

EKN

Registerstelle

Epidemiologisches
Krebsregister
Niedersachsen



Auswertung des EKN zur Häufigkeit von Krebsneuerkrankungen in der Samtgemeinde Bothel

Oldenburg, September 2014

Registerstelle des EKN
Joachim Kieschke MPH – Ärztlicher Leiter der Registerstelle
Claudia Vohmann
Industriestr. 9
26121 Oldenburg
Tel. 0441 361056-12

Inhaltsverzeichnis

1	Hintergrund.....	1
2	Strukturen des EKN.....	1
2.1	Meldeverfahren des EKN	2
2.2	Verschlüsselung der Personendaten – Kontrollnummern und Chiffre.....	2
3	Methodik.....	2
3.1	Hypothese und Vorgehen.....	2
3.2	Untersuchungs- und Vergleichsregion.....	4
3.3	Geschätzte Vollzähligkeit der erfassten Krebsneuerkrankungen im Bezirk Lüneburg	5
3.4	Epidemiologische Maßzahlen.....	5
3.5	Berücksichtigung des multiplen Testens	6
3.6	Ergänzende Statistiken außerhalb der Hauptfragestellung.....	8
4	Ergebnisse	8
4.1	Beschreibung: Häufigkeit von Krebsneuerkrankungen.....	8
4.2	Statistische Bewertung, ob eine Erhöhung der Krebsneuerkrankungen vorliegt (statistische Induktion).....	10
5	Ergänzende Analysen: Leukämien und Lymphome bei Männern.....	10
5.1	Einzeldiagnosen	10
5.2	Zeitliche Entwicklung der Neuerkrankungshäufigkeit.....	11
5.3	Altersverteilung der an Leukämien und Lymphomen neu erkrankten Männer ...	11
6	Diskussion	13
7	Zusammenfassung.....	14
8	Literatur.....	15
9	Anhang	16
	Schließende Statistik.....	16

Auswertung des EKN zur Häufigkeit von Krebsneuerkrankungen in der Samtgemeinde Bothel

1 Hintergrund

Diese Sonderauswertung wurde auf Anfrage des Landkreises Rotenburg vom 13.06.2014 durchgeführt. Das Epidemiologische Krebsregister Niedersachsen (EKN) wird darin um eine Untersuchung der Häufigkeit von Krebserkrankungen in der Samtgemeinde (SG) Bothel gebeten. Hintergrund ist eine Anfrage und eine Unterschriftensammlung einer Bürgerinitiative im Landkreis Rotenburg zu einer vermuteten Krebshäufung in der Bevölkerung, ausgelöst durch die Erdgasförderung in Söhlingen (einem Ortsteil der Gemeinde Hemslingen, die Mitgliedsgemeinde der Samtgemeinde Bothel ist). Eine Arbeitsgruppe, bestehend aus Mitgliedern mehrerer Bürgerinitiativen, Mitarbeiterinnen und Mitarbeitern des Gesundheitsamtes des Landkreises, des Niedersächsischen Landesgesundheitsamtes (NLGA) und der Vertrauensstelle des EKN, hat die Anfrage gemeinsam erarbeitet.

In der Anfrage werden die zu betrachtenden Diagnosegruppen, der Untersuchungszeitraum und das Untersuchungsgebiet festgelegt. Für die Samtgemeinde Bothel soll untersucht werden, ob für zwölf verschiedene Diagnosegruppen über den Zeitraum von 2003 bis 2011 bzw. 2012 eine erhöhte Krebshäufigkeit vorliegt. Die drei häufigsten Diagnosegruppen sollen geschlechtsspezifisch betrachtet werden. Die Ergebnisse der Auswertung werden in diesem Bericht vorgestellt.

2 Strukturen des EKN

Im Jahr 2000 hat das Land Niedersachsen das Niedersächsische Krebsregister eingerichtet. Der Aufbau des EKN erfolgte in einem Stufenausbau – im jährlichen Abstand wurden die Bezirke Weser-Ems (2000), Lüneburg (2001), Braunschweig (2002) und Hannover (2003) in die flächendeckende Erfassung von Krebsneuerkrankungen integriert.

Das EKN besteht aus zwei räumlich, organisatorisch und personell getrennten Arbeitseinheiten: Die Vertrauensstelle (VST), die organisatorisch am NLGA in Hannover angesiedelt ist, erfasst und prüft die eingehenden Meldungen zu Krebserkrankungen; sie entscheidet darüber hinaus über Anträge auf Herausgabe und Nutzung von Daten und koordiniert Anfragen aus der Bevölkerung. Die Aufgaben der Registerstelle (RST) in Oldenburg bestehen in der weiteren Bearbeitung und Zusammenführung von Meldungen, der langfristigen Speicherung der Registerdaten sowie der Erstellung epidemiologischer Routine- und Sonderauswertungen [2].

Der Arbeitsbereich Umweltepidemiologie des NLGA unterstützt VST und RST bei Anfragen mit möglichen Umweltfaktoren als Auslöser einer vermuteten Häufung von Krebserkrankungen.

2.1 Meldeverfahren des EKN

Die Meldeverfahren des EKN sind im Gesetz über das Epidemiologische Krebsregister Niedersachsen (GEKN) geregelt. Im bis Ende 2012 geltenden GEKN (von 1999) war ein Melderecht für ambulant und klinisch tätige Ärztinnen und Ärzte mit Patientenkontakt und mit Einwilligung der Betroffenen geregelt (Melderecht mit Einwilligung). Zusätzlich bestand in Niedersachsen für Ärztinnen und Ärzte, die keinen Patientenkontakt hatten, eine Meldepflicht (z. B. Pathologien, Zytologien). Aus Datenschutzgründen wird für die Pflichtmeldungen nur ein reduzierter Datensatz dauerhaft im EKN gespeichert und eine kleinräumige Zuordnung dieser Meldungen zum genauen Wohnort der Betroffenen (unterhalb der Samtgemeindeebene) ist nicht möglich. Ab dem 1. Januar 2013 besteht in Niedersachsen für alle Ärztinnen und Ärzte, die eine Tumorerkrankung feststellen oder behandeln, eine Meldepflicht (Neufassung des GEKN); eine kleinräumige Zuordnung ist für alle Meldungen möglich. Darüber hinaus erhält das EKN von den Gesundheitsämtern die Todesbescheinigungen und von den Meldeämtern die Angaben zu Verstorbenen.

2.2 Verschlüsselung der Personendaten – Kontrollnummern und Chifftrate

Aus Datenschutzgründen werden alle Personenangaben in der Vertrauensstelle des EKN verschlüsselt (pseudonymisiert). Dafür werden nicht dechiffrierbare Kontrollnummern gebildet. Die dauerhafte Speicherung der verschlüsselten Meldungen findet in der Registerstelle des EKN statt. Über die Kontrollnummern können in der Registerstelle die verschiedenen Meldungen zu einer Person auf pseudonymem Weg zusammengeführt werden. Für Patientinnen bzw. Patienten, die nach dem alten GEKN (von 1999) mit deren Einwilligung gemeldet wurden, wurde zusätzlich ein Chifftrat aus den Personendaten gebildet, welches für spätere Studien ggf. dechiffriert werden kann. Nur diese Erkrankten können später um ihr Einverständnis zur Teilnahme an Studien zu Krebserkrankungen gebeten werden. Wenn im EKN zu einem nach altem GEKN gemeldeten Krebsfall nur eine Pathologiemeldung vorliegt, ist es nicht möglich, die Patientin oder den Patienten zu kontaktieren und um weitere Informationen, z.B. zu möglichen Risikofaktoren, zu bitten.

Die in dieser Auswertung betrachteten Erkrankungen umfassen ausschließlich Neuerkrankungen, die nach dem alten GEKN gemeldet wurden.

3 Methodik

Zur Auswertung kamen die dem EKN gemeldeten Krebsneuerkrankungsfälle der Diagnosejahre 2003 – 2012 einschließlich der DCO-Fälle ('death certificate only', d.h. Fälle, die dem Krebsregister ausschließlich über eine Sterbemeldung bekannt werden und für die als Diagnosejahr das Sterbejahr angenommen wird). Datenstand ist der 02.07.2014.

3.1 Hypothese und Vorgehen

Die durchgeführten Analysen sind auf die formulierte Fragestellung der Anfrage ausgerichtet. Demnach werden die zwölf Diagnosegruppen untersucht, die in Tabelle 1 aufgelistet sind. Diese decken rund 98% aller beobachteten Krebsfälle (ohne nicht-melano-

tischen Hautkrebs) ab. Geschlechtsspezifische Untersuchungen werden für die drei häufigsten Diagnosegruppen (auf Basis der erwarteten Neuerkrankungszahlen) 2, 4 und 12 sowie für die Gruppen 6, 7 und 8 vorgenommen. Für die anderen Diagnosegruppen werden Männer und Frauen zusammen ausgewertet.

Tabelle 1: Zu untersuchende Diagnosegruppen gemäß der Anfrage des Landkreises Rotenburg

Diagnosegruppe	Krebslokalisierung bzw. -diagnose	ICD-10-Code	Geschlecht
1	Mund und Rachen	C00-C14	m + w
2a 2b	Verdauungsorgane (ohne Leber)	C15-C26 (ohne C22)	m w
3	Leber	C22	m + w
4a 4b	Kehlkopf, Lunge	C32-C34	m w
5	Knochen, Haut, Weichteilgewebe (ohne nicht-melanotischen Hautkrebs)	C40-C49 (ohne C44)	m + w
6	Brustdrüse	C50	w
7	Weibliche Genitalorgane	C51-C58	w
8	Prostata, Hoden	C61-C62	m
9	Niere, Harnblase	C64-C68	m + w
10	Hirntumoren	C70-C72	m + w
11	Endokrine Drüsen	C73-C75	m + w
12a 12b	Leukämien & Lymphome*	C81-C96	m w

m = Männer, w = Frauen, m+w = Männer und Frauen gesamt

** offizielle Bezeichnung lt. ICD-10: Bösartige Neubildungen des lymphatischen, blutbildenden und verwandten Gewebes*

Als statistische Nullhypothese wurde formuliert:

„Die Anzahl an Krebsneuerkrankungen in der SG Bothel ist für jede betrachtete Diagnosegruppe kleiner oder gleich der erwarteten Fallzahl“.

Die Alternativhypothese lautet:

„In der SG Bothel treten für mindestens eine Diagnosegruppe mehr Krebsneuerkrankungsfälle auf, als zu erwarten wären.“

Diese Fragestellung ist einseitig formuliert.

Daraus ergeben sich insgesamt 15 durchzuführende Untersuchungen (im Folgenden als 15 Untersuchungsgruppen bezeichnet), die mit einer Gesamtirrtumswahrscheinlichkeit von 5% vorgenommen werden (vgl. Kap. 3.5).

Während die Darstellung des 95%-Konfidenzintervalls des SIR (Erläuterungen siehe Kapitel 3.4) wie üblich auf zweiseitigen Tests beruht und damit auch Abweichungen unterhalb des Erwartungswertes auf lokalem Niveau erkennbar werden, erfolgte die Berechnung des p-Werts einseitig analog zur Fragestellung, ob die Beobachtungswerte erhöht seien.

Deskriptiv wird bei allen derartigen Untersuchungen die Häufigkeit von Krebs insgesamt mit untersucht. Dieses geschieht unter anderem auch, um eventuell Hinweise für Unterschiede im Meldeverhalten und damit zur Vollzähligkeit der Erfassung erkennen zu können.

Die Berechnungen wurden überwiegend mit der für Krebsregister spezifischen Auswertungssoftware CARESS durchgeführt. Die Berechnung der p-Werte erfolgte mit der R-Routine Exact Poisson test¹.

3.2 Untersuchungs- und Vergleichsregion

Untersuchungsregion ist die Samtgemeinde (SG) Bothel. Sie liegt im südlichen Teil des Landkreises (LK) Rotenburg, der zum ehemaligen Regierungsbezirk Lüneburg gehört. In ihr sind die Mitgliedsgemeinden Bothel, Brockel, Hemsbünde, Hemslingen, Kirchwalsede und Westerwalsede zusammengeschlossen. Im Untersuchungszeitraum 2003 - 2012 wohnten durchschnittlich 4.327 Männer und 4.337 Frauen in der SG Bothel. Die Bevölkerungszahlen für die Einzeljahre gehen aus Tabelle 2 hervor. Eine Auswertung für einzelne Mitgliedsgemeinden ist aufgrund der z.T. nicht genügend kleinräumig vorliegenden Daten für die untersuchten Diagnosejahre nicht möglich.

Vergleichsregion für diese Untersuchung ist der ehemalige Regierungsbezirk Lüneburg (im Weiteren ‚Bezirk‘ genannt), dessen altersspezifische Raten zur Berechnung der zu erwartenden Fallzahlen genutzt werden. Die durchschnittliche jährliche Bevölkerung der Vergleichsregion Bezirk Lüneburg beläuft sich auf 835.328 Männer und 861.958 Frauen. Die in die Inzidenzanalyse eingehenden Personenjahre des Diagnosezeitraums 2003 - 2012 sind in Tabelle 2 aufgeführt. Für die SG Bothel liegt der Anteil der Personenjahre bei 0,5% der Vergleichsregion.

Tabelle 2: Bevölkerung der SG Bothel und des Bezirks Lüneburg

Bevölkerung	SG Bothel			Bezirk Lüneburg		
	Männer	Frauen	Männer + Frauen	Männer	Frauen	Männer + Frauen
2003	4.327	4.385	8.712	833.244	862.070	1.695.313
2004	4.360	4.388	8.748	835.823	864.880	1.700.703
2005	4.375	4.379	8.754	837.322	866.231	1.703.552
2006	4.372	4.365	8.737	837.717	865.819	1.703.536
2007	4.366	4.343	8.709	837.324	864.712	1.702.035
2008	4.354	4.318	8.672	836.030	862.858	1.698.888
2009	4.326	4.321	8.647	834.452	860.697	1.695.149
2010	4.289	4.314	8.603	833.913	859.033	1.692.946
2011	4.266	4.296	8.562	833.683	857.293	1.690.976
2012	4.240	4.264	8.504	833.773	855.988	1.689.760
Personenjahre 2003 - 2012	43.274	43.372	86.646	8.353.279	8.619.577	16.972.856

¹ R version 2.11.1 © 2010 The R Foundation For Statistical computing
'poisson.test(x, T = 1, r = 1, alternative = c("greater"))'

3.3 Geschätzte Vollzähligkeit der erfassten Krebsneuerkrankungen im Bezirk Lüneburg

Gemäß der geschätzten Inzidenz des Robert Koch-Instituts (RKI Nds. 2009) liegt der Erfassungsgrad von Krebsneuerkrankungen für den Bezirk Lüneburg für die Diagnosejahre 2003 - 2012 für Krebs insgesamt bei über 95%. Eine Vollzähligkeit von mindestens 90% gilt bundesweit als wesentliche Voraussetzung für wissenschaftlich fundierte Aussagen zur Häufigkeit von Krebsneuerkrankungen. Der Erfassungsgrad für die zehn Einzeljahrgänge geht aus der Tabelle 3 hervor. Internationalem Vorgehen entsprechend wird der nicht-melanotische Hautkrebs (ICD-10 C44) nicht in die Fallzahlen für Krebs insgesamt einbezogen.

Tabelle 3: Geschätzte Vollzähligkeit der Krebsneuerkrankungen im EKN für den Bezirk Lüneburg für Krebs insgesamt (ohne C44; Schätzung nach RKI 2009, EKN-Stand 02.07.2014)

Bezirk Lüneburg Diagnosejahre	Vollzähligkeit der Krebsneuerkrankungen im EKN		
	Männer	Frauen	Männer + Frauen
2003	> 95%	89,9%	93,4%
2004	> 95%	> 95%	> 95%
2005	> 95%	94,7%	> 95%
2006	> 95%	> 95%	> 95%
2007	> 95%	> 95%	> 95%
2008	> 95%	> 95%	> 95%
2009	> 95%	> 95%	> 95%
2010	> 95%	> 95%	> 95%
2011	> 95%	> 95%	> 95%
2012	88,0%	91,1%	89,4%
2003 - 2012	> 95%	> 95%	> 95%

3.4 Epidemiologische Maßzahlen

Altersspezifische Raten

Die altersspezifischen Raten der SG Bothel werden denen der Vergleichsregion Bezirk Lüneburg gegenübergestellt. Die Raten werden gebildet aus der Anzahl von Krebsfällen in einer Altersklasse, dividiert durch die durchschnittliche Bevölkerung der jeweiligen Altersklasse, wobei die Ergebnisse pro 100.000 der Bezugsbevölkerung angegeben werden.

Altersstandardisierte Rate

Für die in der SG Bothel erhöht beobachteten Krebsdiagnosen wird die altersstandardisierte Rate ausgewiesen (Standardbevölkerung Europa). Die altersstandardisierten Raten lassen Vergleiche von verschiedenen Regionen zu, die eine unterschiedliche Altersstruktur aufweisen. Die altersstandardisierte Rate beschreibt, welche Krebshäufigkeit vorliegen würde (auf 100.000 Personen bezogen), wenn die Altersstruktur der beobachteten Bevölkerung derjenigen der Standardbevölkerung (Standardbevölkerung Europa) entsprochen hätte. Die Standardisierung wurde anhand von 18 Fünfjahres-Altersklassen vorgenommen (0-4, ..., 85+).

Erwartete Fallzahl

Die erwartete Fallzahl gibt an, wie viele Krebsneuerkrankungsfälle in der SG Bothel zu erwarten sind, wenn für die einzelnen Altersgruppen der Wohnbevölkerung der SG Bothel die altersspezifischen Krebsneuerkrankungsraten der Vergleichsregion herangezogen werden. Die erwartete Fallzahl für die SG Bothel wird anhand der Vergleichsregion Bezirk Lüneburg berechnet.

SIR

Das standardisierte Inzidenz-Verhältnis (SIR für 'standardized incidence ratio') gibt den Quotienten zwischen der Anzahl beobachteter und erwarteter Krebsneuerkrankungsfälle an. Das SIR ist daher genau dann 1, wenn beobachtete und erwartete Neuerkrankungsfallzahlen übereinstimmen; ein erhöhtes SIR, z.B. von 1,10 besagt, dass in der Untersuchungsregion eine um 10% höhere Krebshäufigkeit vorliegt als erwartet. Ein SIR unter 1 weist auf eine vergleichsweise niedrigere Krebshäufigkeit für die jeweilige Diagnose hin als erwartet.

Der Zeitpunkt des Auftretens einer Erkrankung unterliegt einem Zufallsprozess. Die Aussage allein, dass das SIR ober- oder unterhalb des Wertes 1 liegt, hat keine Aussagekraft, da die beobachteten SIR im Normalfall um die 1 schwanken. Um daher beurteilen zu können, ob es zu statistisch auffälligen Abweichungen der beobachteten Anzahl von der erwarteten gekommen ist, muss eine Annahme zur statistischen Verteilung der Neuerkrankungen getroffen werden. Hierzu wird für die beobachtete Fallzahl in der Untersuchungsregion, die den Zähler des SIR bildet, eine Poissonverteilung angenommen. Mit dieser Annahme können Konfidenzintervalle für das SIR sowie Hypothesen- bzw. Signifikanztests für das SIR abgeleitet werden (vgl. Anhang).

Das beobachtete SIR kann auch als Schätzung für das „wahre SIR“ betrachtet werden, das sich auf einen längeren Zeitraum oder eine größere Population bezieht.

Ergänzend zum Punktschätzer des SIR werden Vertrauens- bzw. Konfidenzintervalle [KI] angegeben, die einen Wertebereich schätzen, der das ‚wahre‘ SIR mit vorgegebener Wahrscheinlichkeit $(1 - \alpha)$ überdeckt. Ein 95%-KI überdeckt somit das wahre SIR mit 95%-iger Wahrscheinlichkeit. Je schmaler ein KI ausfällt, desto präziser ist die Schätzung. Wenn das KI den Wert 1,00 nicht überdeckt, so kann die Hypothese, dass das wahre SIR 1 betrage, mit einer Irrtumswahrscheinlichkeit von α verworfen werden. In diesem Bericht werden ausschließlich zweiseitige 95%-KI angegeben.

Der bei einem SIR angegebene p-Wert gibt die Wahrscheinlichkeit an, dass der geschätzte oder ein noch größerer Wert für das SIR beobachtet wird, obwohl das „wahre“ SIR 1,00 beträgt. Ist der p-Wert kleiner als das für den Einzelvergleich vorgegebene Signifikanzniveau, ist die Hypothese, dass die Inzidenzrate nicht erhöht bzw. das wahre SIR nicht erhöht ist, zu verwerfen.

3.5 Berücksichtigung des multiplen Testens

Der p-Wert (vgl. Kap. 3.4 und Anhang) gilt streng genommen nur für ein hypothesengeleitetes einmaliges Testen bzw. für einen einmaligen Vergleich. Werden beispielsweise zehn unabhängige Tests zu einer jeweiligen Irrtumswahrscheinlichkeit von 0,05 durchgeführt, beträgt die Wahrscheinlichkeit, dass man mindestens eine Hypothese irrtümlich

verwirft, bereits $1 - 0,95^{10} = 0,40$ (entspricht einer Gesamttirrtumswahrscheinlichkeit von 40%).

Werden zahlreiche Tests parallel durchgeführt, wie bei dieser Untersuchung, können sich allein auf Grund der Vielzahl der durchgeführten Vergleiche zahlreiche (zu einer vorgegebenen nominellen Irrtumswahrscheinlichkeit pro Einzelvergleich) „signifikante Ergebnisse“ ergeben. Für diese sollte aber streng genommen allenfalls von „statistisch auffälligen Ergebnissen“ gesprochen werden. Für Aussagen zur Signifikanz ist die Gesamttirrtumswahrscheinlichkeit zu berücksichtigen, die angibt, mit welcher Wahrscheinlichkeit über die Gesamtheit aller durchgeführten Vergleiche mindestens eine Hypothese irrtümlich abgelehnt wird. Soll das Niveau der Fehlerrate für alle durchgeführten Tests gelten, muss dieses für die Signifikanzbestimmung berücksichtigt werden.

Das hierfür verwendete Verfahren ist die Prozedur nach Bonferroni-Holm [3, 4]. Hierbei werden alle p-Werte der Größe nach sortiert und mit wachsenden Schwellenwerten verglichen. Die Prozedur beginnt mit der Untersuchungsgruppe mit dem geringsten p-Wert im Einzelvergleich. Ist dieser p-Wert größer als die gewählte Gesamttirrtumswahrscheinlichkeit dividiert durch die Anzahl der durchgeführten Einzelvergleiche, dann wird die Hypothese beibehalten und es werden keine weiteren Einzelvergleiche durchgeführt. Im anderen Fall wird die Hypothese verworfen und die Untersuchungsgruppe mit dem nächst höheren p-Wert mit der Gesamttirrtumswahrscheinlichkeit dividiert durch die Anzahl der durchgeführten Tests minus eins überprüft, um zu beurteilen, ob auch beim zweiten Einzelvergleich von signifikanten Ergebnissen zu sprechen ist. Dieses Verfahren wird bis zu dem p-Wert fortgeführt, für den sich keine Signifikanz mehr einstellt (Vergleichsschwelle: Gesamttirrtumswahrscheinlichkeit dividiert durch Gesamtanzahl minus bislang signifikante Einzelvergleiche). Sobald ein nicht-signifikantes Einzelergebnis vorliegt, wird das Verfahren abgebrochen; die weiteren Einzelvergleiche sind ebenfalls nicht-signifikant.

Zur Beantwortung der primären Fragestellung „Werden in der SG Bothel mehr Krebserkrankungen beobachtet als zu erwarten sind?“ werden zwölf Diagnosegruppen untersucht, davon drei Diagnosegruppen für Frauen und Männer getrennt. Das entspricht 15 Untersuchungsgruppen. Die Hypothese der Eingangsfragestellung setzt sich somit aus 15 einzelnen Hypothesen zusammen. Zur Einhaltung einer Gesamttirrtumswahrscheinlichkeit von 5% ist daher ein Testergebnis mit dem geringsten p-Wert $< 0,05/15=0,0033$ als „signifikant“ einzustufen. Bei Ablehnung der Hypothese würde der Test mit dem zweit niedrigsten p-Wert $< 0,05/14=0,0036$ als „signifikant“ eingestuft usw.

Mit diesem schließenden („induktiven“) statistischen Vorgehen können nur Zufallseffekte kontrolliert werden. Sofern systematische Fehler wie Untererfassung oder eine unzureichende Validität der Meldungsangaben vorliegen könnten, sind „statistisch auffällige“ Abweichungen vom Erwartungswert umso vorsichtiger zu bewerten. Statistisch zuverlässige Aussagen sind bei wiederholt beobachteten Auffälligkeiten nur durch hypothesengeleitete weiterführende Studien zu erhalten.

3.6 Ergänzende Statistiken außerhalb der Hauptfragestellung

Da die angegebenen Schätzungen des SIR für die 15 Untersuchungsgruppen sowie ergänzend für „Krebs insgesamt“ nur einen einzigen Wert („Punktschätzer“) darstellen, der alleine keine Aussage zum Streuungsverhalten dieses Schätzers erlaubt, hat sich die ergänzende Darstellung von Konfidenzintervallen als Bereichsschätzer (vgl. Kap 3.4) bei der Auswertung von Krebsregisterdaten etabliert.

Um die Vergleichbarkeit mit anderen Berichten (z. B. Krebs in Niedersachsen 2006/2007, Sonderauswertungen des EKN und anderer epidemiologischer Krebsregister) zu erleichtern, wurden - zusätzlich zu dem beschriebenen induktiven Vorgehen - für eine ergänzende Beschreibung der Ergebnisse auch zweiseitige 95%-Konfidenzintervalle angegeben. Auch wenn Konfidenzintervalle grundsätzlich der schließenden Statistik zuzurechnen sind, dürfen sie in diesem Kontext nur rein deskriptiv verstanden werden, da sie die Gesamtzahl der durchgeführten Vergleiche nicht kontrollieren. Deshalb sollten die in der Auswertung zu den Untersuchungsgruppen dargestellten einzelnen Konfidenzintervalle zurückhaltend interpretiert werden.

4 Ergebnisse

Im Folgenden werden die im EKN erfassten Krebsneuerkrankungsfälle für die fünfzehn Untersuchungsgruppen dargestellt und statistisch überprüft.

4.1 Beschreibung: Häufigkeit von Krebsneuerkrankungen

An invasiven Krebserkrankungen (C00-C97 ohne nicht-melanotischen Hautkrebs) erkrankten in den Diagnosejahren 2003 - 2012 in der SG Bothel 533 Personen bei 494,0 erwarteten Fällen. Das bedeutet eine Erhöhung gegenüber der Vergleichsregion von 8% (SIR 1,08). Bei Männer wurden 302 Erkrankungen beobachtet (erwartet 267,3), bei Frauen 231 Erkrankungen (erwartet 226,7).

In Tabelle 4 sind die Ergebnisse für die 15 Untersuchungsgruppen zusammengefasst. Die drei häufigsten bösartigen Neubildungen, Krebs der Verdauungsorgane (C15-C26 ohne C22), Krebserkrankungen von Kehlkopf und Lunge (C32-C34) und Leukämien und Lymphome (C81-C96), wurden entsprechend der Anfrage für Männer und Frauen getrennt ausgewertet.

Für neun Untersuchungsgruppen wurden mehr Neuerkrankungen beobachtet als erwartet (SIR >1), für sechs Untersuchungsgruppen weniger Fälle als erwartet (SIR <1). Mehr Fälle als erwartet wurden beobachtet für bösartige Neubildungen der Verdauungsorgane, Kehlkopf und Lunge bei Männern, Knochen, Haut und Weichteilgeweben, Brustdrüse bei Frauen, Prostata und Hoden, Hirntumoren und endokrine Drüsen sowie bei Leukämien und Lymphomen bei Männern. Weniger Fälle beobachtet als erwartet wurden für bösartige Tumoren von Mund und Rachen, der Leber, an Kehlkopf und Lunge bei Frauen, der weiblichen Genitalorgane, Nieren und Harnorgane sowie bei Leukämien und Lymphomen bei Frauen.

Tabelle 4: Häufigkeit von Krebsneuerkrankungsfällen (inklusive DCO-Fällen) in der SG Bothel für 15 Untersuchungsgruppen in den Diagnosejahren 2003 - 2012 (Vergleichsregion Bezirk Lüneburg, EKN-Stand 02.07.2014)

SG Bothel - Inzidenz 2003-2012			Neuerkrankungen beobachtet	Neuerkrankungen erwartet	SIR (beobachtet / erwartet)	95% Konfidenzintervall des SIR (deskriptiv, zweiseitig)	p-Wert (einseitig)
Krebslokalisation bzw. -diagnose	ICD-10-Code	Geschlecht					
Mund + Rachen	C00-C14	m + w	6	12,1	0,49	0,18 - 1,08	0,9813
Verdauungsorgane außer Leber	C15-C26 (o. C22)	m	66	61,0	1,08	0,84 - 1,38	0,2783
		w	52	49,7	1,05	0,78 - 1,37	0,3916
Leber	C22	m + w	2	5,2	0,38	0,04 - 1,38	0,9667
Kehlkopf, Lunge	C32-C34	m	43	39,3	1,10	0,79 - 1,48	0,2957
		w	10	16,0	0,63	0,30 - 1,15	0,9561
Knochen, Haut, Weichteilgewebe u.ä.	C40-C49 (o. C44)	m + w	32	27,5	1,16	0,80 - 1,64	0,2183
Brustdrüse	C50	w	85	75,2	1,13	0,90 - 1,40	0,1411
Weibliche Genitalorgane	C51-C58	w	28	28,8	0,97	0,65 - 1,40	0,5863
Prostata, Hoden	C61-C62	m	95	82,9	1,15	0,93 - 1,40	0,1036
Niere, Harnorgane	C64-C68	m + w	29	31,5	0,92	0,62 - 1,32	0,6948
Hirntumoren	C70-C72	m + w	7	6,9	1,02	0,41 - 2,09	0,5345
Endokrine Drüsen	C73-C75	m + w	9	4,5	2,02	0,92 - 3,83	0,0385
Leukämien + Lymphome	C81-C96	m	41	21,3	1,93	1,38 - 2,61	0,0001
		w	15	16,8	0,89	0,50 - 1,47	0,7010

m = Männer, w = Frauen, m+w = Männer und Frauen gesamt

Zufallsschwankungen um den Erwartungswert sind zu erwarten. Bei 15 durchgeführten Vergleichen ist damit zu rechnen, dass in etwa eine Untersuchungsgruppe zufällig außerhalb des üblichen 95%-Konfidenzintervalls (95%-KI) liegt. Bei 14 der 15 durchgeführten Untersuchungen lag das Verhältnis der beobachteten zu den erwarteten Fällen innerhalb der Schwankungsbreite des zweiseitigen 95%-KI.

4.2 Statistische Bewertung, ob eine Erhöhung der Krebsneuerkrankungen vorliegt (statistische Induktion)

In dem untersuchten 10-Jahreszeitraum wurden in der Gruppe der Leukämien und Lymphome bei Männern 41 Erkrankungen beobachtet bei 21,3 erwarteten Fällen (SIR 1,93 [95%-KI 1,38-2,61]; $p=0,0001$).

Für Aussagen zur Signifikanz wird - wie in Kapitel 3.5 ausgeführt - die Gesamttirrtumswahrscheinlichkeit kontrolliert. Der p-Wert liegt mit 0,0001 unter der nach Bonferroni-Holm berechneten Signifikanzschwelle von 0,0033. Die Nullhypothese muss daher verworfen werden, d.h. die beobachtete Anzahl an Leukämien und Lymphomen bei Männern im Zeitraum 2003 - 2012 in der SG Bothel ist zur Gesamttirrtumswahrscheinlichkeit von 5% statistisch signifikant erhöht. Damit ist die Hypothese, es liege keine Erhöhung der Krebsneuerkrankungen in der SG Bothel vor, insgesamt zu verwerfen.

Der Einzelvergleich mit dem zweitniedrigsten p-Wert ist der der endokrinen Drüsen. Hier liegt der p-Wert mit 0,0385 jedoch deutlich über der Signifikanzschwelle für den zweiten Vergleich von 0,0036 (vgl. Kap. 3.5). Insofern liegt keine weitere Untersuchungsgruppe vor, die statistisch auffällig vom Erwartungswert abweicht.

5 Ergänzende Analysen: Leukämien und Lymphome bei Männern

Die folgenden Auswertungen gehen über die ursprüngliche Anfrage hinaus. Aufgrund des Ergebnisses einer signifikanten Erhöhung für die Untersuchungsgruppe der Leukämien und Lymphome (ICD-10: C81-C96) bei Männern wurde deskriptiv eine Betrachtung von Einzeldiagnosen bzw. Subgruppen, der zeitlichen Entwicklung und der Altersverteilung vorgenommen, um gegebenenfalls Hinweise zu erhalten, wie eine weitere Abklärung erfolgen könnte.

5.1 Einzeldiagnosen

Tabelle 5 zeigt die beobachteten und erwarteten Neuerkrankungsfälle und die absolute Differenz zwischen beiden Fallzahlen für die Diagnosejahre 2003 - 2012. Die Tabelle ist absteigend sortiert nach der diagnosespezifischen Differenz, d. h. an erster Stelle steht die Diagnose, bei der die Abweichung zwischen beobachteten und erwarteten Fällen am höchsten ist.

Die größte Differenz zeigt sich für das Multiple Myelom (C90) mit 8,5 zusätzlichen Fällen (12 beobachtete und 3,5 erwartete Neuerkrankungsfälle). Non-Hodgkin-Lymphome (C82-C85) folgen an zweiter Stelle mit 7,2 zusätzlichen Fällen (16 beobachtete und 8,8 erwartete Fälle). Für Leukämien (C91-C95) beträgt die Differenz 3,6 zusätzliche Fälle

(11 beobachtete und 7,4 erwartete Fälle). Für die weiteren Diagnosen liegt die Differenz unter 1.

Tabelle 5: Subgruppenanalyse für Leukämien und Lymphome (C81-C96) bei Männern - beobachtete und erwartete Neuerkrankungsfälle sowie die Differenz nach Einzeldiagnosen (Vergleichsregion Bezirk Lüneburg, Diagnosejahre 2003 - 2012)

SG Bothel		Neuerkrankungen beobachtet (a)	Neuerkrankungen erwartet (b)	Übererkrankung (a-b)
Krebsdiagnose	ICD-10-Code			
Multiples Myelom	C90	12	3,5	8,5
Non-Hodgkin-Lymphome	C82 - C85	16	8,8	7,2
Leukämien	C91 - C95	11	7,4	3,6
Hodgkin-Lymphome	C81	2	1,4	0,6
Bösartige immunprolif. Krankheiten	C88	0	0,1	-0,1
Sonstige Erkrankungen des lymphat., blutbild. Gewebes	C96	0	0,1	-0,1
Summe	C81-C96	41	21,3	19,7

5.2 Zeitliche Entwicklung der Neuerkrankungshäufigkeit

Über den 10-Jahres-Zeitraum betrachtet liegt die Neuerkrankungsfallzahl für Leukämien und Lymphome in der SG Bothel bei Männern bei durchschnittlich 4,1 Fällen im Jahr. Es ist kein einheitlicher Anstieg oder Rückgang der beobachteten Fallzahlen über die Zeit zu erkennen. Der Erwartungswert nimmt über die Zeit von 1,8 auf 2,3 Fälle pro Jahr leicht zu. Aus der Tabelle 6 geht hervor, wie sich die in der SG Bothel beobachteten 41 Leukämie- und Lymphomerkrankungen bei Männern über die Zeit verteilen. Aufgrund der geringen Fallzahlen in den einzelnen Diagnosejahren werden die Fallzahlen aus Datenschutzgründen in Zweijahresperioden angegeben.

Tabelle 6: Leukämien und Lymphome (C81-C96) bei Männern - beobachtete und erwartete Neuerkrankungsfälle in der SG Bothel (Vergleichsregion Bezirk Lüneburg)

SG Bothel	Neuerkrankungen beobachtet	Neuerkrankungen erwartet
Diagnosejahre		
2003-2004	8	3,9
2005-2006	9	4,2
2007-2008	11	4,3
2009-2010	9	4,5
2011-2012	4	4,5
2003 - 2012	41	21,3

5.3 Altersverteilung der an Leukämien und Lymphomen neu erkrankten Männer

Die Altersverteilung der im Zeitraum 2003 - 2012 neu an Leukämien und Lymphomen erkrankten Männer aus der SG Bothel wird in Tabelle 7 dargestellt. Ergänzend sind die

altersspezifischen Raten für die SG Bothel je 100.000 Männer ausgewiesen. Angegeben ist jeweils das dem EKN gemeldete Alter der Erkrankten zum Zeitpunkt der Erstdiagnose.

Tabelle 7: Leukämien und Lymphome (C81-C96) bei Männern - Altersverteilung der beobachteten Neuerkrankungsfälle und der altersspezifischen Raten (Fälle je 100.000 Männer) in der SG Bothel, Diagnosejahre 2003 - 2012

SG Bothel	Neuerkrankungen beobachtet	altersspez. Rate pro 100.000
0-14	3	40,8
15-44	4	23,0
45-59	3	32,5
60-74	26	377,4
75+	5	211,8
Gesamt	41	94,7

Die Abbildung 1 zeigt die altersspezifischen Raten für die SG Bothel im Vergleich zu den Raten des Bezirks Lüneburg, wobei zu beachten ist, dass die Raten der SG Bothel z.T. auf sehr geringen Fallzahlen basieren. Für die Altersklassen 0-14, 15-44 und 60-74 Jahre liegen die Raten in der SG Bothel über denen der Vergleichsregion.

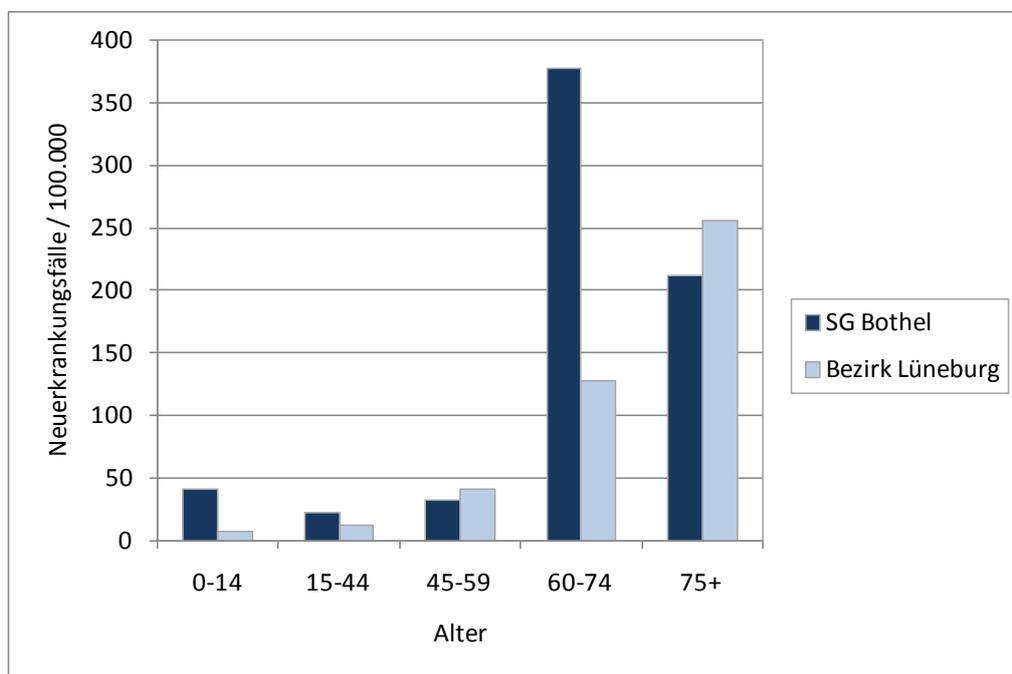


Abbildung 1: Leukämien und Lymphome (C81-C96) Männer – Altersspezifische Neuerkrankungsraten in der SG Bothel im Vergleich zum Bezirk Lüneburg für den Diagnosezeitraum 2003 - 2012

Die altersstandardisierte Neuerkrankungsrate (Standardbevölkerung Europa) für Leukämien und Lymphome bei Männern liegt für den Zeitraum 2003 - 2012 in der SG Bothel bei 77,9 je 100.000 Einwohner. Für die Vergleichsregion Bezirk Lüneburg liegt die alters-

standardisierte Rate bei 40,0. Das Saarland weist für 2003 - 2011 eine altersstandardisierte Rate von 37,1² aus und deutschlandweit gibt das Robert-Koch-Institut für 2003 - 2010 eine Rate von 36,4³ an. Somit ist die altersstandardisierte Rate für Bothel deutlich erhöht.

6 Diskussion

Mit Datenstand 02.07.2014 wurde eine Auswertung der dem EKN gemeldeten Krebsneuerkrankungsfälle für die SG Bothel vorgenommen. Die Ergebnisse wurden mit denen des ehemaligen Regierungsbezirks Lüneburg verglichen.

Für die Auswertung der Krebsneuerkrankungen standen die Diagnosejahre 2003 - 2012 zur Verfügung, der Erfassungsgrad von Krebsneuerkrankungen liegt im Bezirk Lüneburg für den 10-Jahreszeitraum bei über 95%. Aufgrund von nachfolgend eingehenden Meldungen kann es für die einzelnen Diagnosejahre noch zu Veränderungen der Fallzahlen kommen. Für die ausgewerteten Fälle ist eine Verschiebung des Diagnosejahres auf einen Zeitpunkt außerhalb des Diagnosezeitraums 2003 - 2012 jedoch eher unwahrscheinlich.

Ein Selektionsfehler aufgrund eines unterschiedlichen regionalen Meldeverhaltens ist wegen der in Niedersachsen bestehenden Meldepflicht für Pathologien nicht wahrscheinlich. Systematische Meldedefizite sind daher nicht anzunehmen.

In der Anfrage des LK Rotenburg auf Auswertung von 15 Untersuchungsgruppen erfolgte nur eine geringe Einschränkung auf bestimmte Diagnosen(-gruppen). Die angefragten 12 Diagnosegruppen decken etwa 98 % aller beobachteten Krebsfälle ab. Die mit dem multiplen Testen einhergehende Problematik der erhöhten Irrtumswahrscheinlichkeit und der damit rein zufällig auftretenden Abweichungen vom Erwartungswert wurde in den Ausführungen zur Methodik (siehe Punkt 3.5) ausführlich beschrieben. Für die Beurteilung der Signifikanz der beobachteten Rate an Leukämien und Lymphomen bei Männern wurde die Prozedur nach Bonferroni-Holm angewandt; trotz der mit der Anfrage verbundenen „multiplen Vergleiche“ mit 15 Untersuchungsgruppen ist das Ergebnis im Sinne einer Gesamtirrtumswahrscheinlichkeit von 5% als statistisch signifikant einzustufen.

Die Auswertung wurde um die deskriptive Subgruppenanalyse der signifikant erhöhten Untersuchungsgruppe Leukämien und Lymphome für Männer ergänzt. Dabei zeigt sich, dass insbesondere bei Multiplen Myelomen und Non-Hodgkin-Lymphomen mehr Fälle beobachtet wurden, als zu erwarten waren. Ein zeitlicher Trend ist nicht erkennbar.

Für das EKN ist das Diagnosejahr 2003 das früheste Jahr mit ausreichend vollzähligem Datenbestand, so dass keine validen Aussagen zur Inzidenz aus früheren Jahren gemacht werden können. Da jedoch auch in der Altersgruppe 0-14 Jahren mehr Fälle beobachtet worden sind, als zu erwarten waren, könnte eine Anfrage an das Kinderkrebsregister in Mainz klären, ob in deren Datenbestand bereits in früheren Jahren gehäuft

² www.krebsregister.saarland.de, Datenbankabfrage vom 21.08.2014

³ www.krebsdaten.de, Datenbankabfrage vom 21.07.2014

kindliche Krebserkrankungen zu beobachten waren. Das EKN ist gerne bereit, eine entsprechende Anfrage zu koordinieren.

Untersuchungen anhand von Routine-Krebsregisterdaten sind alleine nicht in der Lage, kausale Zusammenhänge zu einzelnen Risikofaktoren zu beweisen. Dem EKN liegen jedoch routinemäßig keine Angaben über konkrete regionale Risikofaktoren oder Umweltexpositionen vor, die in weitergehende Untersuchungen eingebracht werden könnten. Daher wird in diesem Bericht nicht weiter auf bekannte bzw. diskutierte Risikofaktoren der verschiedenen Diagnosen eingegangen.

Auffällig ist, dass nur bei Männern erhöhte Erkrankungsraten bei Leukämien und Lymphomen zu sehen sind, während bei Frauen sogar etwas weniger Fälle beobachtet wurden, als zu erwarten waren. Dies könnte ein Hinweis darauf sein, dass berufliche Expositionen eine Rolle spielen könnten. 34% der im EKN erfassten Leukämie- und Lymphomfälle bei Männern beruhen ausschließlich auf Pathologenmeldungen, für die im EKN nach der alten Fassung des GEKN keine dechiffrierbaren Personendaten für eine Kontaktaufnahme vorliegen.

Weiter zu diskutieren ist, ob und wie Folgeuntersuchungen die erstmaligen Auswertungen des EKN sinnvoll ergänzen können:

- Es sollte geklärt werden, ob die Möglichkeit besteht, nähere Informationen zu möglichen arbeits-, umweltbezogenen oder auch individuellen Risikofaktoren der Erkrankten zu erhalten.
- Das EKN wird die Daten ab dem Diagnosejahr 2013 weiter beobachten und dabei gezielt die Diagnosen Multiples Myelom und Non-Hodgkin-Lymphome betrachten, um zu klären, ob die Erhöhung für diese Diagnosen anhält.

7 Zusammenfassung

Anlass dieser Sonderauswertung ist eine Anfrage des Landkreises Rotenburg zur Häufigkeit von Krebserkrankungen in der Samtgemeinde (SG) Bothel. Im Zusammenhang mit der dort durchgeführten Erdgasförderung wird von der Bevölkerung eine lokale Krebshäufung vermutet. In der Anfrage des Landkreises wurden 15 Untersuchungsgruppen aus 12 Diagnosegruppen definiert, die insgesamt 98% aller Krebsfälle enthielten. Bei der Durchführung vieler paralleler Untersuchungen ist für Aussagen zur Signifikanz der Ergebnisse die Gesamtirrtumswahrscheinlichkeit zu berücksichtigen

In den Diagnosejahren 2003 - 2012 erkrankten in der SG Bothel 533 Personen neu an invasiven Krebserkrankungen bei 494 erwarteten Fällen. Das bedeutet eine Erhöhung gegenüber der Vergleichsregion (Bezirk Lüneburg) von 8% (SIR 1,08).

In dem untersuchten 10-Jahreszeitraum wurden für die Gruppe der Leukämien und Lymphome bei Männern 41 Erkrankungen beobachtet bei 21,3 erwarteten Fällen (SIR 1,93) Der berechnete p-Wert für Leukämien und Lymphome bei Männern liegt mit 0,0001 deutlich unter der nach Bonferroni-Holm korrigierten Signifikanzschwelle von 0,0033, so dass die beobachtete Erhöhung als signifikant einzustufen ist. Somit ist bei

einer Gesamtirrtumswahrscheinlichkeit von 5% die Krebsneuerkrankungsrate für Leukämien und Lymphome bei Männern in der SG Bothel erhöht.

Für Frauen ist mit 15 beobachteten Leukämie- und Lymphom-Erkrankungen bei 16,8 erwarteten Fällen keine Auffälligkeit zu erkennen.

Ebenso ist keine weitere Untersuchungsgruppe statistisch auffällig erhöht.

Die Untersuchung des EKN bestätigt damit die Vermutung einer Krebshäufung in der SG Bothel. Diese betrifft Leukämien und Lymphome bei Männern.

Aussagen zur Ursache von lokalen Krebshäufungen sind mit Analysen, die sich ausschließlich auf Krebsregister-Routinedaten beziehen, nicht möglich.

8 Literatur

- (1) Böhning, D: Allgemeine Epidemiologie. Oldenbourg, München, 1998
- (2) Epidemiologisches Krebsregister Niedersachsen, Registerstelle (Hrsg.): Krebs in Niedersachsen 2011, Jahresbericht, Oldenburg, 2014
- (3) Horn M, Vollandt R: Multiple Tests und Auswahlverfahren, Stuttgart, Jena, New-York, 1995
- (4) Victor A, Elsässer A, Hommel G, Blettner M: Wie bewertet man die p-Wert-Flut? Dtsch Arztebl Int 2010; 107(4): 50–6 DOI: 10.3238/arztebl.2010.0050

9 Anhang

Schließende Statistik

Bei einem statistischen Test oder Hypothesentest wird eine im Vorfeld formulierte Hypothese anhand der vorliegenden Daten überprüft, um eine Aussage über ihre Gültigkeit treffen zu können. Da die Daten Zufallsprozessen unterliegen, wird die Aussage zur Gültigkeit einer Hypothese nie hundertprozentig sicher sein, sondern gewissen Irrtumswahrscheinlichkeiten unterliegen. Ein statistischer Test kann insofern als Entscheidungsregel unter Kontrolle von Unsicherheiten verstanden werden.

Aufstellung einer statistischen Hypothese und Prinzip des indirekten Beweises

Grundsätzlich ist der klassische Hypothesentest eine Form eines „indirekten Beweises“: Eine aufgestellte Hypothese wird so lange beibehalten, wie es die beobachteten Daten zulassen. Wenn hingegen die Beobachtungen unter der Hypothese zu unwahrscheinlich sind, erscheint die Annahme gerechtfertigt, dass die aufgestellte Hypothese nicht stimmt und sie daher verworfen werden kann. Dieses indirekte Vorgehen führt dazu, dass man die (statistische) Hypothese bzw. sogenannte Nullhypothese entsprechend scheinbar umgekehrt formulieren muss:

- Möchte man statistisch nachweisen, dass ein neues Medikament besser ist als das alte, formuliert man die Nullhypothese, das neue Medikament habe höchstens so gute Erfolgsaussichten wie das alte. Die Alternativhypothese besagt, dass das neue Medikament besser sei. Zeigt sich nach der Prüfung anhand der Daten, dass das neue Medikament deutlich bessere Wirkungen erzielt hat, „muss“ die Nullhypothese zugunsten der Alternativhypothese verworfen werden. Dies führt dann zu dem (beabsichtigten) Nachweis, dass das neue Medikament besser ist.
- Um zu prüfen, ob in einer Untersuchungsregion die Wahrscheinlichkeit einer bestimmten Krebserkrankung höher ist als normalerweise, stellt man die Hypothese auf, dass diese Wahrscheinlichkeit höchstens genauso hoch ist. Liegen dann derartig viele Fallmeldungen vor, dass die Hypothese zu verwerfen ist, wird der Schluss gezogen, dass die Wahrscheinlichkeit für diese Krebserkrankung erhöht ist.

Die nachzuweisende Aussage tritt somit in der Alternativhypothese auf.

Irrtumswahrscheinlichkeiten

Die Irrtumswahrscheinlichkeit, die Nullhypothese auf Grund der vorliegenden Daten zu verwerfen, obwohl sie an sich richtig ist, wird als „Fehler 1. Art“ oder auch „ α -Fehler“ bezeichnet. Dieser Fehler wird beim klassischen Hypothesen- bzw. Signifikanztest kontrolliert. Dazu wird vor der Datenauswertung ein sogenanntes Signifikanzniveau des Tests vorgegeben, häufig 1% oder 5%. Das Signifikanzniveau ist der maximal zulässige Wert für den Fehler 1. Art. D.h. die Wahrscheinlichkeit, die Hypothese zu verwerfen, obwohl sie richtig ist, entspricht für den Test höchstens dem Signifikanzniveau.

Daneben gibt es noch den „Fehler 2. Art“ bzw. den „ β -Fehler“. Er gibt entsprechend an, mit welcher Wahrscheinlichkeit die Nullhypothese beibehalten wird, obwohl die Alternativhypothese korrekt ist. Die Wahrscheinlichkeit, die Nullhypothese abzulehnen, wenn sie falsch ist, wird demgegenüber als Güte oder „Power“ des Tests beschrieben.

Über den β -Fehler bzw. die Wahrscheinlichkeit, die Nullhypothese korrekterweise abzulehnen ($1 - \beta$), können Aussagen zur Güte des Tests getroffen werden:

- So kann für eine Untersuchungsregion abgeschätzt werden, mit welcher Wahrscheinlichkeit die Nullhypothese, es läge keine Erhöhung vor, abgelehnt wird, wenn es in Wirklichkeit eine relevante Erhöhung gäbe. Diese Wahrscheinlichkeit fällt umso höher aus, je größer der tatsächliche prozentuale Effekt ist. Aber auch die Größe der Untersuchungsregion spielt eine Rolle: In einer größeren Gemeinde ist bei derselben Effektstärke das Verwerfen der Nullhypothese wahrscheinlicher als bei einer kleineren Gemeinde.

Teststatistik

Eine Teststatistik muss in der Lage sein, für die aufgeworfene Fragestellung möglichst gut zwischen Null- und Alternativhypothesen zu unterscheiden und dabei möglichst alle verfügbaren Informationen zu nutzen:

- Man möchte anhand von Stichproben prüfen, ob die Bewohner einer Stadt A größer sind als die der Stadt B. Hier könnte man jeweils nur die ersten Personen der beiden Stichproben anhand ihrer Körpergrößen vergleichen. Dabei würden aber die übrigen Stichprobeninformationen verschwendet, d.h. die Körpergrößen der übrigen ebenfalls in den Stichproben erhobenen Bewohner blieben bei einem derartigen Test unberücksichtigt. Sinnvoller ist es, die durchschnittliche Körpergröße beider Stichproben miteinander zu vergleichen und darüber ein statistisches Testverfahren abzuleiten.

Für ein Testproblem können somit konkurrierende Teststatistiken bestehen, die anhand verschiedener Gütekriterien verglichen werden können. So sind viele der standardmäßig eingesetzten Testverfahren in dem Sinne optimal, dass es unter den getroffenen statistischen Annahmen keine Tests mit einer höheren Power gibt.

Die Entscheidungsregel und der Begriff der Signifikanz:

Für die Teststatistik wird ein kritischer Wert⁴ berechnet, so dass die Wahrscheinlichkeit dafür, dass die Teststatistik größer als dieser kritische Wert ist, maximal der Irrtumswahrscheinlichkeit entspricht. Überschreitet die Teststatistik diesen kritischen Wert, so kann die Nullhypothese verworfen werden. Das Ergebnis ist signifikant.

Werte unterhalb dieses kritischen Wertes werden auch als Annahmehereich des Tests bezeichnet, da mit ihnen die Nullhypothese nicht verworfen wird. Das Ergebnis des Tests ist nicht signifikant.

Ein signifikanter Test gibt nur an, dass der aus den konkreten Daten resultierende Wert des Tests unter der Nullhypothese unwahrscheinlich ist und zwar höchstens so wahrscheinlich wie die vorgegebene Irrtumswahrscheinlichkeit. Eine Aussage zur Wahrschein-

⁴ Dies bezieht sich auf einseitig formulierte Testprobleme, bei denen allein hohe Werte der Teststatistik gegen die Nullhypothese sprechen. Es gibt aber auch zweiseitig formulierte Testprobleme, bei denen sowohl sehr kleine wie große Werte gegen die Nullhypothese sprechen. So werden Abweichungen von Normwerten, die bspw. in der Produktion eingehalten werden müssen, meist zweiseitig überprüft. D.h. Abweichungen in beide Richtungen führen zur Ablehnung der Hypothese, 'die Normwerte werden eingehalten', und somit zur Neu-Adjustierung der Maschine.

lichkeit von Null- oder Alternativhypothese wird damit nicht getroffen. Entweder wird anhand der Daten die Nullhypothese verworfen oder aber sie muss beibehalten werden. Die Nullhypothese wird damit aber keineswegs „bestätigt“ oder gar „bewiesen“.

Ein „signifikantes Ergebnis“ kann auch zustande kommen, wenn unzutreffende Annahmen über den, den Daten zu Grunde liegenden Zufallsprozess gemacht wurden. Auch können systematische Erhebungsfehler, die die Daten beeinflussen könnten, nicht kontrolliert werden. Eine Testentscheidung hängt somit ab von den Hypothesen, der gewählten Teststatistik, dem vorgegebenen Signifikanzniveau, dem beobachteten Effekt sowie dem gewählten Stichprobenumfang.

p-Wert – „Signifikanztest“ als Alternative zum „Hypothesentest“

Das Konzept des p-Wertes bzw. des „Signifikanztests“ stellt ein induktives Verfahren dar, von der Stichprobe auf die Grundgesamtheit zu schließen. Es wurde parallel zum Konzept des Hypothesentests entwickelt. Diese beiden Konzepte werden allerdings in den allermeisten Statistikeinführungen nicht klar voneinander abgegrenzt, sondern vermischt und als einheitliches Konzept dargestellt.

Dabei benötigt jedoch der Signifikanztest keine ausformulierte Alternativhypothese, sondern beschränkt sich auf Aussagen zur Hypothese. Der p-Wert gibt dabei die Wahrscheinlichkeit der beobachteten oder extremeren Daten unter der Hypothese an. Je kleiner der p-Wert ist, desto mehr Evidenz spricht gegen die aufgestellte Hypothese. Ein p-Wert von 0,001 scheint mehr Sicherheit in die Entscheidung, eine Hypothese zum Signifikanzniveau von 5 % zu verwerfen, zu vermitteln als ein deutlich größerer p-Wert von 0,049, der aber zur selben Entscheidung führt.

Der enge Zusammenhang zum Hypothesentest besteht darin, dass ein p-Wert unter der vorgegebenen Irrtumswahrscheinlichkeit (1. Art) zur Ablehnung der Hypothese führt. Somit kann anhand des p-Wertes bewertet werden, ob ein entsprechend formuliertes Testproblem zu einem signifikanten oder nicht-signifikanten Ergebnis führt.

Der p-Wert darf aber keinesfalls, wie häufig fälschlich angegeben, als Irrtumswahrscheinlichkeit für die getroffene Entscheidung oder gar als die Wahrscheinlichkeit der Hypothese selbst interpretiert werden: Es wird schließlich die Wahrscheinlichkeit von Beobachtungen unter einer Hypothese, nicht jedoch die Wahrscheinlichkeit einer Hypothese unter den Beobachtungen betrachtet.