© 2012 Verlag Hans Huber, Hogrefe AG, Bern
Pflege 2012; 25 (3): 197–206 DOI 10.1024/1012-5302/a000202

Clinical Research Center for Hair and Skin Science, Klinik für Dermatologie, Venerologie und Allergologie, Berlin
Jan Kottner (PD Dr.)

Vergleichende Qualitätsmessungen Teil 1: Laufdiagramme

- **Was ist (zu dieser Thematik) schon bekannt?**
  Qualitätsmessungen basieren im Allgemeinen auf Qualitätskennzahlen. Um vergleichende Bewertungen vornehmen zu können, müssen systematische und zufällige Messfehler minimiert werden.
- **Was ist neu?**
  Laufdiagramme stellen einfache grafische Möglichkeiten dar, fortlaufende Prozesse und deren Ergebnisse mittels Kennzahlen über die Zeit zu analysieren. Laufdiagramme können für Gruppen und Einzelpersonen verwendet werden.
- **Welche Konsequenzen haben die Ergebnisse für die Pflegepraxis?**
  Die kontinuierliche Darstellung und Interpretation gesammelter Qualitätsdaten in der Praxis hilft, Qualitätsverbesserungen, Verschlechterungen oder Stabilität zu erkennen.

Qualitätsbeurteilungen basieren häufig auf Vergleichen von Qualitätskennzahlen. Dabei werden zwei Wege verfolgt: die Kennzahlen werden über eine Zeitspanne (Selbstvergleich) oder gleichzeitig mit anderen Organisationen verglichen (Fremdvergleich). Neben Risikoadjustierungen und Standardisierungen von Messungen müssen zufallsbedingte Variationen in empirisch ermittelten Kennzahlen adäquat berücksichtigt werden. In der Theorie der Statistischen Prozesskontrolle (statistical process control) werden innerhalb komplexer Prozesse natürlicherweise zu erwartende allgemeine Variabilität und besondere unnatürlich bedingte Variabilität in Messwerten unterschieden (common cause und special cause variation). Prozesse innerhalb statistischer Kontrolle sind stabil und vorhersagbar. Befinden sich Prozesse außerhalb statistischer Kontrolle, dann sollte das Management die Ursachen dafür identifizieren und handeln. Im vorliegenden Artikel werden Laufdiagramme als einfache Werkzeuge, um Kennzahlen von Prozessen und Ergebnissen über die Zeit darzustellen, vorgestellt. Diese eignen sich zum Selbstvergleich. In den nachfolgenden zwei Teilen dieser Serie werden Regelkarten und Funnel Plots beschrieben.

Schlüsselbegriffe: Pflege, Qualitätskontrolle, Qualitätsverbesserung, Qualitätsindikatoren, Laufdiagramme

Manuskript erstmals eingereicht am 22.08.2011
Endgültige Fassung eingereicht am 03.01.2012

Einleitung

Das Thema Qualität in der Pflege und Gesundheitsversorgung spielt seit Jahrzehnten in Praxis, Ausbildung und Forschung eine bedeutende Rolle. Auch wenn es keine einheitlichen Begriffsbestimmungen und Wahrnehmungen von Qualität gibt (Attree, 2001; Campbell, Roland & Buetow, 2000), so charakterisiert man damit im Pflege- und Gesundheitsbereich häufig eine bestimmte Güte von Strukturen, Prozessen, Interaktionen oder Ergebnissen. Klassischerweise wird ein für erstrebenswert oder für gut befundener Zustand («Soll») mit der tatsächlichen Situation («Ist») verglichen. So zum Beispiel definierte das Institute of Medicine (USA) Qualität als «the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge» (Committee on Quality Health Care in America & Institute of Medicine, 2001: 232).

Qualitätsbeurteilungen können unter anderem auf quantifizierbaren Ausprägungen (Kennzahlen) von Qualitätsindikatoren beruhen. Qualitätsindikatoren sind explizit definierte Maße und Methoden, von denen angenommen wird, dass sie mit der tatsächlich erbrachten Versorgungsqualität zusammenhängen und diese durch Kennzahlen abbilden können (Ärztliches Zentrum für Qualität in der Medizin, 2009; Baartmans & Geng, 2006; Ovretveit, 2001). Qualitätsmessungen sind somit indirekt und beziehen sich auf Teilaspekte der gesamten komplexen Versorgungsrealität. Einen international anerkannten Referenzstandard für die Entwicklung und Bewertung von Qualitätsindikatoren gibt es nicht, doch es liegen zahlreiche weit entwickelte und etablierte Vorschläge und Verfahren vor (zum Beispiel Ärztliches Zentrum für Qualität in der Medizin, 2009; Reiter, Fischer, Kötting, Geraedts, Jäckel & Döbler, 2008; Rubin, Pronovost & Diette, 2001). International sind indikatorenbasierte pflegerische Qualitätsmesssysteme teilweise schon seit vielen Jahren etabliert oder werden entwickelt (Nakrem, Guttormsen, Harkless, Paulsen & Seim, 2009; West, 2010). Im deutschsprachigen Raum gibt es ebenfalls entsprechende Initiativen (Bundesministerium für Gesundheit; Bundesministerium für Familie, Senioren, Frauen und Jugend, 2011; Burla, Schaffert, Mylaeus & Rüesch, 2010).

https://econtent.hogrefe.com/doi/pdf/10.1024/1012-5302/a000202 - Saturday, May 04, 2024 1:38:30 AM - IP Address:18.188.70.255

https://econtent.hogrefe.com/doi/pdf/10.1024/1012-5302/a000202 - Saturday, May 04, 2024 1:38:30 AM - IP Address:18.188.70.255

Die auf Kennzahlen beruhenden Qualitätsbeurteilungen erfolgen im Allgemeinen auf zwei Wegen: (1) ermittelte Werte und Zahlen von leistungserbringenden Personen (zum Beispiel Case Manager) oder organisatorischen Einheiten (zum Beispiel Wohnbereiche, Krankenhäuser, ambulante Pflegedienste) werden miteinander über die Zeit verglichen, oder (2) die ermittelten Kennzahlen werden gleichzeitig mit anderen Personen oder Organisationen verglichen (Ovretveit, 2001; Wollersheim, Hermens, Hulscher, Braspenning, Ouwens, Schouten et al., 2007). Im ersten Fall geht man der Frage nach, ob beispielsweise einrichtungsinterne Prozesse und Ergebnisse stabil sind, oder ob es Veränderungen gibt («Selbstvergleich»). Im zweiten Fall möchte man wissen: Wo stehen wir im Vergleich zu anderen («Fremdvergleich»)? Unabhängig davon, welchen Weg man wählt, es geht immer um Vergleiche von Zahlen. Werden Unterschiede zwischen den erhobenen Werten festgestellt, dann kann das auf Veränderungen oder Unterschiede in der Qualität hindeuten. Bevor man jedoch zu dieser Schlussfolgerung kommt, müssen andere Faktoren, die ebenfalls die numerische Ausprägung von Kennzahlen beeinflussen können, ausgeschlossen werden. Dazu zählen: (1) Unterschiede zwischen den leistungsempfangenden Personen, die nicht unmittelbar durch die erbrachte Leistung beeinflussbar sind (case mix), (2) Unterschiede in der Messung und (3) Zufall (Mant, 2001).

Unterschiede zwischen Personengruppen

Der verzerrende Einfluss des ersten Faktors ist allgemein bekannt. Unterscheiden sich Personengruppen hinsichtlich ihres Risikoprofils oder anderer Eigenschaften, die Versorgungsprozesse und Ergebnisse beeinflussen können, dann dürfen diese nicht ohne Weiteres verglichen werden. Abhilfe schaffen sogenannte Risikoadjustierungen. Grundsätzlich kommen zwei Methoden zur Anwendung. Bei der einen Methode werden nur bestimmte vergleichbare Patienten, Bewohner oder Klienten verglichen, die nichtvergleichbaren werden ausgeschlossen. Mögliche Vorgehensweisen sind risikostandardisierte Fallkonstellationen und Stratifizierungen. Die andere Methode basiert auf multivariaten (und gegebenenfalls Mehrebenen-)Regressionsmodellen. Dabei werden Verhältnisse oder Differenzen zwischen den in der Gruppe tatsächlich beobachteten Ausprägungen zu den rechnerisch zu erwartenden Ausprägungen ermittelt (AQUA-Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH, 2010; Jones, Hirdes, Poss, Kelly, Berg, Fries et al., 2010; Szklo & Nieto, 2007).

Unterschiede in der Messung

Standardisierungen der Datenerhebungen vermindern systematische und zufällige Variationen. Nur bei geringen Fehleranteilen sind Kennzahlen vergleichbar. Zur Standardisierung gehören unter anderem: genaue Definitionen von Zählern und Nennern; Anwendung bestimmter diagnostischer Verfahren; Festlegung der Personen, die Daten erheben; Datenquellen; Zeit(punkte) der Datenerhebung usw. (Lloyd, 2004; Rubin et al., 2001).

Zufall

Wie jede empirisch ermittelte Zahl, so unterliegen auch Qualitätskennzahlen zufallsbedingten Fehlern (random error). Diese Fehler sind umso größer, je kleiner die betrachteten Fallzahlen sind. Um die Präzision einzelner Punktschätzungen zu erhöhen, ist es gängige Praxis, «kleine» Fallzahlen aus Qualitätsvergleichen auszuschließen. In der Literatur finden sich beispielsweise Mindestnennergrößen von 15 (Karon, Sainfort & Zimmerman, 1999), 20 (AQUA-Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH, 2010) oder 100 (Berlowitz, Anderson, Ash, Brandeis, Brand & Moskowitz, 1998) für externe Einrichtungsvergleiche. Nachteilig an diesem Vorgehen ist, dass die Beurteilung von «klein» subjektiv ist (Kottner, Dassen & Heinze, 2011). Eine höhere «Stabilität» von Kennzahlen kann auch durch die Zusammenfassung mehrerer Messungen erreicht werden (Berlowitz et al., 1998; Rantz, Hicks, Petroski, Madsen, Mehr, Conn et al., 2004). Doch auch hier gilt, dass es die Entscheidung, wie viele Messungen über welchen Zeitraum zu betrachten sind, eher pragmatisch als auf einer Theorie basierend, erfolgt.

Zielsetzung

Zur Entwicklung von Qualitätsindikatoren, zu Risikoadjustierungen und Messverfahren wurden bis heute zahlreiche Methodenpapiere und empirische Ergebnisse publiziert und diskutiert. Der Berücksichtigung des Zufallsfehlers (random error) in vergleichenden Qualitätsmessungen wurde bislang in der deutschsprachigen Literatur weniger Aufmerksamkeit geschenkt. Das Ziel des vorliegenden Beitrags ist es daher, praktikable Methoden aufzuzeigen, wie mit Schwankungen und Variabilität in standardisierten und risikoadjustierten Kennzahlen adäquat umgegangen werden kann. Dabei fokussieren die ersten beiden Teile dieser Beitragsserie auf Kennzahlenvergleiche innerhalb einer Organisation oder von einzelnen Personen über die Zeit (»Selbstvergleich»), der dritte Teil hingegen auf den Vergleich von Organisationen oder Einzelpersonen zu einem Zeitpunkt («Fremdvergleich»).

Statistische Prozesskontrolle

Eine mögliche Grundlage, sich zufällig bedingten Fehlern und Schwankungen

https://econtent.hogrefe.com/doi/pdf/10.1024/1012-5302/a000202 - Saturday, May 04, 2024 1:38:30 AM - IP Address:18.188.70.255

in Qualitätskennzahlen zu nähern, bietet die Theorie der Statistischen Prozesskontrolle (Statistical Process Control, SPC). Der englische Begriff «control» weist in diesem Zusammenhang eher auf Inhalte wie «Lenkung» oder «Steuerung» hin. Somit wird SPC im Deutschen auch mit «Statistischer Prozessregelung» übersetzt (Schmitz, 2004), doch im Folgenden wird die wortwörtliche Übersetzung beibehalten. Die SPC geht auf die Arbeiten von Walter A. Shewhart zurück und hat ihre Ursprünge in der industriellen Produktion der 1920er-Jahre (Deming, 1986; Shewhart, 1931, 1939). In den 1980er- und 1990er-Jahren wurde diese Methode zunehmend im medizinischen Kontext diskutiert und angewendet (Berwick, 1991; Mohammed, Cheng, Rouse & Marshall, 2001) und hat heute im internationalen Raum eine erhebliche Verbreitung erfahren (Thor, Lundberg, Ask, Olsson, Carli, Härenstam et al., 2007). Zentrale Annahme der SPC ist, dass Variationen und Schwankungen von Kennzahlen innerhalb laufender Prozesse und hervorgebrachter Ergebnisse in allgemeine (common cause) und spezielle Variationen (special cause variation) unterteilt werden können.

Allgemeine Variationen (common cause variation)

Allgemeine Variationen und Schwankungen entstehen im klinischen Alltag durch das Zusammenspiel vieler unüberschaubarer interdependenter Faktoren und komplexer Unterprozesse, die zusammen ein Ergebnis hervorbringen. Diese Variationen sind inhärente Bestandteile des Gesamtsystems, teilweise unbekannt, zufällig und durch die Komplexität natürlicherweise zu erwarten. Funktioniert das System stabil und zuverlässig, dann sind die Ergebnisse jedoch innerhalb bestimmter Grenzen vorhersagbar.

Ein Beispiel soll diese Annahme veranschaulichen: Laut Leitungsbeschluss sollen in einen Krankenhaus A alle Entlassungspatienten um 10:00 Uhr vormittags entlassen werden (Sollvorgabe). Wahrscheinlich werden jedoch nur die wenigsten Patienten exakt um 10:00 Uhr das Krankenhaus verlassen. Einige werden früher, später oder gegebenenfalls erst am Nachmittag verabschiedet. Selbst bei tagelangen konzentrierten Überlegungen wäre es unmöglich, alle Faktoren vollständig und erschöpfend zu beschreiben, die diese Abweichungen von 10:00 Uhr erklären: Arbeitsanfall am Entlassungstag, Erfahrung und Kompetenzen der verantwortlichen Pflegekräfte, Zahl der gleichzeitigen Entlassungen, Verfügbarkeit von Ärzten, Verfügbarkeit von Entlassungsmedikamenten, Vollständigkeit der Dokumentation, Gesprächs- und Beratungsbedarf der zu entlassenden Patienten, Notfälle und so weiter. All diese Faktoren, inklusive der unbekannten, ergeben im Zusammenspiel das «System». Die genauen Entlassungszeitpunkte sind somit niemals mit hundertprozentiger Genauigkeit vorhersagbar, sie variieren zufällig. Solange das System stabil arbeitet, ist dieser zufällige Prozess in den für dieses System typischen Grenzen vorhersagbar. Im Krankenhaus A können das ± 1 Stunde sein, das heißt, Patienten werden zwischen 9:00 und 11:00 Uhr entlassen. In einem Krankenhaus B könnten das ± 3 Stunden sein, das heißt, Patienten werden bereits ab 7:00 bis spätestens 13:00 Uhr entlassen. An dieser Stelle wird bereits deutlich, dass allein die Tatsache, Ergebnisse durch allgemeine stabile Variation hervorzubringen, nichts darüber aussagt, ob Prozesse an sich akzeptabel oder gut sind.

Spezielle Variationen (common cause variation)

Besondere oder spezielle Variationen entstehen durch Prozesse, die nicht mehr durch die natürliche Variation erklärt werden können. Sie sind nicht vorhersagbar und liegen außerhalb der natürlich zu erwartenden Grenzen. In der Regel gibt es dafür identifizierbare Auslöser und Gründe. Diese können erwünscht (zum Beispiel Implementierung eines umfassenden Dekubitusmanagementsystems führt zu einer Verringerung der Dekubitusinzidenz; verbessertes Selbstmanagement bei Diabetikern führt zu niedrigeren Blutzuckerwerten) oder unerwünscht sein (zum Beispiel zunehmende Vernachlässigung der Händehygiene führt zu vermehrten Kathetherinfektionen). Eine spezielle Variation im verwendeten Ausgangsbeispiel könnte sein, dass alle Patienten regulär um 10:00 Uhr ± 1 Stunde entlassen werden. Aufgrund eines lang anhaltenden hohen Krankenstandes verschiebt und erweitert sich das Zeitfenster auf 10:00 Uhr bis 13:00 Uhr. In Tabelle 1 sind weitere fiktive Beispiele für zufällige und nichtzufällige Variationen, wie sie in der Pflegepraxis auftreten können, veranschaulicht.

Konsequenzen für die Praxis

Im Sinne der SPC gilt es, beide Variationsarten innerhalb von Prozessen zu identifizieren. Es wird der Frage nachgegangen, ob sich die beobachteten Prozesse und Ergebnisse, ausgedrückt als Kennzahlen, innerhalb «statistischer Kontrolle» oder außerhalb befinden. Eine zentrale Annahme der Theorie der SPC ist, eine gewisse Unvorhersagbarkeit von Ergebnissen in komplexen Prozessen grundsätzlich zu unterstellen. Es gibt keinen Bereich des Lebens, in dem Prozesse stets gleich ablaufen oder Ergebnisse vollkommen identisch sind. Basierend auf der Arbeit von Shewhart (1931) wird zur Veranschaulichung dieser These immer wieder gern das Beispiel der eigenen Unterschrift herangezogen. Die eigene Unterschrift ist stets sehr ähnlich, aber aus unzähligen Gründen niemals identisch. Gleiches gilt für pflegerische Pro-

https://econtent.hogrefe.com/doi/pdf/10.1024/1012-5302/a000202 - Saturday, May 04, 2024 1:38:30 AM - IP Address:18.188.70.255

Tabelle 1: Beispiele für systembedingte (zufällige) und spezielle (nichtzufällige) Schwankungen im Pflegealltag.

Beispiele	allgemeine Variationen (common cause variation)	spezielle Variationen (special cause variation)
24h-Flüssigkeitbilanz	Die innerhalb eines Tages errechneten Bilanzen schwanken innerhalb bestimmter Grenzen, da eine bestimmte Genauigkeit (z. B. Ablesung der Urinausscheidung) nicht erreicht wird oder Einfuhrmengen auf- und abgerundet werden.	Einige Pflegekräfte rechnen Kurzinfusionen als Einfuhr, andere nicht.
Vitalwerte	Vitalwerte (Herzfrequenz, Blutdruck) einer Person weisen im Tagesverlauf Schwankungen auf.	Vitalwerte werden vor und nach einer blutdrucksenkenden Therapie gemessen.
Häufigkeiten schwerer Sturzverletzungen	In einem Pflegeheim weisen einige wenige Bewohner nach einem Sturz schwere Verletzungen auf, andere nicht.	Nach Umgebungsanpassungen und Umstellungen der Medikation (Reduktion der Polypharmazie) kommt es zu weniger schweren Verletzungen.
Krankenstand auf Station	In einem Team sind immer wieder einige Kollegen krank.	Aufgrund einer Verdichtung der Arbeit und zunehmend schlechter Führung steigt der Krankenstand an.

zesse und Ergebnisse. Auch diese sind niemals deckungsgleich, und es gibt immer Abweichungen und Variationen. Daraus folgt, dass der Ausgang und die Ergebnisse komplexer Pflegehandlungen niemals exakt vorherzusagen sind. Shewhart bezeichnete diesen Umstand treffenderweise als «inability to do exactly what we want to do» (Shewhart, 1931: 5). Die Frage, die die SPC zu beantworten sucht, lautet nun: Wie viel Abweichung ist zu viel?

Die Unterscheidung in natürlich bedingte (common cause) und unnatürlich bedingte (special cause) Variationen ist von hoher Wichtigkeit, weil davon adäquate Managemententscheidungen abhängen. Greift man beispielsweise in vermeintlich nichtstabile Prozesse ein, die in Wirklichkeit natürliche Variationen aufweisen, dann können diese außer Kontrolle gebracht werden, vermehrt Variationen aufweisen und unkontrollierbar werden (Überreaktion). Es kommt zu einer Verschwendung von Ressourcen (Deming, 1986; Lloyd, 2004). Bezogen auf das Ausgangsbeispiel wäre es beispielsweise ein Fehler, Pflegekräfte im Krankenhaus A (systembedingte Entlassungszeit 10:00 Uhr ± 1 Stunde) anzuweisen, keinen Patienten später als 10:30 Uhr zu entlassen. Dieses Vorhaben wäre zum Scheitern verurteilt, da die Gestaltungsspielräume der Pflegekräfte bezogen auf das Gesamtsystem begrenzt sind. Die Folgen wären Schuldzuweisungen, Frust und Ressourcenvergeudung. Soll das Zeitfenster der Entlassungen tatsächlich auf maximal 10:30 Uhr begrenzt werden, dann müsste das Krankenhausmanagement das gesamte System von Grund auf verändern, um die bislang unkontrollierte Variation im Entlassungsprozess auf ± 30 Minuten zu verringern.

Eine weitere Annahme der SPC besagt, dass Prozesse, die chaotisch ablaufen und «außer Kontrolle» sind, nicht zielgerichtet verändert werden können. Prozesse sind nicht planbar und steuerbar, solange sie nicht innerhalb bestimmter Grenzen vorhersagbar sind (Deming, 1986; Lloyd, 2004).

Kennzahlen im zeitlichen Verlauf

Qualitätsverbesserungen erfolgen in der Regel über zwei Wege: Standardisierung und Optimierung (Baartmans & Geng, 2006). Standardisierung vereinheitlicht Prozesse in der Annahme, dass diese dann besser funktionieren, planbar und vorhersagbar sind. Standardisierung bedeutet Reduktion von Variabilität und Schwankungen. Bei der Optimierung geht es darum, Prozesse und Ergebnisse zu verändern und auf ein neues Niveau zu bringen. Beide Herangehensweisen können kombiniert werden. Nach Durchführung entsprechender Qualitätsverbesserungen stellt sich die Frage, waren die Maßnahmen erfolgreich? Diese Frage kann mithilfe von SPC-Werkzeugen beantwortet werden. Eine Möglichkeit bieten sogenannte Laufdiagramme.

Laufdiagramme

Laufdiagramme (run charts) sind sehr einfach anzuwendende Werkzeuge, um Prozesse und Ergebnisse im zeitlichen Verlauf darzustellen. Erhobene Kennzahlen – ganz gleich welcher Art (Raten, Proportionen, Mengen, Ereignisse) – werden zeitlich nacheinander geordnet und gegen die jeweils gemessene Zeiteinheit (Tage, Wochen, Monate) in ein Diagramm eingetragen. Nach Eintragung aller verfügbaren Datenpunkte (mindestens 10) wird der Median als waagerechte Linie eingetragen. Basierend auf einem fiktiven Beispieldatensatz (Anzahl von Beschwerden pro Woche in einem Klinikum) ist in Abbildung 1 ein entsprechendes

https://econtent.hogrefe.com/doi/pdf/10.1024/1012-5302/a000202 - Saturday, May 04, 2024 1:38:30 AM - IP Address:18.188.70.255

Laufdiagramm dargestellt. Neben dem Median sind die Anzahl, Lage und Abfolgen von «Läufen» (runs) relevant. Ein Lauf sind zusammenhängende Datenpunkte auf einer Seite des Medians. Ein Lauf besteht aus mindestens einem Datenpunkt. Datenpunkte, die direkt auf dem Median liegen, zählen nicht. Somit sind in Abbildung 1 zehn Läufe auszumachen.

Ist das Diagramm erstellt und sind die Läufe und der Median identifiziert, werden im nächsten Schritt nach Anzeichen für nichtzufällige (special cause) Variationen gesucht. Dafür eignen sich etablierten Regeln, welche auf der Wahrscheinlichkeitstheorie beruhen. Die im Folgenden dargestellten Regeln legen einen p-Wert < 0,05 zugrunde, das heißt, von nichtzufälliger Variation wird dann ausgegangen, wenn die Wahrscheinlichkeit für ein zufälliges Zustandekommen der Anordnung der beobachteten Datenpunkte innerhalb aller betrachteten Datenpunkte kleiner als 5 Prozent ist.

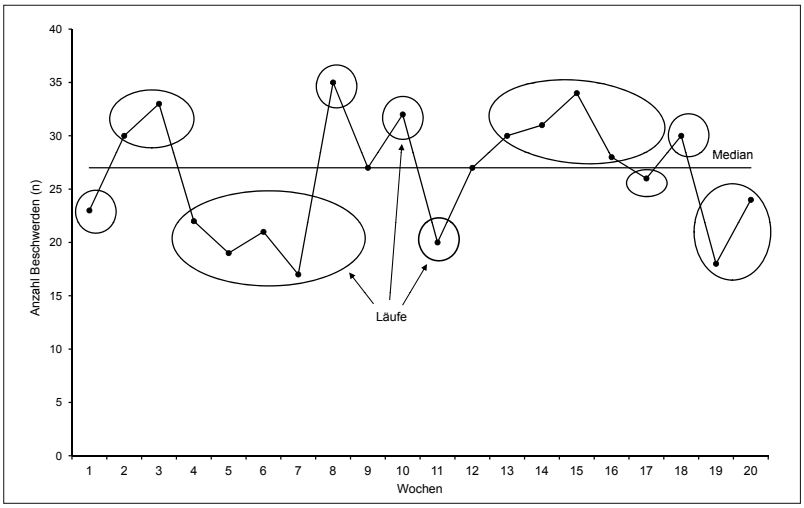

Abbildung 1: Anzahl der Beschwerden pro Woche (fiktives Beispiel) im Laufdiagramm (common cause variation).

Tabelle 2: Minimal und maximal zu erwartende Anzahl von Läufen bei zufällig bedingter Variabilität (modifiziert nach Lloyd, 2004; Perla, Provost & Murray, 2011).

Anzahl Datenpunkte, die nicht auf dem Median liegen	zu erwartendes Minimum an Läufen	zu erwartendes Maximum an Läufen
15	4	12
20	6	15
25	9	17
30	11	20
35	13	23
40	15	26
45	17	30
50	19	33

Zu viele oder zu wenige Läufe

Gibt es zu wenige oder zu viele Läufe im Datensatz, dann ist von nichtzufälligen Veränderungen über die Zeit auszugehen (National Health Service Scotland, 2007; Lloyd, 2004). Die Anzahl zufällig zu erwartender Läufe ist abhängig von der Gesamtzahl aller Datenpunkte, die Läufe darstellen, das heißt, Werte auf dem Median zählen nicht. Wann die Grenzen überschritten sind, lässt sich in Tabelle 2 ablesen. Bezogen auf das Ausgangsbeispiel (Abbildung 1) ist festzustellen, dass zehn Läufe von 18 infrage kommenden Datenpunkten innerhalb natürlich zu erwartender Variabilität liegen. Hinweise auf nichtzufällige Variabilität sind in Abbildung 2 beispielhaft dargestellt.

Zu lange Läufe

In der Regel kann davon ausgegangen werden, dass 6 bis 8 zusammenhängende Datenpunkte auf einer Seite des Medians eine nichtzufallsbedingte Verschiebung (shift) darstellen (Deming, 1986; Lloyd, 2004; National Health Service Scotland, 2007; Perla et al., 2011). Aus statistischer Sicht hängt jedoch auch hier die Länge der zufällig zu erwartenden Läufe von der Gesamtanzahl vorhandener Datenpunkte ab (Tabelle 3). Die maximale Lauflänge im Beispiel (Abbildung 1) beträgt vier. Somit weist auch dieses Kriterium auf einen kontrollierten Prozess hin. Eine statistisch signifikante Verschiebung in der Anzahl der Beschwerden ist in Abbildung 3 dargestellt.

Trend

Ein nichtzufälliger Trend besteht aus einer bestimmten Anzahl konsekutiv an- oder absteigender Datenpunkte, wobei die Punkte auf dem Median in diesem Fall mitgezählt werden. Benachbarte, identische Werte werden als ein Datenpunkt im Trend gezählt. Die Länge eines nichtzufälligen Trends hängt von der Gesamtanzahl der betrachteten Werte ab (Tabelle 4). Im

https://econtent.hogrefe.com/doi/pdf/10.1024/1012-5302/a000202 - Saturday, May 04, 2024 1:38:30 AM - IP Address:18.188.70.255

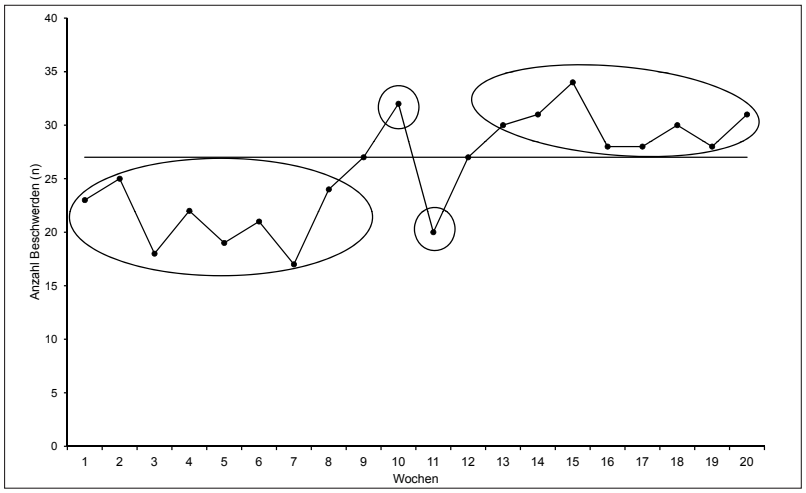

Abbildung 2: Anzahl der Beschwerden pro Woche: vier Läufe bei 18 Datenpunkten (special cause variation).

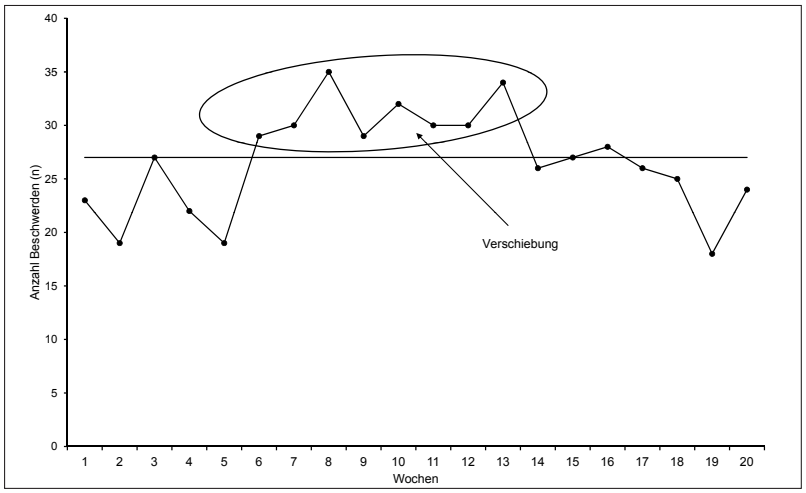

Abbildung 3: Anzahl der Beschwerden pro Woche: Verschiebung zwischen Woche 6 und 13 (special cause variation).

Tabelle 3: Maximal zu erwartende Länge von Läufen bei zufällig bedingter Variabilität (modifiziert nach Mosteller, 1941).

Anzahl aller Datenpunkte	Maximal zu erwartende Lauflänge
10	5
20	7
30	8
40	9
50	10

Tabelle 4: Minimale Anzahl aufeinanderfolgender auf- oder absteigender Datenpunkte, die einen nichtzufälligen Trend darstellen (modifiziert nach Lloyd, 2004; Olmstead, 1946).

Anzahl aller Datenpunkte	minimale Anzahl auf- oder absteigender Werte
8	5
9 bis 20	6
21 und mehr	7

Ausgangsbeispiel sind maximal fünf aufeinanderfolgende ansteigende Werte von insgesamt 20 Datenpunkten zu erkennen. Somit liegen auch in diesem Fall keine Hinweise für «echte» Trends vor. Hinweise auf einen Trend finden sich in Abbildung 4.

Anwendung von Laufdiagrammen

Kontinuierlich gesammelte Daten im zeitlichen Verlauf gehören zu den Standardwerkzeugen der Qualitätsverbesserungspraxis (Pronovost, Nolan, Zeger, Miller & Rubin, 2004). Prozesse und Ergebnisse inklusive der natürlich zu erwartenden Variabilität können grafisch sichtbar gemacht werden. Jeder verfügbare Datenpunkt wird verwendet und erhöht damit den Informationsgehalt sowie die Aussagekraft der Darstellung. Damit stehen Laufdiagramme im Gegensatz zu den weit verbreiteten Mittelwertvergleichen und einfachen Vorher-Nachher-Messungen (zwei Datenpunkte), die für die Evaluation von Qualitätsinitiativen eher ungeeignet sind (Baumgarten, Shardell & Rich, 2009; Kottner & Halfens, 2010; Perla et al., 2011). In Anlehnung an Deming (1986) soll ein abschließendes Beispiel diese Aussage verdeutlichen: Angenommen, zwei Qualitätsverantwortliche zweier vergleichbarer Krankenhäuser (A und B) möchten wissen, ob es Unterschiede in der Beschwerdehäufigkeit innerhalb der letzten 20 Wochen zwischen beiden Einrichtungen gab. Dafür tauschen Sie Mittel- und Streuwerte untereinander aus und stellen fest, dass diese identisch sind: Median = 27, Spannweite = 17 bis 35, Interquartilsabstand = 21,5 bis 30,5. Basierend auf diesen Parametern schlussfolgern sie, dass es zwischen beiden Krankenhäusern keine Unterschiede gab. Nun gelten diese Parameter für alle vier bisher dargestellten Beispiele (Abbildungen 1 bis 4) und auch für einen hypothetischen Ext-

https://econtent.hogrefe.com/doi/pdf/10.1024/1012-5302/a000202 - Saturday, May 04, 2024 1:38:30 AM - IP Address:18.188.70.255

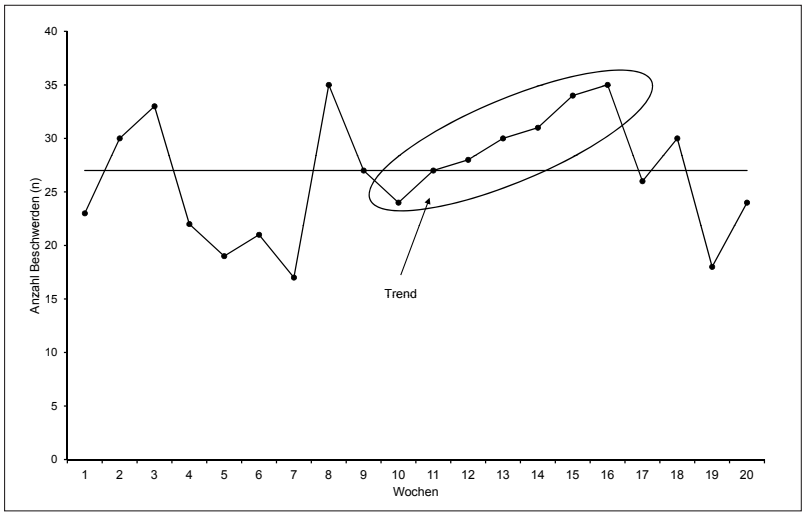

Abbildung 4: Anzahl der Beschwerden pro Woche: sieben kontinuierlich ansteigende Datenpunkte bilden einen Trend (special cause variation).

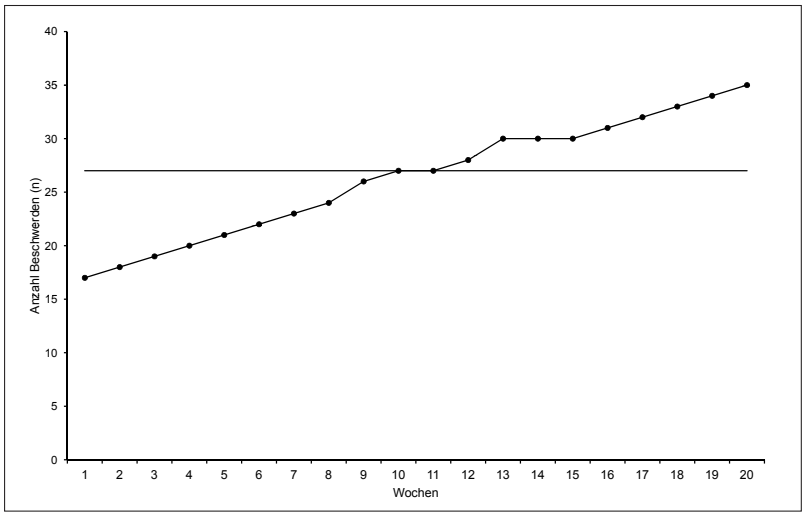

Abbildung 5: Anzahl der Beschwerden pro Woche (special cause variation).

remfall in Abbildung 5. Angenommen der Verlauf der Beschwerden pro Woche entspräche in dem einen Haus A der Abbildung 1, in dem anderen Haus B Abbildung 5, dann wären die Urteile beider Qualitätsverantwortlichen wahrscheinlich anders ausgefallen. Während im Haus A das Beschwerdeniveau vorhersagbar und stabil zu sein scheint, greift im Haus B ein Mechanismus (special cause), der die Anzahl der Beschwerden extrem ansteigen ließ. Der Prozess ist «außer Kontrolle» und das Management sollte dringend handeln.

Auf institutioneller Ebene gibt es mehrere Einsatzmöglichkeiten. Zunächst können zurückliegende Prozesse und Ergebnisse dargestellt und beurteilt werden. Laufdiagramme können dynamisch eingesetzt werden, indem aktuelle Datenpunkte zu den bereits vorhandenen hinzugefügt werden. Auf diese Art und Weise werden signifikante Veränderungen im Prozess unmittelbar sichtbar und es kann umgehend reagiert werden (Überwachung). Schließlich können mittels Laufdiagrammen Prozesse und Ergebnisse zunächst charakterisiert werden (Baseline). Basierend auf diesen Ausgangswerten können Veränderungswünsche oder Ziele formuliert werden und entsprechende Interventionen durchgeführt werden (zum Beispiel Schulungen, Implementierung von Standards). Während der Implementierungsphase werden weiterhin kontinuierlich Daten erhoben, und es kann überprüft werden, ob die intendierten Veränderungen tatsächlich zu messbaren Veränderungen geführt haben. Laufdiagramme stellen somit typische Bausteine im PDCA-Zyklus dar (Boaden, Harvey, Moxham & Proudlove, 2008). In diesem Sinn wurden sie beispielsweise erfolgreich zur Überwachung chirurgischer Wundinfektionen (Philips, 2009), zur Evaluation eines Händehygieneprogramms (Waters, 2009) oder zur Evaluation erfolgreicher Darmvorbereitungen vor der Koloskopie eingesetzt (Crump, 2010).

Laufdiagramme können für einzelne Personen genutzt werden. Die Leistungen einzelner Leistungserbringer oder -empfänger können im zeitlichen Verlauf beurteilt werden (zum Beispiel Anteil von Komplikationen nach immer gleichen Interventionen, individuelle Fortschritte in der Rehabilitation). Es liegen erste positive Berichte vor, dass durch Laufdiagramme, die von Patienten selbst geführt wurden, Bluthochdruckthera pien begleitet und auf ihre Wirksamkeit hin evaluiert wurden (Herbert & Neuhauser, 2004). Weitere Einsatzgebiete in der Pflege sind denkbar, denn gerade Pflegeinterventionen (zum Beispiel Beratungen, Selbstmanagementtrainings) zeigen, sofern sie effektiv waren, eher längerfristige Ergebnisse als unmittelbar schnell feststellbare Veränderungen.

https://econtent.hogrefe.com/doi/pdf/10.1024/1012-5302/a000202 - Saturday, May 04, 2024 1:38:30 AM - IP Address:18.188.70.255

Vor- und Nachteile von Laufdiagrammen

Der offensichtlich größte Vorteil von Laufdiagrammen ist deren Einfachheit. Sie können mithilfe verbreiteter Computerprogramme oder auch handschriftlich angefertigt werden. Die Diagrammfunktion in Microsoft Excel® bietet sich sehr gut an. Schritt-für-Schritt-Anleitungen finden sich beispielsweise im «Tutorial Guide Statistical Process Control» des National Health Service Scottland (2007). Handschriftlich erstellte Laufdiagramme haben den Nachteil, dass sich bei jedem neu hinzukommen Datenpunkt der Median verändert und somit das Diagramm neu gezeichnet werden müsste. Microsoft Excel® oder vergleichbare Programmen passen den Median automatisch an. Unter Anwendung der vorgestellten Regeln können Laufdiagramme ohne statistische Formeln oder Berechnungen ausgewertet werden. Somit versetzen Laufdiagramme eine große Anwendergruppe in die Lage, selbstvergleichende Qualitätsmessungen unkompliziert durchzuführen.

Periodische Messungen lassen sich leicht in die tägliche Routine integrieren und stehen dem Management unmittelbar zur Verfügung. Damit sind sie deutlich weniger aufwändig, und entscheidungsrelevante Informationen sind schneller verfügbar als die Ergebnisse umfassender Audits und Studien. Je nach Anwendungszweck kann jedoch gerade dieser Vorteil einen Nachteil darstellen, denn Routinemessungen sind häufig fehleranfälliger als methodisch hochwertig durchgeführte Studien.

Für die sinnvolle Interpretation von Laufdiagrammen sind mindestens zehn Datenpunkte notwendig (Perla et al., 2011), 15 Datenpunkte und mehr wären noch besser (National Health Service Scotland (2007). Je mehr Datenpunkte verfügbar sind, umso zuverlässiger können Prozesse und Ergebnisse interpretiert werden. Wie die Anwendung jedes statistischen Tests, können auch die hier vorgestellten Regeln zum Herausfinden nichtzufälliger Variationen falsch-positiv oder falsch-negativ sein. Bei sehr schiefen Verteilungen oder ausgeprägten Boden- und Deckeneffekten können einfache Laufdiagramme nicht erstellt werden, weil der Median in diesen Fällen kein sinnvoller Mittelwert ist.

Datenanalysen mittels Laufdiagrammen können zu anderen Ergebnissen führen als andere Auswertungsmethoden und statistischer Zugänge (zum Beispiel klassische Mittelwertvergleiche) hervorbringen würden. Dabei sollte sich nicht die Frage gestellt werden, welche Methode die «richtigeren» Ergebnisse liefert, sondern welche Fragen beantwortet werden sollen. Dieses Thema wird im zweiten Teil dieser Serie vertieft.

Grundsätzlich sind Laufdiagramme wertfreie Werkzeuge. Man kann mit ihnen Prozesse und Ergebnisse statistisch analysieren und kontrollieren, doch die Ergebnisse erlauben keine inhaltliche Bewertung, das heißt, sie sagen nichts darüber aus, ob die Prozesse und Ergebnisse als gut oder schlecht zu beurteilen sind. Inwiefern Laufdiagramme tatsächlich zur Entwicklung pflegerischer Qualität beitragen, ist noch unbekannt. Bislang scheint dieser Ansatz keinen Eingang in das deutschsprachige pflegerische Qualitätsmanagement gefunden zu haben.

Ausblick

Menschen möchten und sollen pflegerische und gesundheitsbezogene Leistungen erhalten, die dem aktuellen Stand der Kunst und des Wissens entsprechen, die für die Personen zu positiven Ergebnissen führen und keinen Schaden anrichten. Inwiefern Prozesse und Ergebnisse tatsächlich diesen Vorstellungen entsprechen, ist meistens nicht einfach zu sagen. Dennoch sind Antworten auf diese Frage wichtig. Adäquate Gütebeurteilungen von pflegerischen oder medizinischen Leistungen sind die Voraussetzung für Veränderungen, um festgestellte Qualitätsdefizite abzubauen oder Leistungen zu verbessern.

Die Feststellung von auf Kennzahlen basierten Qualitätsunterschieden ist eine Ausschlussdiagnose. Erst wenn systematische und zufällige Messfehler adäquat berücksichtigt wurden, sind vergleichende Urteile zulässig. Laufdiagramme bieten einfache Möglichkeiten, Kennzahlen im zeitlichen Verlauf darzustellen, zu analysieren und für intendierte Veränderungen zu nutzen. Eine Weiterentwicklung von Laufdiagrammen stellen sogenannte Regelkarten (control charts) dar, auf die im Teil 2 dieser Reihe eingegangen wird.

Danksagung

Vielen Dank an Herrn Armin Hauss, Abteilung Qualitätsmanagement der Charité-Universitätsmedizin Berlin, für die kritische Durchsicht des Manuskripts aus Qualitätsmanagement-Perspektive.

Comparative quality measurements part 1: run charts

Quality assessment may be based on data of quality indicators. There are two main approaches for comparative quality measurements: comparison of data of the same service at different points over time or comparison of data of different services at the same time. Risk adjustment and standardisation must be performed and random variation must be adequately

https://econtent.hogrefe.com/doi/pdf/10.1024/1012-5302/a000202 - Saturday, May 04, 2024 1:38:30 AM - IP Address:18.188.70.255

taken into account. In Statistical Process Control (SPC) theory common cause and special cause variation are distinguished. Processes in statistical control are stable and predictable. If processes exhibit special cause variation the management should investigate the reasons for this and manage the causes. Run charts as simple tools to display and analyse data of processes and outcomes over time are discussed in this article. They can be used for self-comparison. The following two parts of this three-part series explain control charts and funnel plots.

Key words: nursing, quality control, quality improvement, quality indicators, run charts

Literatur

AQUA-Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH (2010). Allgemeine Methoden. Göttingen.

Ärztliches Zentrum für Qualität in der Medizin (2009). Programm für nationale VersorgungsLeitlinien von BÄK, KBV und AWMF: Qualitätsindikatoren. Neukirchen: Make a Book.

Attree, M. (2001). Patients' and relatives' experiences and perspectives of «Good» and «Not so Good» quality care. Journal of Advanced Nursing, 33 (4), 456 – 466.

Baumgarten, M.; Shardell, M.; Rich, S. (2009). Methodological issues in studies of the effectiveness of pressure ulcer prevention interventions. Advances in Skin and Wound Care, 22 (4), 180 – 188.

Baartmans, P. C. M.; Geng, V. (2006). Qualität nach Maß. Bern: Hans Huber.

Berlowitz, D. R.; Anderson, J. J.; Ash, A. S.; Brandeis, G. H.; Brand, H. K.; Moskowitz, M. A. (1998). Reducing random variation in reported rates of pressure ulcer development. Medical Care, 36 (6), 818 – 825.

Berwick, D. M. (1991). Controlling variation in health care: a consultation from Walter Shewhart. Medical Care, 29 (12), 1212 – 1225.

Boaden, R.; Harvey, G.; Moxham, C.; Proudlove, N. (2008). Quality Improvement: Theory and Practice in Healthcare. Coventry: NHS Institute for Innovation and Improvement.

Bundesministerium für Gesundheit; Bundesministerium für Familie, Senioren, Frauen und Jugend (2011). Entwicklung und Erprobung von Instrumenten zur Beurteilung der Ergebnisqualität in der stationären Altenhilfe. Berlin: Druckerei im Bundesministerium für Arbeit und Soziales.

Burla, L.; Schaffert, R.; Mylaeus, M.; Rüesch, P. (2010). Entwicklung und Erprobung von Qualitätsindikatoren für die ambulante Pflege in der Schweiz. Gesundheitswesen, 72 (2), 106 – 113.

Campbell, S. M.; Roland, M. O.; Buetow, S. A. (2000). Defining quality of care. Social Science & Medicine, 51 (11), 1611 – 1625.

Committee on Quality Health Care in America; Institute of Medicine (2001). Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, D. C.: National Academy Press.

Crump, B. (2010). Interactive improvement clinic: increase the success of your projects. International Forum on Quality & Safety in Healthcare, Nice, France. http://internationalforum.bmj.com/2010-forum/presentation-slides/friday/G9%20Bernard%20Crump.pdf [26.10.2011].

Deming, W. E. (1986). Out of the Crisis. Cambridge, Massachusetts: Cambridge University Press.

Herbert, C.; Neuhauser, D. (2004). Improving hypertension care with patient-generated run charts: physician, patient, and management perspectives. Quality Management in Healthcare, 13 (3), 174 – 177.

Jones, R. N.; Hirdes, J. P.; Poss, J. W.; Kelly, M.; Berg, K.; Fries, B. E.; Morris, J. N. (2010). Adjustment of nursing home quality indicators. BMC Health Service Research, 15 (10), 96.

Karon, S. L.; Sainfort, F.; Zimmerman, D. R. (1999). Stability of nursing home quality indicators over time. Medical Care, 37 (6), 570 – 579.

Kottner, J.; Dassen, T.; Heinze, C. (2011). Funnel-Plots zum Vergleich von Dekubitus- und Sturzkennzahlen in 76 Pflegeheimen. Gesundheitswesen, 73 (6), e98–e102.

Kottner, J.; Halfens, R. (2010). Using statistical process control for monitoring the prevalence of hospital-acquired pressure ulcers. Ostomy Wound Management, 56 (5), 54 – 59.

Lloyd, R. (2004). Quality Health Care: A Guide to Developing and Using Indicators. Sudbury, Massachusetts: Jones and Bartlett Publishers.

Mant, J. (2001). Process versus outcome indicators in the assessment of quality health care. International Journal for Quality in Health Care, 13 (6), 475 – 480.

Mohammed, M. A.; Cheng, K. K.; Rouse, A.; Marshall, T. (2001). Bristol, Shipman, and clinical governance: Shewhart's forgotten lessons. The Lancet, 357 (9254), 463 – 467.

Mosteller, F. (1941). Note on an application of runs to quality control charts. Annals of Mathematical Statistics, 12, 228 – 232.

Nakrem, S.; Vinsens, A. G.; Harkless, G. E.; Paulsen, B.; Seim, A. (2009). Nursing sensitive quality indica-

https://econtent.hogrefe.com/doi/pdf/10.1024/1012-5302/a000202 - Saturday, May 04, 2024 1:38:30 AM - IP Address:18.188.70.255

tors for nursing home care: International review of literature, policy and practice. International Journal of Nursing Studies, 46 (6), 848 – 857.

National Health Service Scotland (2007). Tutorial Guide Statistical Process Control. http://www.indicators.scot.nhs.uk/SPC/Statistical_Process_Control_Tutorial_Guide_V5.pdf [20.07.2011].

Olmstead, P. S. (1946). Distribution of sample arrangements for runs up and down. Annals of Mathematical Statistics, 17, 24 – 33.

Ovretveit, J. (2001). Quality evaluation and indicator comparison in health care. International Journal of Health Planning and Management, 16 (3), 229 – 241.

Perla, R. J.; Provost, L. P.; Murray, S. K. (2011). The run chart: a simple tool for learning from variation in healthcare processes. BMJ Quality and Safety, 20 (1), 46 – 51.

Philips, J. (2009). Introduction of Safety Briefings and ‹surgical site infection bundle› into operating theatres. International Forum on Quality & Safety in Healthcare, Berlin, Germany. http://internationalforum.bmj.com/forum-resources/F5%20phillips.pdf [26.10.2011].

Pronovost, P. J.; Nolan, T.; Zeger, S.; Miller, M.; Rubin, H. (2004). How can clinicians measure safety and quality in acute care? Lancet, 363 (9414), 1061 – 1067.

Rantz, M. J.; Hicks, L.; Petroski, G. F.; Madsen, R. W.; Mehr, D. R.; Conn, V.; Zwygart-Staffacher, M.; Maas, M. (2004). Stability and sensitivity of nursing home quality indicators. The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 59 (1), 79 – 82.

Reiter, A.; Fischer, B.; Kötting, J.; Geraedts, M.; Jäckel, W. H.; Döbler, K. (2008). QUALIFY: Ein Instrument zur Bewertung von Qualitätsindikatoren. Zeitschrift für ärztliche Fortbildung und Qualitätsmanagement im Gesundheitswesen, 101 (10), 683 – 688.

Rubin, H. R.; Pronovost, P.; Diette, G. B. (2001). From a process of care to a measure: the development and testing of a quality indicator. International Journal for Quality in Health Care, 13 (6), 489 – 496.

Schmitz, P. (2004). Was-Ist-SPC-De. Rheinische Fachhochschule Köln e. V. http://www.was-ist-spc.de/index.php?section=historie [27.07.2011].

Shewhart, W. A. (1931). Economic Control of Quality of Manufactured Product. New York: Van Nostrand.

Shewhart, W. A. (1939). Statistical Method from the Viewpoint of Quality Control. Washington, D.C.: Graduate School of the Department of Agriculture.

Szklo, M.; Nieto, J. (2007). Epidemiology Beyond the Basics. Sudbury, Massachusetts: Jones and Bartlett Publishers.

Thor, J.; Lundberg, J.; Ask, J.; Olsson, J.; Carli, C.; Härenstam, K. P.; Brommels, M. (2007). Application of statistical process control in healthcare improvement: systematic review. Quality and Safety in Health Care, 16 (5), 387 – 399.

Waters, L. (2009). Why don't we wash our hands? We know it works. International Forum on Quality & Safety in Healthcare, Berlin, Germany. http://internationalforum.bmj.com/forum-resources/presentation-files/I6%20waters.pdf [26.10.2011].

West, D. (2010). First outcome indicators allow national comparison. Nursing Times, 106 (13), 1.

Wollersheim. H.; Hermens, R.; Hulscher, M.; Braspenning, J.; Ouwens, M.; Schouten, J.; Marres, H.; Dijkstra, R.; Grol, R. (2007). Clinical indicators: development and applications. The Netherlands Journal of Medicine, 65 (1), 15 – 22.

Korrespondenzadresse

PD Dr. Jan Kottner
Clinical Research Center for Hair and Skin Science
Klinik für Dermatologie, Venerologie und Allergologie
Charité-Universitätsmedizin Berlin
Campus Charité Mitte
Charitéplatz 1
DE-10117 Berlin
Tel.: +49-030-450 518 122
Fax: +49-030-450 518 952

jan.kottner@email.de