



Elektronischer Sonderdruck für J. Kieschke

Ein Service von Springer Medizin

Bundesgesundheitsbl 2014 · 57:33–40 · DOI 10.1007/s00103-013-1872-5

© Springer-Verlag Berlin Heidelberg 2013

J. Kieschke · M. Hoopmann

Aktives Monitoring kleinräumiger Krebshäufungen

Vorstellung eines neuen Monitoringkonzeptes für das Epidemiologische
Krebsregister Niedersachsen

Diese PDF-Datei darf ausschließlich für nichtkommerzielle Zwecke verwendet werden und ist nicht für die Einstellung in Repositorien vorgesehen – hierzu zählen auch soziale und wissenschaftliche Netzwerke und Austauschplattformen.

Aktives Monitoring kleinräumiger Krebshäufungen

Vorstellung eines neuen Monitoringkonzeptes für das Epidemiologische Krebsregister Niedersachsen

Das Epidemiologische Krebsregister Niedersachsen (EKN) hat die gesetzliche Aufgabe, das Krebsgeschehen in Niedersachsen zu beobachten, einschließlich etwaiger regionaler oder zeitlicher Häufungen bei spezifischen Krebsdiagnosen. Unterhalb der Landkreisebene erfolgen derartige Untersuchungen bisher nur „anlassbezogen“ aufgrund spezifischer Anfragen. Nach einem Bericht des EKN zu Leukämieerkrankungen in der Samtgemeinde Assen erhielt das EKN Ende 2010 vom Niedersächsischen Sozialministerium den Auftrag, ein Monitoringkonzept zu entwickeln, das ganz Niedersachsen kontinuierlich nach in der Krebsinzidenz auffälligen Gemeinden prüfen soll.

Regionale Häufungen registrierter Krebserkrankungen können verschiedene Ursachen haben:

- Es liegt eine echte Erhöhung vor:
 - systematisch bedingt, z. B. durch Unterschiede im Meldeverhalten, von Screeningmaßnahmen oder bei der Kodierung von Diagnosen;
 - bedingt durch außergewöhnliche Verteilung individueller Risikofaktoren, z. B. von „Lifestyle“-Faktoren oder von genetischen Risiken;
 - bedingt durch umwelt- oder arbeitsplatzassoziierte Faktoren.
- Die Erhöhung ist als zufällig einzu-stufen.

Anders als bei anlassbezogenen Krebsclusteranfragen, bei denen in der Regel eine mögliche Umweltextposition vermutet wird, sind für die im Rahmen eines Monitorings identifizierte Gemeinden zunächst keinerlei verdächtige Expositionen bekannt. Die Möglichkeit rein zufälliger Erhöhungen (sog. Fehlalarme) führt zur Frage, welche Konsequenzen diese haben und wie viele Fehlalarme durch das Monitoring zu erwarten sind.

Im ersten Teil des hier vorliegenden Beitrages werden Problemfelder der statistisch-epidemiologischen Auswertung und der Interpretation regionaler Häufungen beschrieben, bevor im zweiten Teil auf das derzeitige Konzept des niedersächsischen Monitorings eingegangen wird. Die Diskussion um methodische Aspekte des Monitorings, aber auch zur Relevanz der gesellschaftlichen Folgen etwa von Fehlalarmen sowie ein kurzes Fazit beschließen den Artikel.

Hintergrund

Clusteranalysen

Eine Clusterstudie basiert auf einer nicht geplanten Beobachtung einer scheinbar erhöhten Zahl an Fällen einer bestimmten Erkrankungsart in einem kleinräumigen Gebiet zu einer bestimmten Zeit. Das Auftreten von Clustern ist bei seltenen Krankheiten ein erwartetes Phänomen: Aufgrund zufälliger Variatio-

nen und der kleinen Anzahl ist unter der Vielzahl der möglichen regionalen Bezüge ein zufälliges Zusammentreffen einiger Fälle in *irgendeiner* Region sicher, selbst wenn sie für jede einzelne *vorab* („a priori“) definierte kleinräumige Region höchst unwahrscheinlich wäre. Es gibt keine statistische oder andere Methode, um zu entscheiden, ob ein Cluster zufällig bedingt ist oder nicht, wenn eine Clusteruntersuchung im Nachhinein aufgrund einer scheinbaren Erhöhung von Krankheitsfällen durchgeführt wird. Untersuchungen zu Clustern sind im wissenschaftlichen Sinne häufig nicht zielführend und werden eher als öffentliche Reaktion auf eine gesellschaftliche Besorgnis eingesetzt. Im Einzelfall können jedoch neue Forschungsthemen zu einem bislang unbekanntem Zusammenhang abgeleitet werden (vgl. [1, 2, 3, 4, 5]).

Bei einer Clusteranalyse wird zunächst untersucht, ob es sich tatsächlich um eine „statistisch signifikante“ Häufung von Fällen handelt. Die dabei vorgegebene Irrtumswahrscheinlichkeit sagt aber auch aus, dass eine auffällige Erhöhung dennoch mit dieser Wahrscheinlichkeit zufällig vorkommen kann.

In seinem grundlegenden Aufsatz zum Nutzen von Clusteruntersuchungen führte K. Rothman 1990 aus [2], dass sie epidemiologisch kaum von Nutzen und selbst als Instrument der Risikokommunikation meist nicht zielführend seien.

Seine „epidemiologischen Hauptkritikpunkte“ lauten:

- Cluster sind in der Regel zu klein, um damit kontrollierte epidemiologische Untersuchungen durchführen zu können.
- Den Fällen liegen häufig vage Fallabgrenzungen zugrunde.
- Die Abgrenzung der zugrunde gelegten Bezugspopulation erfolgte nach der Fallerfassung („a posteriori“), häufig zur Illustration des Clusters.
- Die Expositionsabgrenzung bzw. -erfassung ist häufig unzureichend.
- Aufgrund des geweckten öffentlichen Interesses ist eine unverzerrte Datenerfassung (Befragung) fast unmöglich.

Der dritte Punkt, der zu einer Überschätzung der relativen Fallzahl führt, ist auch unter der Metapher des „texanischen Scharfschützen“ bekannt: Zuerst schießt der Schütze auf ein Scheunentor, und anschließend legt er um das Zentrum der Einschüsse eine Zielscheibe, sodass einem späteren Betrachter eine hohe Treffsicherheit suggeriert wird.

Insbesondere den wissenschaftlichen Nutzen von Clusteranalysen schätzt Rothman als sehr gering ein: Es gibt zwar umweltepidemiologische Analysen, die zeigen, dass nach einer Exposition gegenüber Schadstoffen das Erkrankungsrisiko für damit in Verbindung stehende Krebsdiagnosen angestiegen ist (z. B. [6]), oder auch vereinzelt Berichte über Krebscluster, bei denen es durchaus gelang, eine relevante Exposition gegenüber bekannten Humankanzerogenen nachzuweisen, die die regionale Erhöhung verursacht haben könnten [7]. Doch wurde tatsächlich weltweit erst ein einziges Mal beschrieben, dass durch ein regionales Cluster Hinweise auf bislang unbekannte umweltbezogene Ursachen von Krebserkrankungen gefunden wurden (Häufung von Mesotheliomen in der türkischen Provinz Kappadokien aufgrund einer Eronit-Exposition [5]).

Demgegenüber wird die Auffassung vertreten, dass derartige Untersuchungen dennoch unumgänglich seien, um mit der öffentlichen Beunruhigung umzugehen und insbesondere das Vertrauen in die Institutionen des öffentlichen Ge-

sundheitsdienstes zu bewahren [5]. Somit liege der Nutzen außerhalb der Wissenschaft ebenso wie auch die Diskussion über den Cluster selbst *außerhalb* stattfindet.

Monitoring als Alternative?

Auf ein prospektives Monitoring durch ein etabliertes Register, das zu (geplanten) Clusterstudien führt, treffen einige der aufgeführten Kritikpunkte gegen Krebsclusteruntersuchungen nicht oder nur modifiziert zu, da es ein einheitliches, standardisiertes Vorgehen gewährleistet:

- Das Problem des „texanischen Scharfschützen“ ist weniger kritisch, da die Bezugspopulation vorab bereits beim Monitoring definiert wird (z. B. durch Gemeindegrenzen).
- Die Einschätzung erfolgt auf Grundlage bereits validierter Krebsregisterfälle.
- Die nicht korrigierbare Verzerrung, die durch die implizite Selektion von *auffälligen* Regionen entsteht, tritt bei geplanten (!) Clusteruntersuchungen nicht auf.

Weitere Argumente für ein Monitoring sind:

- Es unterstützt die Aufklärung über die Häufigkeit von Krebserkrankungen, deren Risiken und Präventionsmöglichkeiten.
- Bei seltenen Diagnosen kann durch Zusammenfassung verschiedener auffälliger Regionen evtl. eine ausreichende Fallanzahl für weitergehende evaluierende Studien erreicht werden.
- Bis zum Zeitpunkt eines Berichtes über eine mögliche regionale Erhöhung von Krebsfällen ist keine Verzerrung der erhobenen Daten durch eine öffentliche Diskussion zu erwarten.
- Viele Einzelanfragen werden damit vorweggenommen: „Alle Gemeinden sind unter Beobachtung“.
- Schnelle Auskünfte über die regionale Situation können helfen, mögliche öffentliche Diskussionen zu möglichen regionalen „Krebsauslösern“ zu versachlichen.

Demgegenüber behalten viele der seit den 1990er-Jahren vorgebrachten Einwände gegen ein prospektives Monitoring von Krebserkrankungen [2] ihre Berechtigung. Bei Krebserkrankungen beträgt die „Latenzzeit“, die zwischen einer mutmaßlichen Exposition gegenüber einem kanzerogenen Stoff und der Erstdiagnose der Erkrankung liegt, oft viele Jahre. Der aktueller Wohnort und der Expositions-ort unterscheiden sich daher oft. Zudem liegen im Register meist nur Informationen zu dem Fall und der Form der Krebserkrankung vor, nicht aber zu möglichen „Risiken“, die über eine reine Wohnort-historie hinausgehen. Es müssen jedoch plausible Expositionsszenarien vorliegen, um zu belegen, dass ein umweltbezogenes Cluster anzunehmen ist. Auch werden Häufungen unterhalb der Gemeindeebene kaum aufgedeckt werden können.

Die zentralen Kritikpunkte sind, dass einerseits – abhängig von der Anzahl der durchgeführten Vergleiche und der gewählten Irrtumswahrscheinlichkeiten – zu viele „falsche Alarme“ gegeben werden, die mit gesellschaftlichen Folgekosten verbunden sind, unter die auch ein durch die Verunsicherung einhergehender Verlust an Lebensqualität zu fassen ist. Andererseits wird ein „Routinemonitoring“ aufgrund der geringen statistischen Power kaum in der Lage sein, umweltbezogene Risiken, die lediglich einen Teil der Gemeinde betreffen, mit ausreichender Wahrscheinlichkeit aufdecken zu können. Zudem ist mit Krebsclusteranalysen die Identifizierung bislang unbekannter umweltbezogener Risikofaktoren praktisch unmöglich, während bei verdächtigen, aber bekannten potenziell kanzerogenen Stoffen die Frage der Regulation und Kontrolle primär zu beantworten wäre, weniger jedoch die exakte Schätzung des möglichen Schadensmaßes in Form vermehrter Krebserkrankungen.

Auch nach Einführung eines routinemäßigen Monitorings muss die Möglichkeit zu anlassbezogenen Untersuchungen bestehen, insbesondere bei Fragen nach möglichen Häufungen innerhalb eines Ortsteils oder auch für Krebsdiagnosen, für die das Monitoring nicht hinreichend sensitiv ausgelegt ist.

Grundlegende Konzeption des niedersächsischen Monitorings

Anforderungsprofil

Die Forderung nach einem routinemäßigen Monitoring kommt aus der Politik und nicht aus der Wissenschaft. Daher müssen unabhängig von den konkret zu benennenden Verfahren und Teststatistiken die konzeptionellen Grundlagen eines solchen Monitorings zum Teil auch gesellschaftspolitisch festgelegt werden. Wichtige festzulegende Eckpunkte zeigt

■ Abb. 1.

Zwei zusammenhängende, antagonistische Forderungen an ein ideales Monitoring verdeutlichen die Problematik der Suche nach „optimalen“ Stellschrauben:

- Das Monitoring sollte möglichst viele „echte“ Inzidenzerhöhungen (regional; spezifische Diagnose) aufdecken.
- Die Wahrscheinlichkeit für einen „falschen Alarm“ bzw. die erwartete Zahl an falschen Alarmen sollte möglichst niedrig sein.

Da diese Forderungen nicht gleichzeitig zu erfüllen sind, müssen die gesellschaftlichen Kosten eines Fehlalarms sowie der potenzielle Nutzen eines aufgedeckten Clusters gegeneinander abgewogen werden, um schließlich zu einem tragfähigen Kompromiss zu kommen.

Insbesondere sind im Vorfeld folgende Fragen zu beantworten:

- Wie viele „Fehlalarme“ sind aufgrund der mit ihnen einhergehenden Beunruhigung der Bevölkerung aber auch eines möglichen Glaubwürdigkeitsverlustes des EKN zu verkraften?
- Welche (echte) Inzidenzerhöhungen müssen mit hinreichend hoher Wahrscheinlichkeit aufgedeckt werden können?
- Welche Krebsdiagnosen sollen als „Leitdiagnosen“ für das Monitoring herangezogen werden?
- Wann und wie werden Ergebnisse des Monitorings kommuniziert?
- Wie sehen die Strukturen und Zuständigkeiten für eine vertiefende Krebsclusteruntersuchung aus?

Bundesgesundheitsbl 2014 · 57:33–40 DOI 10.1007/s00103-013-1872-5
© Springer-Verlag Berlin Heidelberg 2013

J. Kieschke · M. Hoopmann

Aktives Monitoring kleinräumiger Krebshäufungen. Vorstellung eines neuen Monitoringkonzeptes für das Epidemiologische Krebsregister Niedersachsen

Zusammenfassung

Hintergrund. Krebsregister erhalten oft Anfragen zu möglichen regionalen Häufungen von Krebserkrankungen. Nach einem Bericht des Epidemiologischen Krebsregisters Niedersachsen (EKN) zu Leukämieerkrankungen in der Samtgemeinde (SG) Asse 2010 bekam es vom Land Niedersachsen den Auftrag, ein regionales Krebsmonitoring zu konzipieren. Das Konzept sieht 2 Phasen vor: eine fixe Suchphase über alle regionalen Beobachtungseinheiten (ReBe) sowie für in der Suchphase auffällige ReBe eine sich anschließende Beobachtungsphase. Ein positiver Bestätigungstest gilt als Anfangsverdacht, der eine vertiefende Evaluation einleiten kann.

Methodik. Die Suchphase nutzt das Standardized Incidence Ratio (SIR). Für den Bestätigungstest wird der Sequential Probability Ratio Test (SPRT) eingesetzt, mit dem zu festen Zeitpunkten alle neuen Beobachtungen betrachtet werden. Die Teststatistik des SPRT prüft neben der Nullhypothese (keine Erhöhung) eine Alternativhypothese (Erhöhung), sodass 3 Ergebnisse möglich sind: Entwarnung, Warnung oder Weiterbeobachtung.

Ergebnisse. Bei der Vielzahl von statistischen Tests ist mit einer hohen Anzahl von Fehlalarmen zu rechnen, bei denen die Auffälligkeiten rein zufällig zustande kommen (Problem des multiplen Testens). Die Anzahl hängt ab von der Zahl an getesteten Regionen, von Diagnosen, Subgruppenanalysen sowie von der Periodizität. Vorüberlegungen gehen von 400 regionalen Einheiten, 3 Diagnosegruppen, keinen Subgruppenanalysen und einer Suchphase alle 2 Jahre aus. Mit Vorgaben zu Irrtumswahrscheinlichkeiten sowie aufzudeckender Inzidenzerhöhung ließen sich die entsprechenden Parameter des Monitorings festlegen.

Diskussion. Vor Beginn des Monitorings sind Fragen politisch zu beantworten, z. B. dazu, wie viele Fehlalarme bewältigt werden können, welche Auswirkungen diese auf die Bevölkerung haben und welche Strukturen für die Risikokommunikation vorliegen.

Schlüsselwörter

Monitoring · Clusteranalysen · Multiples Testen · Risikokommunikation · Krebsregister

Active surveillance of small-scale spatial cancer clusters. Presentation of a new monitoring approach for the Epidemiological Cancer Registry of Lower Saxony

Abstract

Background. Cancer registries often receive inquiries on possible cancer clusters. In 2010, the Cancer Registry of Lower Saxony reported on a spatial leukemia accumulation in the community of Asse. As a result, the Cancer Registry was engaged in developing an approach for a regional cancer monitoring program by the federal government. The modular approach involves a two-stage process. First, all regional monitoring areas are under surveillance, and in the second phase, noticeable areas only are observed further. A positive confirmation test is regarded as an initial concern and may lead to further investigations.

Methods. The standardized incidence ratio (SIR) is used in the first stage. For confirmation, the Sequential Probability Ratio Test (SPRT) is applied in the second stage. All new observations are tested at a given time. The SPRT tests the null hypothesis (no increase) and an alternative hypothesis (increase) leading to three possible conclusions: warning, all-clear, or further surveillance necessary.

Results. A high false alarm rate due to clustering by chance is to be considered with multiple statistical testing. The number of false alarms depends on the number of tested areas, diagnoses, subgroup analyses, and time periods. Preliminary considerations comprise 400 spatial areas, three diagnosis groups, no subgroup analyses, and a biannual first-stage testing. Guidelines for the level of significance and for detecting an increase in incidence will be necessary for the assessment of the monitoring parameters.

Discussion. Before running the monitoring program, a number of questions have to be handled politically; e.g., how many false alarms can be handled, the impact on the population involved, and existing risk communication structures.

Keywords

Surveillance · Cluster analysis · Multiple statistical testing · Risk communication · Cancer registry

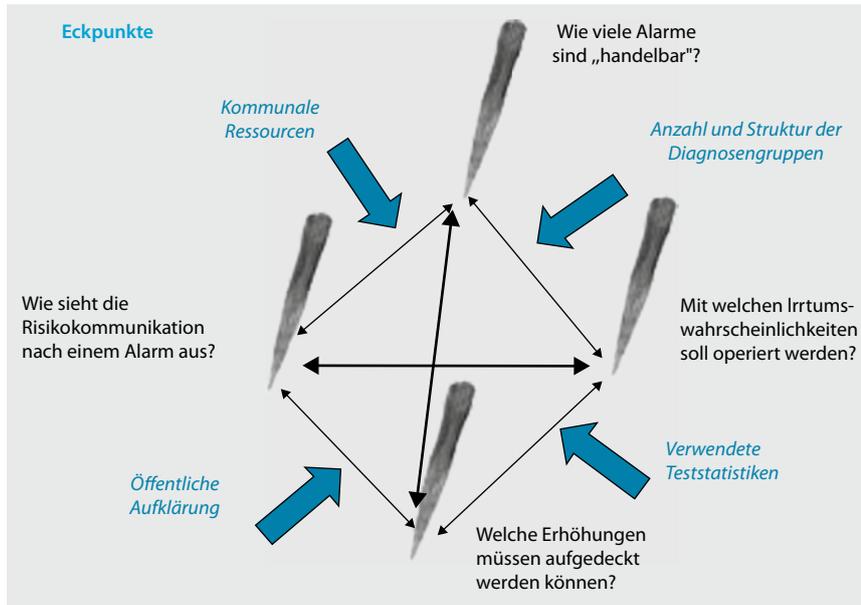


Abb. 1 ▲ Festzulegende Eckpunkte bei der Konzeptionierung eines Monitorings regionaler Krebshäufungen

Das Monitoring-Konzept

Beim Monitoring wird auf eine zweistufige Konzeption mit Such- und Bestätigungsphase zurückgegriffen, dessen Grundgedanke bereits 1999 publiziert wurde [8].

In der vorgelagerten Suchphase werden Gemeinden ermittelt, deren altersstandardisierte Inzidenzrate für Krebserkrankungen über einen Zeitraum von 5 Jahren erhöht ist (einseitig getestet, 5% Signifikanzniveau). Diese Suchphase erzeugt quasi die Hypothesen darüber, in welchen Gemeinden Krebserkrankungen vermehrt auftreten. In einer nachgeschalteten prospektiven Beobachtungsphase wird für diese indizierten Gemeinden überprüft, ob weiterhin eine Erhöhung gegenüber dem Erwartungswert vorliegt. Überschreitet eine regionale Einheit auch in der folgenden Beobachtungsphase den kritischen Wert der Teststatistik, so wird dies als Bestätigung eines zeitlich-räumlichen Krebsclusterverdachts durch das Krebsregister interpretiert.

Zur Konzeption eines Monitorings gehört die Festlegung auf die

- regionalen Einheiten,
- zu analysierenden Diagnosen(gruppen) und Subgruppen (z. B. Geschlecht und Alter),

- Häufigkeit der Durchführung des Monitorings (Periodizität) und die Zeitnähe.

Zudem sind geeignete Testverfahren festzulegen. Während für die Suchphase die Standardverfahren, die auch in den Jahresberichten beschrieben sind, eingesetzt werden, wird in der Beobachtungsphase auf sequenzielle Testverfahren „Sequential Probability Ratio Test“ (SPRT) [9, 10] zurückgegriffen. Hierbei muss im Gegensatz zu anderen Statistiken weder die Anzahl der einzubeziehenden Fälle noch die Beobachtungsdauer vorher festgelegt werden. Sowohl eine deutliche Erhöhung (Ablehnung der Nullhypothese nicht erhöhter Fallzahl) als auch eine zu geringe Anzahl an beobachteten Fällen (Ablehnung der Alternativhypothese einer vorgegebenen Erhöhung) können und sollen zu einer Entscheidung und damit zur Beendigung der Beobachtungsphase führen.

Auf Fragen zu den Kommunikationsprozessen der Ergebnisse wird in der Diskussion eingegangen.

Regionale Einheiten

Gemeinden weisen große Unterschiede in ihrer Bevölkerungsgröße auf. In Großstädten können vergleichsweise gering erhöhte relative Risiken aufgrund des mit der hohen Fallzahl verbundenen gerin-

geren Variationskoeffizienten signifikant sein, während in kleinen Gemeinden selbst hohe relative Risiken wegen der geringen Fallzahl als nicht signifikant eingestuft werden müssen. Um bei regionalen Analysen eine Mindestanzahl von Personen zu erhalten, wurden „Regionale Beobachtungseinheiten“ („ReBe“) mit mindestens 5000 Einwohnern festgelegt. Dabei wurden Gemeinden mit einer Zahl von unter 5000 Einwohnern mit der kleinsten Nachbargemeinde desselben Kreises vereinigt. Eine Untergliederung der (Groß-)Städte in vergleichbare Größenordnungen ist geplant. Zurzeit ist Niedersachsen in 389 ReBe mit durchschnittlich etwa 20.000 Einwohnern aufgeteilt.

Diagnosen- und Subgruppen

Zur Begrenzung der Einzelvergleiche sind anfänglich nur 3 Diagnosengruppen im Routinemonitoring vorgesehen:

1. Leukämien, Lymphome und Schilddrüsenkrebs bilden die erste Gruppe, da sie immer wieder im Fokus umweltpolitischer Diskussionen stehen und bei ihnen umweltbezogene Risiken (z. B. ionisierende Strahlung) bekannt sind.
2. Die zweite Gruppe setzt sich aus sehr seltenen Krebserkrankungen zusammen, zu deren Ätiologie bisher wenig bekannt ist, die aber im Zusammenhang mit Umwelttoxinen stehen könnten. Für diese Gruppe könnten Clusteruntersuchungen hypothesengenerierend und bezüglich etwaiger Risikofaktoren sogar wissenschaftlich von Nutzen sein [2].
3. Die dritte zusammenfassende „Rest“-Gruppe umfasst die übrigen Krebsdiagnosen mit Ausnahme derjenigen, die durch Früherkennungsuntersuchungen detektiert werden (z. B. Brust-, Prostata- und Darmkrebs) oder die vornehmlich mit individuellen Risikofaktoren in Verbindung stehen (wie Rauchen und Lungenkrebs). Mit diesem Ausschluss soll die Möglichkeit einer Maskierung reduziert werden, d. h., dass eine Erhöhung einer selteneren Krebsdiagnose aufgrund der Zufallsschwankungen von zahlenmäßig häufigeren Krebserkrankungen überdeckt wird.

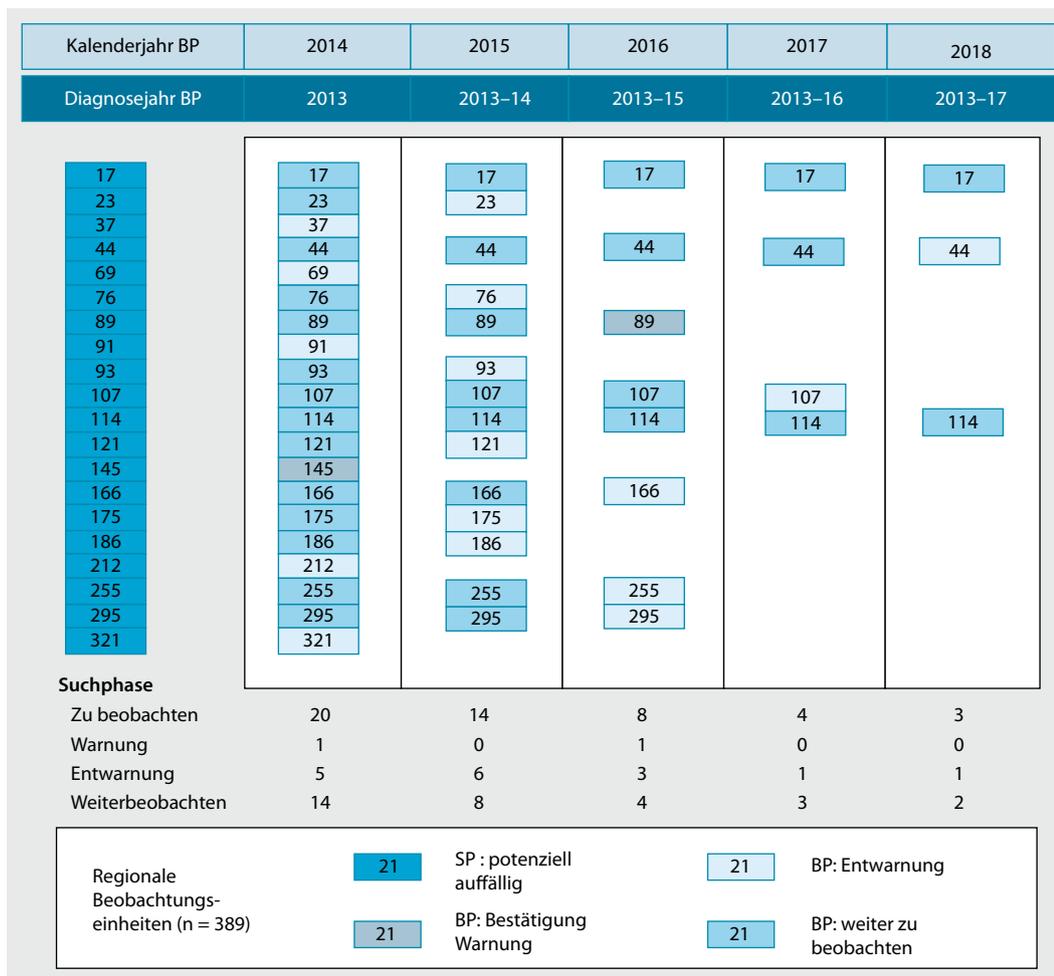


Abb. 2 ◀ Exemplarischer Ablauf von Such- und Beobachtungsphase des regionalen Krebsmonitorings des Epidemiologischen Krebsregisters Niedersachsens (eine Diagnosengruppe, 389 ReBe, SP Suchphase, BP Beobachtungsphase)

Auf geschlechts- und altersspezifische Subgruppenanalysen wird beim Routine-monitoring verzichtet. Allerdings würden Subgruppen bei einer vertiefenden Evaluation [1] im Rahmen einer Krebsclusteruntersuchung gezielt untersucht werden.

Periodizität und Zeitnähe

Bei der Festlegung der zeitlichen Abstände ist auf ein ausgewogenes Verhältnis zwischen Aktualität, Sensitivität und Unabhängigkeit der Untersuchungen zu achten.

Daher soll die Suchphase nur alle 2 Jahre durchgeführt werden; dabei werden die beobachteten Erkrankungsfälle der letzten 5 Jahre betrachtet. Als (formal) vollzählig erfasste Diagnosejahrgänge gelten diejenigen, die 2 und mehr Jahre vor dem aktuellen Erfassungsjahr liegen. Der aus ihnen berechnete Erwartungswert für den Zeitraum der Suchphase wird auch für die Beobachtungs-

phase verwendet. In der Beobachtungsphase können hingegen auch die aktuellen, noch nicht vollzähligen Erfassungsjahrgänge insofern genutzt werden, als dass auch bei einer Untererfassung bei stark erhöhten Inzidenzen die Öffentlichkeit zeitnah informiert werden kann. Eine Entwarnung im Sinne einer nicht auffällig erhöhten Inzidenz ist hingegen erst auf Basis der vollzähligen Diagnosejahrgänge möglich.

■ **Abb. 2** zeigt die zeitlichen Abläufe in der Such- und Beobachtungsphase. Eine erste Suchphase im Jahr 2014 mit den Diagnosejahren 2008 bis 2012 ergab 20 zu beobachtende ReBe, von denen eine in der Beobachtungsphase bereits anhand der Daten des Diagnosejahres 2013 bestätigt wird. Für 5 weitere kann eine Entwarnung gegeben werden. 14 verbleiben in der Beobachtungsphase und werden 2015 anhand der Daten der Diagnosejahre 2013 bis 2014 weiter überprüft usw. Alle 2 Jahre erfolgt eine neue Suchphase, un-

abhängig davon, ob bereits für alle ReBe eine Entscheidung getroffen worden ist.

Sequenzielle Testverfahren

In klassischen Testverfahren wird allein die Nullhypothese („keine Erhöhung“) aus statistischer Sicht betrachtet: Entweder wird die Nullhypothese verworfen oder sie kann nicht verworfen werden und wird daher beibehalten. Zwar kann man berechnen, mit welcher Wahrscheinlichkeit die Nullhypothese verworfen wird, wenn tatsächlich eine Alternativhypothese gültig ist, doch wird zur Konstruktion der Teststatistik keine Alternativhypothese benötigt.

Damit kann zwar die Grundforderung an ein Monitoring, nicht zu viele Fehlalarme auszulösen, eingelöst werden, die zweite Grundforderung jedoch, dass Inzidenzerhöhung zumindest ab einer gewissen Größenordnung aufgedeckt werden müssen, wird nicht formalisiert.

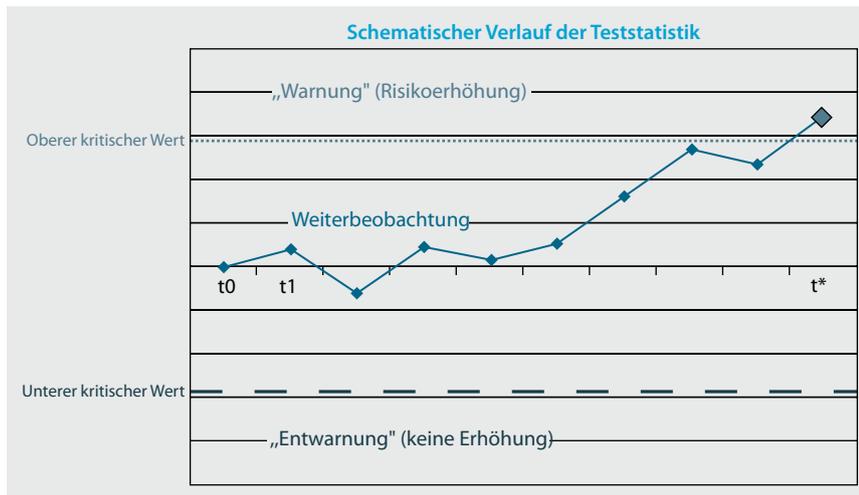


Abb. 3 ▲ Schematischer Verlauf des sequenziellen Testens im regionalen Krebsmonitoring des Epidemiologischen Krebsregisters Niedersachsens

Zur Ableitung des SPRT muss nicht nur die Nullhypothese, sondern auch eine Alternativhypothese – der Grad der Risikoerhöhung – vorgegeben werden. Insofern werden hier beide Grundforderungen in der Wahl des Testverfahrens direkt parallel aufgegriffen.

Beim SPRT wird zu jedem Beurteilungszeitpunkt im Lauf der Beobachtungsperiode getestet, ob die Information ausreicht, um sich für eine der beiden Hypothesen zu entscheiden, oder aber ob weiter beobachtet werden muss (■ **Abb. 3**). Dabei wird die Teststatistik mit 2 kritischen Werten verglichen, die sich aus den beiden Irrtumswahrscheinlichkeiten für die beiden Hypothesen ableiten [10, 11]. In die von einem Poisson-Prozess abgeleitete sequenzielle Teststatistik [11] geht u. a. die Summe aller bis zum Bewertungszeitpunkt beobachteten Fälle ein, wobei jahrgangsweise getestet wird („group sequential test“).

Bei dem SPRT ist weder die Gesamtzahl der Fälle noch die Dauer der Beobachtungsphase vorab festgelegt. Die Verfahren verwenden die fortlaufenden Fallinformationen und bauen sie in die bisherigen Teststatistiken ein. Im Vergleich zu „klassischen Testverfahren“ mit einer fest vorgegebenen Beobachtungsdauer führen sequenzielle Verfahren in der Regel zu einer schnelleren Entscheidung.

Problem des multiplen Testens

In einem Aufsatz zum möglichen Missbrauch von automatisierten geografischen Analysen kommt Neutra [12] in seinem Szenario („80 varieties of cancer ... some 5000 census tracts state wide, a probability value of 0.01“) auf jährlich zu erwartende 4000 falsch positive Ergebnisse. Das zentrale Problem bei einem geplanten kleinräumigen Monitoring ist dabei das multiple Testen [13, 14].

Werden – wie bei einem Routinemonitoring – zahlreiche Teststatistiken parallel berechnet, können allein aufgrund der Vielzahl der durchgeführten Vergleiche zahlreiche Einzelvergleiche nominell betrachtet „signifikante Ergebnisse“ ergeben, wenn auch die Gesamtzahl derartiger Ereignisse durchaus im erwarteten Bereich liegt. Der p-Wert eines Einzelvergleiches darf nicht selektiv herausgegriffen und wie ein singuläres Testergebnis ohne Bezug auf die Gesamtzahl der durchgeführten Vergleiche interpretiert werden. Für Aussagen zur Signifikanz sollte daher die „Gesamtirrtumswahrscheinlichkeit“ berücksichtigt werden. Sie gibt an, mit welcher Wahrscheinlichkeit über die Gesamtheit aller durchgeführten Vergleiche mindestens eine Hypothese irrtümlich abgelehnt wird. Werden beispielsweise 10 unabhängige Tests zu einer jeweiligen Irrtumswahrscheinlichkeit von 0,05 durchgeführt, besteht bereits eine 40%ige Wahrscheinlich-

keit, dass man mindestens eine Hypothese irrtümlich verwirft:

$$1 - 0,95^{10} = 0,40.$$

Für die Risikokommunikation bietet es sich wegen der Problematik der Gesamtirrtumswahrscheinlichkeit an, von „statistisch auffälligen Einzelergebnissen“ zu sprechen und damit den oft falsch gebrauchten Signifikanzbegriff zu vermeiden.

Soll das Niveau der Fehlerrate für alle durchgeführten Tests gelten, muss dies für die Signifikanzbestimmung berücksichtigt werden, z. B. durch die Methode nach Bonferroni [13]. Bei ihr muss der berechnete p-Wert kleiner sein als die gewählte nominelle Signifikanzschwelle dividiert durch die Anzahl der durchgeführten Tests.

Um die Parameter eines kleinräumigen Monitorings in Niedersachsen einschätzen zu können, haben EKN und NLGA die Anzahl von erwarteten Fehlalarmen berechnet. In ■ **Tab. 1** wird in Abhängigkeit von den Irrtumswahrscheinlichkeiten in den beiden Phasen die Anzahl von Fehlalarmen ausgewiesen, die für jede untersuchte Diagnosegruppe bei einem Monitoring über 400 Gemeinden zu erwarten wären. Zusätzlich ist der ungefähre 95%-Vertrauensbereich für die zu erwartende Anzahl von Fehlalarmen angegeben.

Aus ■ **Tab. 1** wird ersichtlich, dass für Niedersachsen mit der zweistufigen Monitoringkonzeption bei einem „üblichen“ Signifikanzniveau von $p=0,05$ (einseitig getestet) für Such- und Beobachtungsphase mit etwa einem falsch positiven Ergebnis pro Diagnose zu rechnen ist.

Diskussion

Die dargestellte Konzeption für Niedersachsen basiert auf „fixen“ Beobachtungseinheiten (Gemeinden bzw. genauer ReBe) und damit auf Verwaltungsgrenzen, während in der Infektionsepidemiologie eingesetzte Clusteralgorithmen grenzüberschreitend suchen. Neue gesetzliche Möglichkeiten des EKN erlauben die Zuordnung von Meldungen auch unterhalb der Gemeindeebene und dürften zukünftig auch alternative Clustersuchstrategien zulassen [15].

Tab. 1 Erwartete Fehlalarme in Abhängigkeit von Irrtumswahrscheinlichkeiten der beiden Phasen (Suchphase/Beobachtungsphase) – (eine Diagnosengruppe; 400 Regionen, einseitig getestet)

α -Fehler Suchphase	Erwartungswert „zufällig auffällige Gemeinden“ nach Suchphase (400 Tests)	α -Fehler Beobachtungsphase	Erwartungswert „Falsche Warnungen“ nach Beobachtungsphase	Ungefähres 95%-KI	Exakte Wahrscheinlichkeit für KI
0,05	20	0,05	1	[0; 3]	0,984
0,04	16	0,04	0,64	[0; 2]	0,976
0,03	12	0,03	0,36	[0; 1]	0,951
0,025	10	0,025	0,25	[0; 1]	0,975
0,02	8	0,02	0,16	[0; 1]	0,990
0,01	4	0,01	0,04	0	0,961
0,1	40	0,01	0,4	[0; 1]	0,939

Die in den sequenziellen Verfahren geforderte konkrete Benennung einer Alternativhypothese in Form einer zu konkretisierenden Risikoerhöhung führt dazu, dass das EKN beide Fehlerwahrscheinlichkeiten abschätzen und auch die voraussichtliche Länge des Beobachtungszeitraums bis zu einer Entscheidung steuern kann. Allerdings könnte die notwendige Festlegung dieser Risikoerhöhung dahingehend missinterpretiert werden, dass geringere Risikoerhöhungen wohl als akzeptabel angesehen werden würden, was zur Forderung führen könnte, dass auch geringere Erkrankungsrisiken aufgedeckt werden müssten.

Die zumeist indirekt entstehenden gesellschaftlichen Belastungen und Kosten durch das Monitoring sind schwer zu fassen. Die größte Herausforderung des Monitorings liegt im Umgang mit und in der Interpretation der Ergebnisse, gerade auch für die Öffentlichkeit. Bei einem mehrstufigen Monitoring ist das Recht der Bevölkerung auf frühzeitige Information mit dem möglichen Schaden durch Fehlalarme abzuwägen. Eine öffentliche Warnung des Krebsregisters kann nicht einfach ungeschehen gemacht werden. Negative Informationen, wie eine lokal erhöhte Krebshäufigkeit, werden von den Medien bevorzugt aufgegriffen, anders als Berichte über Entwarnungen. Neben einer unnötigen Beunruhigung der Bevölkerung, kann es auch zu Werteverlusten von Immobilien und zu einer sinkenden touristischen Attraktivität der Gemeinden kommen.

Die Entscheidung, die Öffentlichkeit zu informieren, darf daher nicht leicht-

fertig getroffen werden. Laut Hardy [16], müssen folgende Aspekte gegeneinander abgewogen werden:

1. die Nachteile der Erklärung, es bestünde eine empirisch auffällige Erhöhung, wenn es in Wirklichkeit keine Risikoerhöhung gibt;
2. die Nachteile, die damit verbunden sind, dass ein bestehendes Gesundheitsrisiko nicht aufgedeckt wird;
3. die Wahrscheinlichkeit, dass die alarmierten Behörden zu der richtigen Feststellung der Ursache für die Erhöhung kommen werden;
4. der relative Nutzen für die Gesellschaft jetzt und in Zukunft, der daraus erwächst, dass ein Gesundheitsrisiko identifiziert werden konnte.

Die unter Punkt 1 und 2 genannten Kosten führen zur Gewichtung von Spezifität und Sensitivität des Monitorings und haben unmittelbaren Einfluss auf die anzusetzenden Fehlerwahrscheinlichkeiten. Der unter Punkt 4 angesprochene Nutzen entfällt, wenn ein die Inzidenzerhöhung (möglicherweise) verursachender Emitent bereits stillgelegt worden ist (keine Risikominimierung mehr möglich). Bei bekannten kanzerogenen Stoffen kann ein Nutzen an sich nur erwartet werden, wenn die bestehenden behördlichen Überwachungssysteme zur Vermeidung potenzieller Gefährdungen für die Belegschaft und Anwohner versagt haben.

Aus Gründen der „Transparenz“ können die aufgeführten Punkte nicht erst bei Unterrichtung der kommunalen Behörden bewertet werden. Parallel zu den Kriterien des Monitorings selbst ist da-

her ein Bewertungsschema darüber zu erarbeiten, wie mit den Ergebnissen umgegangen wird, insbesondere wer auf welcher Stufe und in welchem Umfang zu informieren ist. Eine standardisierte Unterstützung des Öffentlichen Gesundheitsdienstes vor Ort ist für die Risikokommunikation unerlässlich.

Auch Kriterien für etwaige vertiefende Vor-Ort-Untersuchungen sind vorher festzulegen, um standardisierte Vorgehensweisen zu erreichen. Für eine weitere Abklärung sind rechtliche Grundlagen der Krebsregistrierung nötig zur vollzähligen Erfassung der Fälle (Meldepflicht), zur Durchführbarkeit kleinräumiger Subanalysen unterhalb der Gemeindeebene sowie zur Möglichkeit einer direkten Kontaktaufnahme mit den Betroffenen.

Eine wesentliche Hoffnung ist, dass sich mit dem Monitoring mögliche Clusteranfragen auf einer frühen Ebene beantworten lassen. Ob sich diese Hoffnung erfüllen wird, bleibt abzuwarten. Eine Evaluation der Auswirkungen des Monitoringkonzeptes ist spätestens 5 Jahre nach Beginn des Routinebetriebs vorgesehen.

Fazit

Als Erklärung für eine erhöhte regionale Krebsinzidenz wird von Bürgern häufig eine lokale umweltbezogene Risikoerhöhung angenommen. Die Forderung nach einem routinemäßigen regionalen Monitoring mit Krebsregisterdaten kommt aus dem politischen Bereich und nicht aus der Wissenschaft, die den Nutzen im Sinne eines ursächlichen Erkenntnisgewinns als (sehr) gering einschätzt. Mit dem hier vorgestellten Konzept eines zweistufigen Monitorings lassen sich viele der früher im Vordergrund stehenden methodisch-statistischen Probleme kontrollieren. Eine neue gesetzliche Grundlage verbessert für das EKN zudem die Möglichkeiten, im Falle einer bestätigten Auffälligkeit eine vertiefende Evaluation durchzuführen. Unabhängig von den konkret zu benennenden Verfahren und Teststatistiken müssen wichtige Fragen auch gesellschaftspolitisch festgelegt bzw. bewertet werden, vor allem: Wie hoch sind die gesellschaftlichen Kosten von Fehlalarmen

im Gegensatz zum Nutzen eines aufgedeckten Clusters und welche Strukturen für die Risikokommunikation liegen vor? Hierzu stehen Antworten noch aus. Die größte Herausforderung des Monitorings liegt nicht in der statistisch-epidemiologischen Konzeption, sondern im Umgang und der Interpretation der Ergebnisse.

Jede Einführung eines prospektiven kleinräumigen regionalen Monitorings birgt Risiken, weshalb sie immer evaluiert werden sollte. Ohne sorgfältige Konzeption – auch für den Bereich der Risikokommunikation – kann sie nicht vorbehaltlos empfohlen werden. Vor Einführung des Monitorings sind seine Möglichkeiten öffentlich zu kommunizieren, um unrealistische Erwartungen von vornherein zu reduzieren. Man hofft, dass sich mit dem Monitoring Clusteranfragen auf einer frühen Ebene beantworten lassen und dass auffällige Inzidenzerhöhungen in einem standardisierten Verfahren abgeklärt werden können.

Korrespondenzadresse

J. Kieschke

Epidemiologisches Krebsregister
Niedersachsen, Registerstelle
Industriestr. 9, 26121 Oldenburg
kieschke@krebsregister-niedersachsen.de

Einhaltung ethischer Richtlinien

Interessenkonflikt. J. Kieschke und M. Hoopmann geben an, dass kein Interessenkonflikt besteht.

Dieser Beitrag beinhaltet keine Studien an Menschen oder Tieren.

Literatur

1. Schümann M, Bolte G, Heinzow B et al (2009) Empfehlungen für den Umgang mit Beobachtungen von räumlich-zeitlichen Krankheitsclustern. Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz 52:239–255
2. Rothman K (1990) A sobering start for the cluster busters' conference. Am J Epidemiol 132(Suppl 1):6–13
3. Centers for Disease Control and Prevention (CDC) (1990) Guidelines for investigating clusters of health events. MMWR Morb Mortal Wkly Rep 39(RR-11):1–17.

4. Washington State Department of Health (2007) (rev.) Guidelines for investigating clusters of chronic disease and adverse birth outcomes. Washington State Department of Health, Washington, DC. <http://www.doh.wa.gov/Portals/1/Documents/1200/EP2-2-2Whatcom.pdf>. Zugegriffen: 28. November 2013
5. Neutra R (1990) Counterpoint from a clusterbuster. Am J Epidemiol 132:1–8
6. Lampi P, Tuomisto J, Hakulinen T et al (2008) Follow-up study of cancer incidence after chlorophenol exposure in a community in southern Finland. Scand J Work Environ Health 34:230–233
7. Thun M J, Sinks T (2004) Understanding cancer clusters. CA Cancer J Clin 54:273–280
8. Kieschke J, Wellmann I (1999) Conception for monitoring in an epidemiological cancer registry. In: Victor N et al (Hrsg) Medical informatics, biostatistics and epidemiology for efficient health care and medical research. Urban und Vogel, München, S 101–104
9. Wald A (1947) Sequential analysis. Wiley, New York
10. Spiegelhalter DJ, Grigg OA, Kinsman R et al (2003) Sequential probability ratio tests (sprts) for monitoring risk-adjusted outcomes. Int J Qual Hlth Care 15:1–7
11. Haggstrom GW (1979) Sequential tests for exponential populations and poisson processes. The Rand Cooperation, Santa Monica. <http://www.rand.org/pubs/papers/2008/P6336.pdf>. Zugegriffen: 26. November 2013
12. Neutra RR (1999) Computer geographic analysis: a commentary on its use and misuse in public health. In: Lawson A et al (Hrsg) Disease mapping and risk assessment for public health. Chichester
13. Horn M, Vollandt R (1995) Multiple Tests und Auswahlverfahren. G. Fischer, Stuttgart
14. Victor A, Elsässer A, Hommel G et al (2010) Wie bewertet man die p-Wert-Flut? Dtsch Arztebl Int 107(4):50–60
15. Neufassung des Gesetzes über das Epidemiologische Krebsregister Niedersachsen (GEKN) vom 7.12.2012; Nds. GVBl. 31/2012, ausgegeben am 13.12.2012, S 550–557
16. Hardy RJ, Schroder GD, Cooper SP et al (1990) A surveillance system for assessing health effects from hazardous exposures. Am J Epidemiol 132:S32–S42