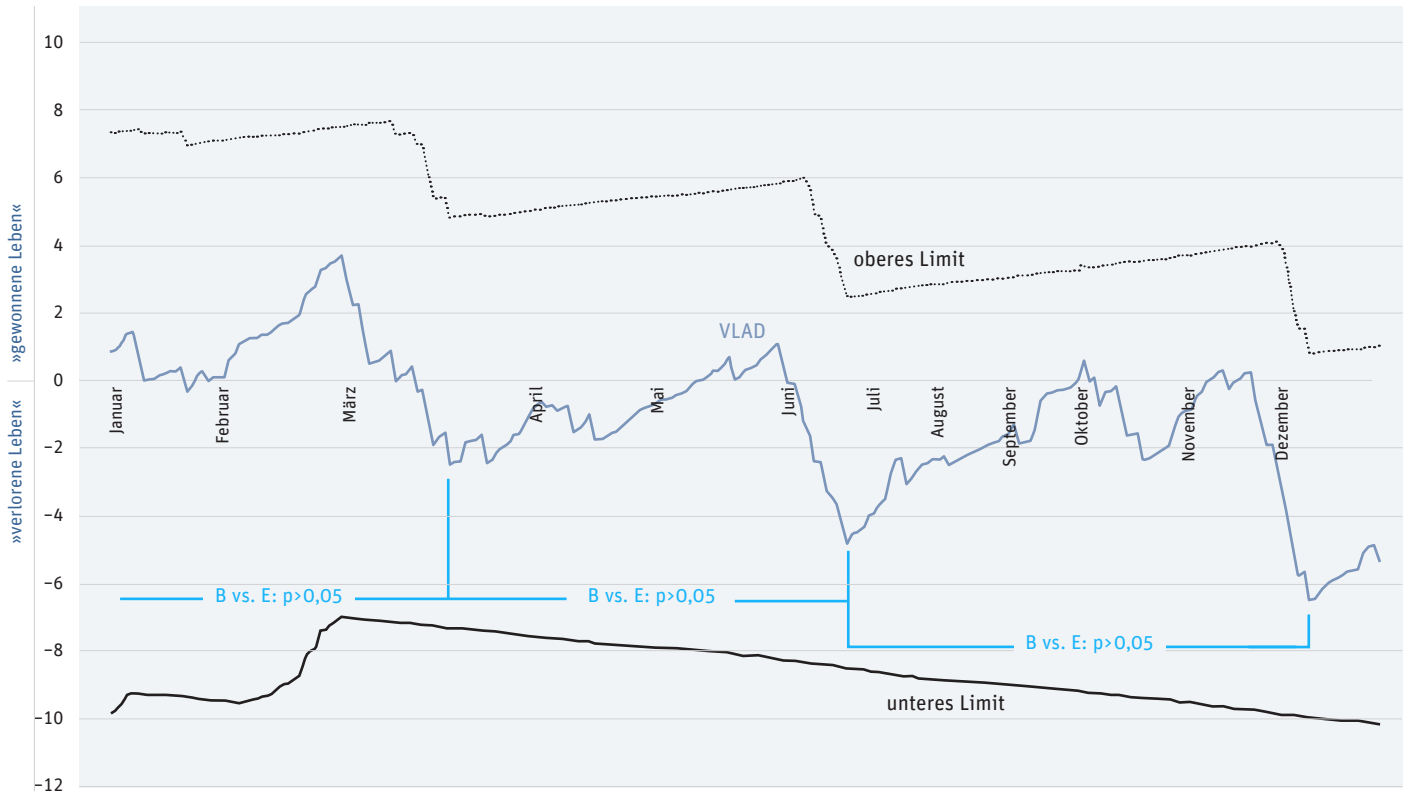
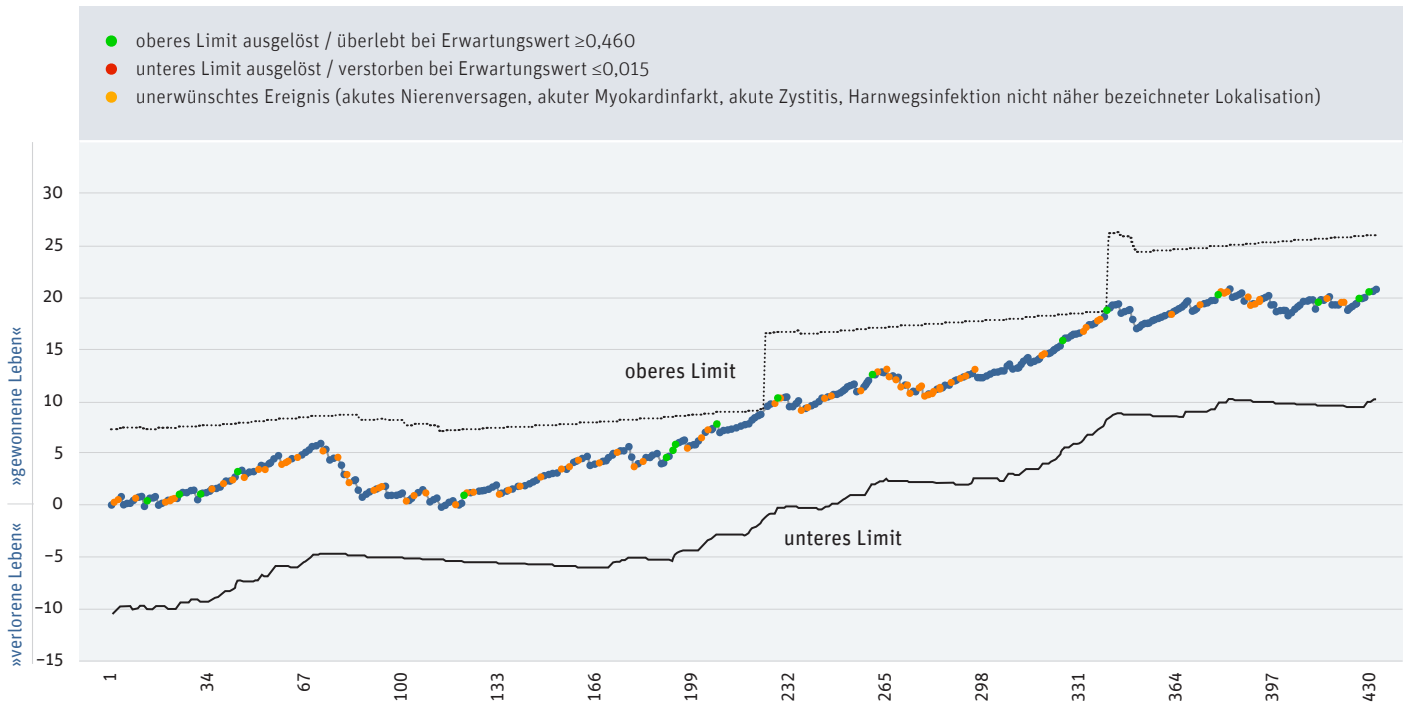


Abb. 23:
Weiterentwicklung Variable Life Adjusted Display (VLAD) nach Pagel et al. (2013)
Krankenhaus A | Daten 2012 | 432 Patienten



Tabellen

Tab. 1:
 Berechnung CUSUM (Krankenhaus A | Daten 2012 | 432 Patienten)

Patient	E	B	E - B	CUSUM
1	0,127	0	0,127	0,127
2	0,127	0	0,127	0,254
3	0,127	0	0,127	0,381
4	0,127	0	0,127	0,508
5	0,127	1	-0,873	-0,365
6	0,127	0	0,127	-0,238
7	0,127	0	0,127	-0,111
8	0,127	0	0,127	0,016
9	0,127	0	0,127	0,143
10	0,127	0	0,127	0,270
:	:	:	:	:
:	:	:	:	:
430	0,127	0	0,127	2,610
431	0,127	0	0,127	2,737
432	0,127	0	0,127	2,864

E erwartete Sterblichkeit (hier: Wert aus der externen Qualitätssicherung 12,7% im Jahr 2011)
 B Beobachtungswert: verstorben ja (1) nein (0)
 CUSUM kumulative Summe der E - B Werte
 (bei Patient Nr. 1 ist CUSUM = E - B)

Tab. 2:
Berechnung VLAD (Krankenhaus A | Daten 2012 | 432 Patienten)

Patient	E	B	E – B	VLAD
1	0,146	0	0,146	0,146
2	0,218	0	0,218	0,364
3	0,255	0	0,255	0,619
4	0,257	0	0,257	0,876
5	0,253	1	-0,747	0,129
6	0,045	0	0,045	0,174
7	0,074	0	0,074	0,248
8	0,211	0	0,211	0,459
9	0,242	0	0,242	0,701
10	0,090	0	0,090	0,791
:	:	:	:	:
:	:	:	:	:
430	0,008	0	0,008	20,649
431	0,136	0	0,136	20,785
432	0,098	0	0,098	20,883

E erwartete Sterblichkeit, für jeden Patienten individuell berechnet
 B Beobachtungswert: verstorben ja (1) nein (0)
 VLAD kumulative Summe der E – B Werte
 (bei Patient Nr. 1 ist VLAD = E – B)

Tab. 3:
Zusammenfassung der Ergebnisse: nicht risikoadjustierte Daten

Grafik	Krankenhaus A	Krankenhaus B
Balkendiagramm (2)		
Lage zu Gruppenwert	▼	▲
Lage zu externem Wert	▼	▲
Zeitreihe (6–9)		
Trendlinie: Verlauf	fällt	steigt
Trendlinie: Anpassung	ungenügend	ungenügend
Einfluss Fallzahlen	ja	ja
Wert(e) > 95%-KI	ja	ja
Wert(e) < 95%-KI	ja	ja
Trendlinie: innerhalb 95%-KI	ja (a)	nein (b)
Serie > 95%-KI	nein	ja
Serie < 95%-KI	ja	nein
p-Regelkarte (10–11)		
Wert(e) > obere Warngrenze	nein	ja
Wert(e) > obere Kontrollgrenze	nein	ja
Wert(e) < obere Warngrenze	nein	ja
Wert(e) < obere Kontrollgrenze	nein	ja
Signal: Verschiebung	nein	nein
Signal: zwei von drei	nein	nein
Signal: Trend	nein	nein
Cumulative Sum (CUSUM) (12–13)		
gewonnene Leben	ja	nein
verlorene Leben	nein	ja
Signal(e): oberes Limit	ja	nein
Signal(e): unteres Limit	ja	ja
guter Lauf ohne Signal	ja	ja
schlechter Lauf ohne Signal	ja	ja
(a) 33 von 36 Monate innerhalb 95%-KI (b) 12 von 36 Monate innerhalb 95%-KI		

Tab. 4:
Zusammenfassung der Ergebnisse: risikoadjustierte Daten

Grafik	Krankenhaus A	Krankenhaus B
Balkendiagramm: beobachtet versus erwartet (14)		
beobachtet \leq erwartet	ja	nein
beobachtet \leq erwartet 95%-KI	ja	nein
beobachtet $>$ erwartet	nein	ja
beobachtet $>$ erwartet 95%-KI	nein	nein
Balkendiagramm: risikoadjustierte Krankenhausletalität (15)		
Lage zu Gruppenwert	▼	▲
Lage zu externem Wert	(a)	(a)
Zeitreihe (16–17)		
Wert(e) $>$ 95%-KI	ja	ja
Wert(e) $<$ 95%-KI	ja	ja
Trendlinie: innerhalb 95%-KI	nein (b)	nein (b)
Serie $>$ 95%-KI	nein	ja
Serie $<$ 95%-KI	ja	ja
Funnel Plot (18)		
Wert(e) $>$ oberes 99,8%-Limit	nein	nein
Wert(e) $>$ oberes 95%-Limit	nein	nein
Wert(e) $<$ unteres 95%-Limit	nein	nein
Wert(e) $<$ unteres 99,8%-Limit	nein	nein
Variable Life Adjusted Display (VLAD) (19–22)		
gewonnene Leben	ja	nein
verlorene Leben	nein	ja
Signal(e): oberes Limit	ja	nein
Signal(e): unteres Limit	nein	nein
guter Lauf ohne Signal	ja	ja
schlechter Lauf ohne Signal	ja	ja
(a) nicht vorhanden (b) 25 von 36 Monaten innerhalb 95%-KI		

Tab. 5:
Falltabelle (Auszug)
Krankenhaus A | Daten 2012 | 432 Patienten

Lfd. Nr.	Fall Aufnahme	Entlassung	expected	tot	VLAD	O-Limit	U-Limit	Ereignis
335			0,071	n	17,503	18,579	6,831	
336			0,060	n	17,563	18,588	6,891	
337			0,305	n	17,868	18,624	7,196	Harnwegsinfektion, Lokalisation nicht näher bezeichnet
338			0,131	n	17,999	18,643	7,327	Harnwegsinfektion, Lokalisation nicht näher bezeichnet
339			0,197	n	18,196	18,669	7,524	
340			0,719	n	18,915	18,707	8,243	überlebt bei Erwartungswert $\geq 0,46$
340								oberes Limit
340								akuter Myokardinfarkt
341			0,123	n	19,038	26,222	8,366	
342			0,278	n	19,316	26,256	8,644	
343			0,080	n	19,396	26,268	8,724	
344			0,083	n	19,479	26,280	8,807	
345			0,172	j	18,651	25,941	8,787	

Literaturverzeichnis

- Altman DG, Machin D, Bryant T, Gardner MJ (2000).** Statistics with confidence. 2000 (Second Edition). British Medical Journal Books
- Andrianopoulos N, Jolley D, Evans SM, Brand CA, Cameron PA (2012).** Application of Variable Life Adjusted Displays (VLAD) on Victorian Admitted Episodes Dataset (VAED). *BMC Health Serv Res.* 2012 Aug 28; 12:278.
PubMed-ID: [22925089](#)
- Ash A, Shwartz M, Peköz EA (2003).** Comparing outcomes across providers. In: Risk Adjustment for Measuring Health Care Outcomes. Iezzoni LI (Editor). 2003 (Third Edition). Health Administrative Press
- Bateson G (1972).** Form, substance and difference. In: Steps to an ecology of mind. Bateson G (Hrsg.). 1972. Bantane Verlag. Deutsche Ausgabe: Form, Substanz und Differenz. In: Ökologie des Geistes. Bateson G (Hrsg.). 1985. Suhrkamp Verlag
- Becker A, Raskop AM, Beck U (2003).** CLINOTEL-Krankenhausverbund: Optimale Kodierung erfordert optimale Unterstützung. *das Krankenhaus.* 2003; 95 (6): 463–468.
Quelle: http://www.clinotel.de/files/art_2003-06-00-das-krankenhaus.pdf (letzte Einsicht 28.08.2012)
- Becker A, Mantke R, Beck U (2005).** Qualitätssicherung mit Routinedaten im CLINOTEL-Krankenhausverbund. *das Krankenhaus.* 2005; 97 (12): 1093–1102.
Quelle: http://www.clinotel.de/files/art_2005-12-00-das-krankenhaus.pdf (letzte Einsicht 28.08.2012)
- Becker A, Beck U, Pfeuffer B, Mantke R (2006).** Qualitätssicherung mit Routinedaten – Ergebnisqualität und Kosten. *das Krankenhaus.* 2006; 98 (9): 748–755.
Quelle: http://www.clinotel.de/files/art_2006-09-00-das-krankenhaus_1.pdf (letzte Einsicht 28.08.2012)
- Becker A (2012).** Present-On-Admission-Kennzeichen (POA) für administrative Routinedaten in Krankenhäusern. Literaturübersicht und Handlungsempfehlungen. Eine Ausarbeitung für die Gesellschaft für Qualitätsmanagement in der Gesundheitsversorgung e.V. (GQMG).
Quelle: http://gqmg.de/gqmg_about/Position_Empfehlung.htm (letzte Einsicht 28.08.2012)
- Becker A, Schwacke H, Perings C, Kamp T (2012a).** Qualitätssicherung mit Routinedaten (QSR) in der Kardiologie. *Interdisciplinary Contributions to Hospital Management: Medicine, Patient Safety and Economics.* 02.07.2012 #001. Quelle: <http://www.clinotel-journal.de/article-id-001.html> (letzte Einsicht 23.05.2013)
- Becker A, Ochs G, Thies C, Lefering R (2012b).** Qualitätssicherung mit Routinedaten (QSR) in der Neurologie. *Interdisciplinary Contributions to Hospital Management: Medicine, Patient Safety and Economics.* 02.07.2012 #004. Quelle: <http://www.clinotel-journal.de/article-id-004.html> (letzte Einsicht 23.05.2013)
- Becker A, Eissler U (2013).** Die standardisierte primäre Sectorate (SPSR) und ihre Anwendung im Qualitätsmanagement und für Krankenhausvergleiche. Prädiktoren der primären Sectio als Beitrag zur Versachlichung einer komplexen Diskussion. *Interdisciplinary Contributions to Hospital Management: Medicine, Patient Safety and Economics.* 17.04.2013 #010.
Quelle: <http://www.clinotel-journal.de/article-id-010.html> (letzte Einsicht 24.05.2013)
- Becker A (2013a).** Qualitätskriterien erfolgreicher Morbiditäts- und Mortalitätskonferenzen. *Interdisciplinary Contributions to Hospital Management: Medicine, Patient Safety and Economics.* 23.10.2013 #015.
Quelle: <http://www.clinotel-journal.de/article-id-015.html> (letzte Einsicht 03.11.2013)
- Benchimol EI, Manuel DG, To T, Griffiths AM, Rabeneck L, Guttman A (2011).** Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. *J Clin Epidemiol.* 2011; 64 (8): 821–829. PubMed-ID: [21194889](#)

- Bender R, Lange S (2007).** Was ist ein Konfidenzintervall? Deutsche Medizinische Wochenschrift. 2007; 132 Suppl. 1: e17–e18. PubMed-ID: [17530586](#)
- Bragg F, Cromwell DA, Edozien LC, Gurol-Urganci I, Mahmood TA, Templeton A, van der Meulen JH (2010).** Variation in rates of caesarean section among English NHS trusts after accounting for maternal and clinical risk: cross sectional study. BMJ. 2010 Oct 6; 341: 1–8. PubMed-ID: [20926490](#)
- Bratzler DW, Normand SL, Wang Y, O'Donnell WJ, Metersky M, Han LF, Rapp MT, Krumholz HM (2011).** An administrative claims model for profiling hospital 30-day mortality rates for pneumonia patients. PLoS One. 2011; 6(4): e17401. Quelle: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3075250/> (letzte Einsicht: 26.05.2013)
- Cook DA, Duke G, Hart GK, Pilcher D, Mullany D (2008).** Review of the application of risk-adjusted charts to analyse mortality outcomes in critical care. Crit Care Resusc. 2008; 10 (3): 239–251. PubMed-ID: [18798724](#)
- Coory M, Duckett S, Sketcher-Baker K (2008).** Using control charts to monitor quality of hospital care with administrative data. Int J Qual Health Care. 2008; 20 (1): 31–39. PubMed-ID: [18065757](#)
- Dimick JB, Welch HG, Birkmeyer JD (2004).** Surgical Mortality as an Indicator of Hospital Quality. The Problem With Small Sample Size. JAMA. 2004; 292 (7): 847–851. PubMed-ID: [15315999](#)
- DIN EN 15224 (2012).** Dienstleistungen in der Gesundheitsversorgung – Qualitätsmanagementsysteme – Anforderungen nach EN ISO 9001:2008. Deutsche Fassung EN 15224:2012. Ausgabedatum: 2012–12. DIN Deutsches Institut für Normung
- Dixon-Woods M, Bosk CL, Aveling EL, Goeschel CA, Pronovost PJ (2011).** Explaining Michigan: Developing an Ex Post Theory of a Quality Improvement Program. Milbank Mem Fund Q. 2011; 89 (2): 167–205. PubMed-ID: [21676020](#)
- Duckett SJ, Coory M, Sketcher-Baker K (2007).** Identifying variations in quality of care in Queensland hospitals. Med J Aust. 2007; 187 (10): 571–575. PubMed-ID: [18021046](#)
- Duclos A, Touzet S, Soardo P, Colin C, Peix JL, Lifante JC (2009).** Quality monitoring in thyroid surgery using the Shewhart control chart. Br J Surg. 2009; 96 (2): 171–174. PubMed-ID: [19160350](#)
- du Prel JB, Hommel G, Röhrig B, Blettner M (2009).** Konfidenzintervall oder p-Wert? Teil 4 der Serie zur Bewertung wissenschaftlicher Publikationen. Dtsch Arztebl Int 200. 2009; 106 (19): 335–339. Quelle: <http://www.aerzteblatt.de/archiv/64533/Konfidenzintervall-oder-p-Wert-Teil-4-der-Serie-zur-Bewertung-wissenschaftlicher-Publikationen> (letzte Einsicht: 29.05.2013)
- EQS (2011).** Ambulant erworbene Pneumonie. Bundesauswertung zum Erfassungsjahr 2011. Erstellt am 31.05.2012. AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH Göttingen. Quelle: <https://www.sgg.de/ergebnisse/leistungsbereiche/ambulant-erworbene-pneumonie.html> (letzte Einsicht 24.05.2013)
- Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, Coley CM, Marrie TJ, Kapoor WN (1997).** A prediction rule to identify low-risk patients with community-acquired pneumonia. N Engl J Med. 1997; 336 (4): 243–250. Quelle: <http://www.nejm.org/doi/full/10.1056/NEJM199701233360402> (letzte Einsicht 26.05.2013)
- Flaatten H (2005).** Effects of a major structural change to the intensive care unit on the quality and outcome after intensive care. Qual Saf Health Care. 2005; 14 (4): 270–272. PubMed-ID: [16076791](#)
- Fusco D, Barone AP, Sorge C, D'Ovidio M, Stafoggia M, Lallo A, Davoli M, Perucci CA (2012).** P.Re.Val.E.: outcome research program for the evaluation of health care quality in Lazio, Italy. BMC Health Serv Res. 2012; Jan 27; PubMed-ID: [22283880](#)

Goeschel CA (2011). Nursing leadership at the crossroads: evidence-based practice ‘Matching Michigan-minimizing catheter related blood stream infections’. *Nurs Crit Care.* 2011; 16 (1): 36–43. PubMed-ID: [21199553](#)

Grunkemeier GL, Wu YX, Furnary AP (2003). Cumulative Sum Techniques for Assessing Surgical Results. *Ann Thorac Surg.* 2003; 76 (3): 663–637. PubMed-ID: [12963172](#)

Guest RV, Chandrabalan VV, Murray GD, Auld CD (2012). Application of Variable Life Adjusted Display (VLAD) to Risk-Adjusted Mortality of Esophagogastric Cancer Surgery. *World J Surg.* 2012; 36 (1): 104–108. PubMed-ID: [21964818](#)

Hart MK, Lee KY, Hart RF, Robertson JW (2003). Application of Attribute Control Charts to Risk-Adjusted Data for Monitoring and Improving Health Care Performance. *Qual Manag Health Care.* 2003; 12 (1): 5–19. PubMed-ID: [12593370](#)

Henderson GR, Mead GE, van Dijke ML, Ramsay S, McDowall MA, Dennis M (2008). Use of statistical process control charts in stroke medicine to determine if clinical evidence and changes in service delivery were associated with improvements in the quality of care. *Qual Saf Health Care.* 2008; 17 (4): 301–306. PubMed-ID: [18678730](#)

Hüsler J, Zimmermann H (2006). Statistische Prinzipien für medizinische Projekte. 2006 (4., vollständig überarbeitete und erweiterte Auflage). Verlag Hans Huber,

Kottner J, Hauss A (2013). Vergleichende Qualitätsmessungen Teil 2: Regelkarten. *Pflege.* 2013; 26 (2): 119–127

Kreienbrock L, Pigeot I, Ahrens W (2012). Epidemiologische Methoden. 2012 (5. Auflage). Springer-Verlag

Lack N, Gerhardinger U (2009). Qualitätsvergleiche mit FunnelpLOTS – Plädoyer für eine einheitliche Methodik. *Z Evid Fortbild Qual Gesundheitswes.* 2009; 103 (8): 536–541. PubMed-ID: [19998782](#)

Leitlinie Pneumonie (2009). Epidemiologie, Diagnostik, antimikrobielle Therapie und Management von erwachsenen Patienten mit ambulant erworbenen tiefen Atemwegsinfektionen (akute Bronchitis, akute Exazerbation einer chronischen Bronchitis, Influenza und andere respiratorische Virusinfektionen) sowie ambulant erworbener Pneumonie. Leitlinie der Paul-Ehrlich-Gesellschaft für Chemotherapie, der Deutschen Gesellschaft für Pneumologie und Beatmungsmedizin, der Deutschen Gesellschaft für Infektiologie und des Kompetenznetzwerks CAPNETZ. Registernummer 082-001. Stand 01.07.2009, gültig bis 30.06.2014. Quelle: <http://www.awmf.org/leitlinien/detail/ll/082-001.html> (letzte Einsicht 24.05.2013)

Mayer EK, Bottle A, Aylin P, Darzi AW, Vale JA, Athanasiou T (2011). What is the role of risk-adjusted funnel plots in the analysis of radical cystectomy volume-outcome relationships? *BJU Int.* 2011; 108 (6): 844–850. PubMed-ID: [21884357](#)

Mohammed MA, Rathbone A, Myers P, Patel D, Onions H, Stevens A (2004). An investigation into general practitioners associated with high patient mortality flagged up through the Shipman inquiry: retrospective analysis of routine data. *BMJ.* 2004; 328 (7454): 1474–1477. PubMed-ID: [15205291](#)

Mohammed MA, Worthington P, Woodall WH (2008). Plotting basic control charts: tutorial notes for healthcare practitioners. *Qual Saf Health Care.* 2008; 17 (2): 137–145. PubMed-ID: [18385409](#)

Mohammed MA, Worthington P (2013). Why traditional statistical process control charts for attribute data should be viewed alongside an xmr-chart. *BMJ Qual Saf.* 2013; 22 (3): 263–269. PubMed-ID: [23104897](#)

Mohammed MA, Panesar JS, Laney DB, Wilson R (2013). Statistical process control charts for attribute data involving very large sample sizes: a review of problems and solutions. *BMJ Qual Saf.* 2013; 22 (4): 362–368. PubMed-ID: [23365140](#)

- Muche R, Ring C, Ziegler C (2005).** Entwicklung und Validierung von Prognosemodellen auf Basis der logistischen Regression. 2005. Shaker Verlag
- Noyez L (2009).** Control charts, Cusum techniques and funnel plots. A review of methods for monitoring performance in healthcare. *Interact Cardiovasc Thorac Surg.* 2009; 9 (3): 494–499. PubMed-ID: [19509097](#)
- Pagel C, Prost A, Nair N, Tripathy P, Costello A, Utley M (2012).** Monitoring mortality trends in low-resource settings. *Bull World Health Organ.* 2012; 90 (6): 474–476. PubMed-ID: [22690038](#)
- Pagel C, Utley M, Crowe S, Witter T, Anderson D, Samson R, McLean A, Banks V, Tsang V, Brown K (2013).** Real time monitoring of risk-adjusted paediatric cardiac surgery outcomes using variable life-adjusted display: implementation in three UK centres. *Heart.* 2013; 99 (13): 1445–1450. PubMed-ID: [23564473](#)
- Perla RJ, Provost LP, Murray SK (2011).** The run chart: a simple analytical tool for learning from variation in healthcare processes. *BMJ Qual Saf.* 2011; 20 (1): 46–51. PubMed-ID: [21228075](#)
- Poelaert J, Schuepfer G, Umnus A, Bauer M, Schleppers A (2007).** »Statistische Prozesskontrolle« als Monitoring-Tool zur Evaluierung von Reorganisationsmaßnahmen. Untersuchung auf einer Intensivstation. *Anaesthesist.* 2007; 56 (6): 599–603.
- Poloniecki J, Valencia O, Littlejohns P (1998).** Cumulative risk adjusted mortality chart for detecting changes in death rate: observational study of heart surgery. *BMJ.* 1998; 316 (7146): 1697–1700. PubMed-ID: [9614015](#)
- Queensland Government.** Health Service and Clinical Innovation Division. Patient Safety Unit.
Quelle: <http://www.health.qld.gov.au/psu/health-professionals.asp> (letzte Einsicht 30.08.2013)
- Roberts G, Tang CB, Harvey M, Kadirkamanathan S (2012).** Real-time outcome monitoring following oesophagectomy using cumulative sum techniques. *World J Gastrointest Surg.* 2012; 4 (10): 234–237. PubMed-ID: [23443533](#)
- Sherlaw-Johnson C, Lovegrove J, Treasure T, Gallivan S (2000).** Likely variations in perioperative mortality associated with cardiac surgery: when does high mortality reflect bad practice? *Heart.* 2000; 84 (1): 79–82. PubMed-ID: [10862596](#)
- Sherlaw-Johnson C (2005).** A Method for Detecting Runs of Good and Bad Clinical Outcomes on Variable Life-Adjusted Display (VLAD) Charts. *Health Care Manag Sci.* 2005; 8 (1): 61–65. PubMed-ID: [15782513](#)
- Spiegelhalter D, Grigg O, Kinsman R, Treasure T (2003).** Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care.* 2003; 15 (1): 7–13. PubMed-ID: [12630796](#)
- Spiegelhalter DJ (2005).** Funnel plots for comparing institutional performance. *Stat Med.* 2005; 24 (8): 1185–1202. PubMed-ID: [15568194](#)
- Steiner SH, Cook RJ, Farewell VT, Treasure T (2000).** Monitoring surgical performance using risk-adjusted cumulative sum charts. *Biostatistics.* 2000; 1 (4): 441–452. PubMed-ID: [12933566](#)
- Tan HB, Cross SF, Goodacre SW (2005).** Application of variable life adjusted display (VLAD) in early detection of deficiency in trauma care. *Emerg Med J.* 2005; 22 (10): 726–728. PubMed-ID: [16189037](#)
- Tennant R, Mohammed MA, Coleman JJ, Martin U (2007).** Monitoring patients using control charts: a systematic review. *Int J Qual Health Care.* 2007; 19 (4): 187–194. PubMed-ID: [17545672](#)

Victorian Government. Health Information.

Quelle: <http://www.health.vic.gov.au/psi/auspsi/how-to-use-the-auspsi> (letzte Einsicht 30.08.2013)

Winkel P, Zhang NF (2007). Statistical Development of Quality in Medicine. 2007. John Wiley & Sons Ltd

Woodall WH (2006). The Use of Control Charts in Health-Care and Public-Health Surveillance. J Qual Tech. 2006; 38 (2): 89–104

Manuskriptdaten

Interessenkonflikt

Der Autor erklärt, dass kein Interessenkonflikt besteht. Er ist Geschäftsführer der CLINOTEL Krankenhausverband gemeinnützige GmbH. In dieser Eigenschaft vertritt er das Unternehmen bei folgenden Institutionen und Fachgesellschaften: Aktionsbündnis Patientensicherheit e.V., Gesellschaft für Qualitätsmanagement in der Gesundheitsversorgung e.V., Deutsche Gesellschaft für Qualität e.V., European Foundation for Quality Management, International Society for Quality in Health Care, Deutsche Gesellschaft für Medizinische Dokumentation, Informatik und Statistik e.V., Verband der Krankenhausesdirektoren Deutschlands e.V. und Bundesverband Pflegemanagement e.V. Prof. Becker ist persönliches Mitglied der Zertifikatskommission »Medizinische Informatik« der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) e.V. und Fellow der Fachgesellschaft International Society for Quality in Health Care.

Bearbeitung

Manuskript eingereicht am 11.06.2013,
überarbeitete Fassung angenommen am 10.09.2013

Zitierung

Becker A. Von Daten zu Informationen zu Entscheidungen – wie können statistische Grafiken die klinische Qualitätsbewertung im Krankenhaus unterstützen? Eine Übersicht und praktische Anwendung am Beispiel der Krankenhaussterblichkeit bei ambulant erworbener Pneumonie. *Interdisciplinary Contributions to Hospital Management: Medicine, Patient Safety and Economics*. 18.12.2013 #017.
<http://www.clinotel-journal.de/article-id-017.html>

Danksagung

Der Autor bedankt sich ganz besonders bei den drei Reviewern für den sehr konstruktiven und hilfreichen Begutachtungsprozess. Mein Dank gilt auch Dr. Alan Poots PhD (Imperial College London) und Dr. Kris Denhaerynck PhD (University of Basel, Institute of Nursing Science) für die hilfreiche Unterstützung bei der Berechnung der Kontrollgrenzen für die VLAD-Grafiken. Ebenso bedanken möchte ich mich bei Dr. Christina Pagel PhD (University College London) für die wichtigen Literaturhinweise, die einen guten Überblick zur Berechnung der Kontrollgrenzen bei VLAD-Grafiken gegeben haben.

Autoren

Prof. Dr. med. Andreas Becker

Geschäftsführer

CLINOTEL Krankenhausverbund gemeinnützige GmbH

Riehler Straße 36

50668 Köln

www.clinotel.de

