



Bundesministerium
für Umwelt, Naturschutz
und Reaktorsicherheit

SCHRIFTENREIHE REAKTORSICHERHEIT UND STRAHLENSCHUTZ

ANWENDUNGSBEREICH VON EPIDEMIOLOGISCHEN STUDIEN MIT ZUSAMMENGEFASSTEN DATEN ZUR BESTIMMUNG VON RISIKOFAKTOREN

BMU - 2005-676



WIR STEuern UM AUF ERNEUERBARE ENERGIEN.

BMU – 2005-676

**„Anwendungsbereich von epidemiologischen Studien
mit zusammengefassten Daten zur Bestimmung von
Risikofaktoren“**

**P. Jacob, J.C. Kaiser, M. Blettner, H. Bertelsmann,
M. Kutschmann, I. Likhtarev, L. Kovgan, S. Vavilov,
M. Tronko, T. Bogdanova**

IMPRESSUM

Dieser Band enthält einen Abschlussbericht über ein vom Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (BMU) gefördertes Vorhaben. Verantwortlich für den Inhalt sind allein die Autoren. Das BMU übernimmt keine Gewähr für die Richtigkeit, die Genauigkeit und Vollständigkeit der Angaben sowie die Beachtung privater Rechte Dritter. Der Eigentümer behält sich alle Rechte an der weiteren Nutzung oder Vervielfältigung des Berichts vor.

Der Bericht gibt die Auffassung und Meinung des Auftragnehmers wieder und muss nicht mit der des BMU übereinstimmen.

Herausgeber:

Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit

Referat RS I 2

Postfach 12 06 29

53048 Bonn

ISSN 1612-6386

Erscheinungsjahr: 2005

Inhaltsverzeichnis

1	Einleitung: Methodische Probleme zur Bestimmung von Risikofaktoren aus ökologischen Studien	6
2	Grundlagen	7
2.1	Methodische Grundlagen	7
2.2	Literaturübersicht: Methodische Probleme ökologischer Studien	12
2.2.1	Ökologische Studien: Definition und Kritik	12
2.2.2	Aufbau des Review „Ökologische Studien“	14
2.2.3	Neuere ökologische Studien im Überblick.....	15
2.2.4	Mathematische Modelle in ökologischen Studien	17
2.2.5	Ein prominentes Beispiel für einen Bias in ökologischer Studien: Cohens ökologische Analyse zu Radon und Lungenkrebs	20
2.2.6	Methodische Aspekte in ökologischen Studien	23
2.2.7	Weitere Diskussionspunkte zu ökologische Studien.....	30
2.2.8	Vergleichende Analysen auf individueller und ökologischer Ebene	36
2.2.9	Zusammenfassung.....	42
2.3	Ökologische Studien zur Untersuchung des Krebsrisikos durch ionisierende Strahlung	43
2.3.1	Methode der Literatursuche	43
2.3.2	Strahlenquellen.....	44
2.3.3	Ausgewählte Studien.....	44
2.3.4	Zusammenfassung.....	62
3	Simulationen zu ökologischen Studien	63
3.1	Pure specification bias.....	63
3.2	Zweck der Untersuchung	65
3.3	Simulationskonstellationen	65
3.4	Datenauswertung.....	66
3.5	Ergebnisse	67
3.6	Fazit.....	68
4	Schilddrüsenkrebs nach dem Tschernobylunfall.....	70
4.1	Schilddrüsenkrebsrisiko durch ¹³¹ I Exposition während der Kindheit.....	70
4.1.1	Reihenuntersuchungen	72
4.2	Literaturübersicht: Epidemiologie und Morphologie von Karzinomen der Schilddrüse	72
4.2.1	Epidemiologie und Risikofaktoren für Schilddrüsenkrebs	72
4.2.2	Latente papilläre Karzinome der Schilddrüse	73
4.3	Datengrundlagen	73
4.3.1	Studiengebiet.....	74
4.3.2	Personen und Krebsfälle in der vorläufigen Dosisschätzung.....	75
4.3.3	Personen und Krebsfälle in der endgültigen Dosisschätzung	76
4.3.4	Abschätzung des Hintergrundrisikos	77
4.3.5	Statistische Eigenschaften der zeitintegrierten Schilddrüsenaktivitäten aus den Dosisschätzungen der ersten und zweiten Kategorie	79
4.3.6	Morphologische Analyse der Schilddrüsenkrebsfälle in den hoch kontaminierten Gebieten der Ukraine	80
5	Simulationen zu ökologischen Studien zum Schilddrüsenkrebs-risiko nach dem Tschernobylunfall.....	82
5.1	Screening.....	82
5.1.1	Überdiagnose.....	82

5.1.2	Methodenentwicklung mit realitätsnahen Daten.....	83
5.1.3	Szenarienrechnungen mit realitätsnahen Daten	85
5.2	Messfehler	87
5.2.1	Klassischer und Berkson-Fehler.....	88
5.2.2	Ökologischer Messfehler.....	89
5.3	Linear-quadratisches Risikomodell in ökologischen Studien	90
6	Zusammenfassung	92
6.1	Allgemeine Ergebnisse für ökologische Studien	92
6.2	Spezielle Ergebnisse für ökologische Studien zum Schilddrüsenkrebsrisiko nach dem Tschernobylunfall.....	93
6.3	Fazit.....	95
7	Literatur	96

Abbildungsverzeichnis

Abbildung 1: Zusammenfassung des relativen Risikos (RR) einer Metaanalyse für Radon in Wohnräumen, sowie gepoolte Analyse von Studien an Minenarbeitern, transformiert für den Dosisbereich <JHM-3, sowie Cohens ökologische Regression. (aus Beir IV 1999, Abb. 3-2, S.89)	21
Abbildung 2: Schematische Darstellung einer hypothetischen Studie, in der ein Ecologic Bias auftritt. Jedes Kästchen stellt ein hypothetisches Individuum dar, wobei Kästchen mit fettem Rahmen Individuen mit inzidenten Verkehrsunfällen repräsentieren. Die Ziffern in den Kästchen stellen das Jahreseinkommen der jeweiligen Individuen dar. Quelle: Szklo & Nieto (2000).	27
Abbildung 3: Suizidrate (Y, pro 10 ⁵ Personenjahre) in Abhängigkeit von dem Protestantenanteil in preußischen Provinzen, 1883-1890. Aus: Durkheim E. Suicide: a study in Sociology. New York, Free Press, 1951, nach Morgenstern 1997).	32
Abbildung 4: Studiengebiet mit den ukrainischen Oblasts Zhytomyr, Kyiv und Chernihiv (schraffiert) sowie der Stadt Kyiv	74
Abbildung 5: Arithmetische Dosismittelwerte für 743 Ortschaften als Funktion der Anzahl der Personen der Geburtskohorte 1968-85 im Jahre 1986 in der vorläufigen Dosisschätzung	76
Abbildung 6: Arithmetische Dosismittelwerte für 670 Ortschaften als Funktion der Anzahl der Personen der Geburtskohorte 1968-85 im Jahre 1986 in der endgültigen Dosisschätzung	77

Tabellenverzeichnis

Tabelle 1: Ableitungen zum Vergleich von Risiken, mit R_1 = Risiko für Subpopulation $X=x+1$ und R_0 = Risiko für Subpopulation mit $X=x$	12
Tabelle 2: Zusammenhang der oben beschriebenen Modelle	12
Tabelle 3: Medline-Recherche Stichwort „ecologic studies“, angewandte ökologische Studien 2001/2002	15
Tabelle 4 Mittelwerte aus 5000 Simulationen der Risiko-Schätzer und Standardabweichungen in individuellen und ökologischen Analysen. Der wahre Betawert betrug $2,937 \times 10^{-3}$. Die Ergebnisse basieren auf 21 Gruppen mit 10.000 Personen. (nach Sheppard und Prentice 1995, Tab.1, verändert)	34
Tabelle 5: Vergleich des individuell und ökologisch geschätzten zusätzlichen relativen Risikos (ERR) bei 100 Bqm-3 und 95%-Confidenzintervalle (95%-CI) für unterschiedliche Adjustierung für die Covariablen in der schwedischen Fall-Kontroll-Studie zu Radon und Lungenkrebs 1980-1984. (nach Lagarde und Pershagen 1999, verändert)	38
Tabelle 6: Relatives Risiko (RR) und 95%- Confidenzintervalle (95%-CI) für Lungenkrebs bei Radonkonzentration in Innenräumen von 100 Bqm ⁻³ vergleichen mit 0 Bqm ⁻³ bei ökologischer und individueller Regression unter Berücksichtigung unterschiedlicher Confounder. Daten der englischen Fall-Kontroll Studie für Radon in Innenräumen. (Darby et al. 2001, Tab. 2, verändert)	39
Tabelle 7: Ökologische Studien zu regionalen Clustern in der Umgebung von kerntechnischen Anlagen	46
Tabelle 8: Ökologische Studien in der Umgebung mehrerer kerntechnischer Anlagen (multi-site studies)	48
Tabelle 9: Ökologische Studien zu Krebs bei natürlicher Strahlung	53
Tabelle 10: Ökologische Studien zu Krebs nach nuklearem Fallout	55
Tabelle 11: Ökologische Studien zu Kinderleukämie nach Tschernobyl	57
Tabelle 12: Ökologische Studien zu Schilddrüsenkrebs bei Kindern und Erwachsenen nach Tschernobyl	59
Tabelle 13: Varianz innerhalb und zwischen den Gruppen für die Simulationsstudie	65
Tabelle 14: Anzahl und Prävalenz der Schilddrüsenkrebsfälle in Reihenuntersuchungen (mass screening), die in der Ukraine bis zum Jahr 2000 durchgeführt wurden. AR steht für Alter bei der Reihenuntersuchung, AU für Alter zur Zeit des Unfalls	72
Tabelle 15: Anzahl der Siedlungen, geschätzte Anzahl der exponierten Kinder im Jahr 1986, beobachtete Anzahl von Schilddrüsenkrebsfällen von 1990-99 und geschätzte arithmetische Dosismittelwerte in der vorläufigen Dosisschätzung	75
Tabelle 16: Anzahl der Siedlungen, geschätzte Anzahl der exponierten Kinder im Jahr 1986, beobachtete Anzahl von Schilddrüsenkrebsfällen von 1990-99 und geschätzte arithmetische Dosismittelwerte in der endgültigen Dosisschätzung	76
Tabelle 17: Schätzer für das Hintergrundrisiko \bar{h}_{0k} aus der Inzidenz I_k^{90-99} mit den zusätzlichen Anstiegsfaktoren \bar{n}_k^{90-99} , den Dosismittelwerten \bar{D}_k und einem ERRPD γ von 10 Gy^{-1} (Variante 1) und 20 Gy^{-1} (Variante 2)	79
Tabelle 18: Haupttypen von im Zeitraum 1990-99 operierten Schilddrüsenkarzinomen in der Geburtskohorte 1968-85	81
Tabelle 19: Größenverteilung von im Zeitraum 1990-99 operierten Schilddrüsenkarzinomen in der Geburtskohorte 1968-85	81

Anhänge

1. Literaturreview zu ökologischen Studien und zu Schilddrüsenkrebs, Kapitel 4 und 5 des Zwischenberichtes zu diesem Projekt
2. Comparison of the time-integrated thyroid activities for different age-gender-groups estimated in the first and second levels of thyroid dose estimation. Statistical properties of the distributions
3. Morphologic analysis of thyroid carcinomas selected for carrying out investigations in the framework of the StSch4299 project
4. Bias durch Überdiagnose von papillären Mikrokarzinomen
5. A methodology to assess possible effects of enhanced surveillance on the risk estimate from ecologic studies of thyroid cancer after the Chernobyl accident
6. Implications of increased thyroid cancer detection and reporting on risk estimations after the Chernobyl accident
7. Pure specification bias, Tabellen und Abbildungen
8. Klassischer Fehler, Berkson-Fehler und Ökologischer Messfehler
9. Bias in ökologischen Studien bei nichtlinearen Risikomodellen

1 Einleitung: Methodische Probleme zur Bestimmung von Risikofaktoren aus ökologischen Studien

Epidemiologische Studien werden in Studien mit individuellen Daten (Kohortenstudien, Fallkontrollstudien) und solche mit zusammengefassten Daten (Studies with aggregate data, ecologic studies) unterschieden. Die erstgenannten haben den Vorteil, dass Verfahren zur Untersuchung von Risikomodellen, z.B. der Untersuchung von Dosis-Wirkungsbeziehungen auch unter der Berücksichtigung von Confoundern verfügbar und mathematisch abgesichert sind, und den Nachteil großer Kosten. Studien mit zusammengefassten Daten können mit meist geringerem Aufwand große Fallzahlen erfassen. Andererseits haben ökologische Studien erhebliche methodische Probleme, insbesondere wenn kausale Aussagen gefordert sind („ecological bias“). Es ist bekannt, dass die Variation eines Confounding-Faktors (wie das Rauchen beim Lungenkrebs durch Radon) innerhalb der zusammengefassten Daten zu nicht belastbaren Ergebnissen führen kann. Andererseits sind Untersuchungen mit aggregierten Daten in den durch den Tschernobyl-Unfall kontaminierten Gebieten bisher die einzigen, die zu quantitativen Ergebnissen über das Schilddrüsenkrebsrisiko nach ^{131}I Expositionen während der Kindheit geführt haben. Durch diese Ausgangslage motiviert, werden im vorliegenden Abschlussbericht die Fragestellungen der Belastbarkeit und des möglichen Anwendungsbereiches von ökologischen Studien untersucht.

Dabei behandeln die Abschnitte 2 und 3 ökologische Studien im Allgemeinen, während in den Abschnitten 4 und 5 direkt auf die spezielle Situation nach dem Tschernobylunfall eingegangen wird. Zunächst werden im Abschnitt 2.1 die theoretischen Grundlagen skizziert. In einer Reihe von Arbeiten werden die Probleme von ökologischen Studien dargestellt (s. Abschnitt 2.2 und Anhang 1), systematische Untersuchungen, unter welchen Bedingungen quantitative Information über Risikofaktoren aus ökologischen Studien gewonnen werden können, sind aber selten. Deshalb werden im Abschnitt 3 (und Anhang 7) Simulationen zum sogenannten „pure specification bias“ vorgestellt, der ein generelles Problem für ökologische Studien darstellt. Im Abschnitt 4 wird speziell die Situation nach dem Tschernobylunfall diskutiert. Zunächst in Abschnitt 4.1 das Problem der zusätzlichen Schilddrüsenkrebsfälle diskutiert, die durch Reihenuntersuchungen (screening) gefunden wurden. Auch wurde in Abschnitt 4.2 eine Literaturübersicht zur Epidemiologie und Morphologie der Schilddrüsenkarzinome erstellt. Dann werden in Abschnitt 4.3 die Datengrundlagen zur Bevölkerungsverteilung, Exposition und Inzidenz von Schilddrüsenkrebs vorgestellt. Ausgehend von Daten aus der Ukraine zur Strahlenbelastung durch den Reaktorunfall und zu Inzidenzraten für Schilddrüsenkrebs bei Kindern wurden im Abschnitt 5 systematische Analysen durchgeführt, mit dem Ziel exemplarisch Bedingungen zu definieren, unter denen epidemiologische Studien mit zusammengefassten Daten zu belastbaren Ergebnissen für das Risiko führen können. Abschnitt 6 fasst die Ergebnisse des Forschungsvorhabens zusammen und zieht ein Fazit.

2 Grundlagen

2.1 Methodische Grundlagen

Vorbemerkung: Es sollen hier zunächst einige Grundlagen der Risikomodellierung in der Strahlenepidemiologie dargestellt werden. Modellierung heißt Beschreibung der Abhängigkeit der Erhöhung der Krebsraten (Mortalität oder Inzidenz) von der Strahlendosis, wobei auch die Berücksichtigung von Faktoren wie Alter bei der Exposition, Geschlecht, Zeit zwischen Exposition und Auftreten der Erkrankung zur Modellierung gehört. Beobachtungen aus Hiroshima und Nagasaki und vielen andere epidemiologischen Studien haben dazu geführt, dass unterschiedliche Modelle für Leukämien und solide Tumoren betrachtet werden. Prinzipiell gibt es zwei Ansätze: Exposition gegenüber ionisierender Strahlung erhöht die altersabhängigen Spontanraten *proportional*. Das heißt, das Verhältnis der Krebsraten der Personen, die nicht exponiert sind, und der Personen, die exponiert sind, ist konstant. Da die Krebsraten mit dem Alter steigen, steigt das absolute Zusatzrisiko (also die Differenz der Krebsraten der beiden Gruppen) mit dem Alter, während das relative Zusatzrisiko konstant bleibt. Der Proportionalitätsfaktor (d.h. das *relative Risiko*) bleibt zeitlich konstant, hängt aber von der Dosis ab (und außerdem von dem Alter bei der Exposition und vom Geschlecht). Modellierung heißt hier also die Beschreibung der Abhängigkeit des relativen Risikos von der Dosis, dem Alter bei der Exposition und gegebenenfalls anderen Faktoren.

Statistische Aspekte: Regressionsmodelle gehen von der grundsätzlichen Überlegung aus, dass der Erwartungswert einer (Zufalls)Variable Y mittels eines Modells durch eine oder mehrere erklärende Variablen X beschrieben werden kann, also $E(Y|X=x) = f(x)$. Im medizinischen Kontext ist dieser Zusammenhang im Allgemeinen sehr komplex und jede Beschreibung kann nur als Vereinfachung – also als Modell – des „wahren“ Zusammenhangs gesehen werden. Im Bereich der Epidemiologie werden verschiedene Modelle benutzt, um das Auftreten von Krankheiten und die Beziehung zwischen Krankheit und Risikofaktoren zu beschreiben. Im Folgenden sollen diese Modelle in einfacher Form beschrieben werden, um Klarheit in die Nutzung der Begriffe zu bringen und einen Überblick zu schaffen. Diese Arbeit lehnt sich eng an die Arbeiten von Greenland (1997) und Breslow und Day (1980, 1987) an.

(1) Als Maßzahlen zur Beschreibung der Krankheitshäufigkeit (Auftreten) werden hier nur die Inzidenzrate oder das Risiko (Incidence proportions) für Neuerkrankungen oder für Mortalität betrachtet. Auf die Untersuchung von Überlebenszeiten wird am Ende kurz eingegangen. Die Raten beziehen sich hier auf ein Kalenderjahr. Die Inzidenzrate ist der Quotient aus der Zahl der „Fälle“ (Erkrankten einer definierten Krankheit, hier Schilddrüsenkrebs) bezogen auf die Personjahre aller unter Risiko stehenden Personen in einem Zeitraum. Das Risiko bezeichnet die Wahrscheinlichkeit, dass eine bestimmte Person in einer gegebenen Zeit eine Krankheit entwickelt. Inzidenzrate bezieht sich also auf eine Population, Risiko auf eine Person. Manchmal wird der Begriff „Inzidenzrate“ auch als Durchschnittsrisiko aller Personen in der Population bezeichnen. Statistisch kann die Inzidenzrate als ein Schätzer für das Risiko betrachtet werden (wie auch die relative Häufigkeit ein Schätzer für die Wahrscheinlichkeit ist). Im Folgenden wird I für Inzidenzrate und R für Risiko verwendet.

(2) Als Maßzahlen für den Vergleich von Raten und Risiko werden Ratendifferenz und Ratenverhältnis betrachtet, bzw. Risikodifferenz und Risikoverhältnis. Raten- und Risikoverhältnis werden häufig beide „*relatives Risiko*“ genannt. Im Allgemeinen wird dies

damit begründet, dass die Rate als Approximation für das Risiko gesehen werden kann, streng genommen muss jedoch zwischen relativer Rate (rate ratio) und relativem Risiko unterschieden werden. In manchen Publikationen wird auch der Begriff „rate ratio“ (richtig) verwendet. Als weitere Maßzahl wird das odds ratio betrachtet, der Quotient aus zwei Odds (Quotenverhältnissen).

(3) Raten werden häufig als „absolute“ Maßzahlen bezeichnet. Soll die Inzidenzrate einer „exponierten“ Population mit einer „nichtexponierten“ Population betrachtet werden, werden Quotienten von absoluten Maßzahlen berechnet, also die Inzidenzrate der Exponierten (bezeichnet mit I_1) wird in Beziehung mit einer (absoluten) Basisrate I_0 gestellt. Dabei wird entweder die Inzidenzrate der „Exponierten“, I_1 (Quotient = I_1/I_0), manchmal aber auch die Differenz der Inzidenzraten (Quotient = $(I_1 - I_0)/I_0$) im Nenner betrachtet. Der absolute Effekt einer Exposition wird mit $I_1 - I_0$ beschrieben, und der relative Effekt mit $(I_1 - I_0)/I_0 = I_1/I_0 - 1$. Dieses Verhältnis heißt korrekterweise „Relative Exzess Rate“ (*relative excess rate*). Da diese Maßzahlen von I_0 abhängen, kann es vorkommen, dass bei gleicher Ratendifferenz die Relativen Effekte unterschiedlich sind. Zu beachten ist, dass in der Epidemiologie häufig (wenn nicht sogar immer) nur der Quotient I_1 / I_0 angegeben wird (und relatives Risiko genannt wird). Sprachlich wird aber nicht immer klar unterschieden: „das Risiko ist um 60% erhöht“ heißt, dass der Quotient von I_1 und $I_0 = 1,6$ ist. Die relative Exzess-Rate wird häufig auch als „excess relative risk“ bezeichnet. Gerade in der Strahlenepidemiologie wird im Allgemeinen von diesem Wert gesprochen.

(4) Eine etwas andere Betrachtungsweise ist die Modellierung der Inzidenzzeit. Hier wird *die Zeit bis zum Auftreten eines Ereignisses* modelliert, eine Methode, die in klinischen Studien eine lange Tradition hat („survival-time-models“, „failure time models“), aber in epidemiologischen Studien erst in den letzten Jahren an Bedeutung gewinnt. Beachte, hier ist die Zeit T eine Zufallsvariable, die Werte zwischen 0 und dem theoretischen Wert unendlich annehmen kann. Neben der Zeit muss bei der Modellierung die „Zensierung“ betrachtet werden, also ob die Zeit bis zum Auftreten des Ereignisses (also hier bis zum Auftreten eines Schilddrüsenkrebs) tatsächlich beobachtet wurde oder die Beobachtung vorher aus pragmatischen Gründen (Ende der Studie, andere Todesursache, lost to follow-up) vorher „zensiert“ also abgebrochen wurde. Auf die Modellierung der Inzidenzzeit wird hier nicht eingegangen.

(5) Das zugrunde liegende statistische Modell kann auch so beschrieben werden: Sei π die Wahrscheinlichkeit, dass ein Ereignis bei einer Person zu einem bestimmten Zeitpunkt innerhalb einer gegebenen Zeit auftritt, die bis zu diesem Zeitpunkt frei von diesem Ereignis war. Statistisch gesehen handelt es sich um Bernoulli-verteilte Zufallsvariablen. Betrachtet man hingegen eine Gruppe von Personen und die Zählvariable („Anzahl der Fälle d , die in der Zeit t auftreten“), so kann dies durch einen Poissonprozess modelliert werden. Die kumulative Intensität bzw. der Parameter für die Raten wird im Modell durch den Parameter λ beschrieben. Die Wahrscheinlichkeit für das Auftreten eines Ereignisses in einem Intervall der Länge s wird durch $\lambda_s \cdot s$ beschrieben. Die „rohe Rate“, also der Quotient aus Zahl der „Fälle“ und Personenjahre ist ein Schätzer für λ . Seien λ_1 und λ_2 die Raten zweier (homogener) Poissonprozesse, so wird das Relative Risiko (also das Rate ratio) durch den Quotient der beiden Parameter geschätzt.

(6) Wird die Inzidenzrate (I) oder das Relative Risiko (RR) als Funktion einer stetigen Variable (D , hier Strahlendosis) untersucht, so kann z.B. folgendes Modell angenommen werden:

- a. $I(s; D) = f(\text{strata}, \text{Dosis})$,
 b. $RR(s; D) = f(\text{strata}, \text{Dosis})$

wobei s z.B. für Altersgruppen, Geschlecht oder andere Confounder stehen kann, und D die (Strahlen)dosis benennt.

In der Epidemiologie wird klassischerweise das Modell

c. $RR(s; D) = \exp(a + bD)$

genutzt. Dies hat sowohl theoretisch-statistische als auch (teilweise) biologisch-medizinische Gründe. Die methodische Grundlage für dieses Modell ist, dass der Zusammenhang zwischen dem klassischen Ansatz der 2×2 Tafel und dem Modell für den Fall einer binären Exposition ($D = 0$ oder 1) einfach hergeleitet werden kann und dass es der „natürliche“ Link in dem logistischen Modell ist. Außerdem erfüllt dieses Modell die Bedingung, dass das RR zwischen 0 und unendlich liegen muss, also nicht negativ werden kann.

Ein besonders in der Strahlenepidemiologie bevorzugtes Modell dieser Art ist z.B.

d. $RR(s; D) = 1 + \alpha D$,

in diesem Fall hängt die Dosis-Wirkungsfunktion nicht von s ab und das Relative Risiko wird als linear in der Dosis betrachtet (siehe unten). Dieses Modell wird aus strahlenbiologischen Überlegungen bevorzugt.

e. $RR(s; D) = 1 + \alpha_s D + \beta_s D^2$

Hier wird angenommen, dass es eine linear-quadratische Abhängigkeit des Relativen Risikos von der Dosis gibt. Beide Parameter unterscheiden sich für die Strata s (z.B. $s = 1$ für Frauen, $s = 2$ für Männer, wenn die Parameter sich für die beiden Geschlechter unterscheiden). In der Strahlenbiologie wurde folgendes Modell bevorzugt, das in experimentellen Studien den Zusammenhang zwischen Krebsentstehung und Strahlenexposition gut beschreibt.

f. $RR(s; D) = [1 + \alpha D + \beta D^2] \left[\exp(\gamma D + \delta D^2) \right]$

Dieses Modell wird aber in der Strahlenepidemiologie praktisch nicht angewandt, da die Daten nicht hinreichend genau sind, um alle Parameter zu schätzen. Außerdem ist bei kleiner bis mittlerer Dosis die Funktion in Gleichung (f) schon recht gut durch die Gleichung (d) approximiert.

(7) In der Strahlenepidemiologie wird meistens die o.g. Form angegeben, wobei D die Strahlendosis, der lineare Exzess-Relative Risiko (ERR) Koeffizient (pro Sv) und der quadratische ERR Koeffizient (pro Sv^2) ist, wobei dies für das s -te Stratum gilt. Um diese Parameter zu schätzen, betrachtet man die beobachteten Fälle pro Dosisgruppe (Geschlecht oder andere Strata) und geht davon aus, dass die erwartete Anzahl von Fällen sich aus einem Poissonprozess ergibt und so geschrieben werden kann:

$$E(\text{Fälle}, s; d) = \text{PYR}_{s_j} \cdot \lambda_s \cdot (RR(s, D) | d)$$

wobei λ_s die Inzidenzrate in der s -ten Gruppe ist, wenn keine Exposition vorhanden wäre (sogenannte „base-line-incidence“) und PYR die entsprechenden Personenjahre. Konkret für ein Stratum Geschlecht sieht das dann so aus:

$$E(\text{Fälle}, s = 1; d) = \text{PYR}_{1,j} \cdot \lambda_1 \cdot (1 + (\beta_1 \cdot d))$$

Für $s=1$ und ein Modell, in dem das RR als lineares Modell betrachtet werden kann.

(8) Im Folgenden werden die Modelle, die für Risiko R und Relatives Risiko RR und für das odds ratio OR genutzt werden, kurz genannt:

a. Konstantes Modell:

die Risiken werden in allen Subpopulationen als konstant angesehen: Sei x die Dosis, so gilt:

i. $R(x) = \alpha$.

ii. $R(x) = \exp(\alpha')$, hier ist α' der (konstante) Wert für das log Risiko

iii. $\text{Ln}(R(x)) = \alpha'$

$R(x) = \exp(\alpha'' / (1 + \exp(\alpha'')))$, hier ist $\text{logit } R(x) = \alpha''$

b. Lineares Modell:

die Risiken variieren für Subpopulationen und zwar so, dass sie linear von der Exposition x (oder Strahlendosis D) abhängen. Man geht also davon aus, dass das Risiko (oder eine der unten genannten Funktionen davon) der Subgruppe mit Exposition $X=x+1$ von $X=x$ sich um einen konstanten Faktor unterscheiden. Hier ist (im Gegensatz zum konstanten Modell) zu beachten, für welche Funktion von R die lineare Beziehung gilt (siehe auch Tabelle 2)

$$R(x_1) = \alpha + \beta x_1, \text{ also } R(x_1 + 1) - R(x_1) = \beta.$$

Der Parameter β beschreibt den Unterschied zwischen Populationen, bei denen sich X um die Größenordnung 1 verändert. Dieses Modell wird als *lineares Risikomodell* bezeichnet. Manchmal wird nicht die Dosis X , sondern X/c oder $X-c$ betrachtet, damit eine Interpretation der β sinnvoll ist (Beispiel: also nicht Gray sondern $m\text{Gray}$, nicht Alter sondern Zeit nach dem 50. Geburtstag ($X-50$)).

c. Exponentielles Modell:

dies wird geschrieben in der Form

$$R(x_1) = \exp(\alpha + \beta x_1).$$

Der Vorteil zum linearen Modell ist, das $R(x)$ immer positiv ist. Umschreiben ergibt sich:

$$\text{ln}(R(x_1)) = (\alpha + \beta x_1).$$

Daher wird dieses Modell auch oft als log-lineares Modell bezeichnet (woher auch statistisch methodische Eigenschaften hergeleitet werden können). Beim Vergleich der Risiken von Subpopulationen mit unterschiedlichen Werten für X gilt, dass die Differenz der logarithmierten Risiken konstant ist, wenn X um den Betrag 1 wächst, also $\text{ln}(R(x_1+1)) - \text{ln}(R(x_1)) = \text{ln}(R(x_1+1)/R(x_1)) = \beta$. In anderen Worten α ist das ln-Risiko für die Personen mit $X=0$ (also für die nicht exponierten Personen) und $\exp(\beta)$ ist der Risikoquotient („risk ratio“)

zwischen Subpopulationen bei denen X um eine Einheit steigt (also z.B. mit 2 vs. 3 Gray aber auch 5 vs. 6 Gray).

d. Logistisches Modell:

$$R(x_1) = \exp(\alpha + \beta x_1) / (1 + \exp(\alpha + \beta x_1)) \text{ oder}$$

$$\text{Logit } R(x_1) = (\alpha + \beta x_1).$$

Alle Werte für R(x) in diesem Modell liegen zwischen 0 und 1. Man spricht daher auch vom *logit-linear Risikomodell*. Das Modell macht vor allem dann Sinn, wenn man die odds (geschrieben O) betrachtet. Es gilt nämlich gerade, dass

$$O(x_1) = R(x_1) / (1 - R(x_1)) = \exp(\alpha + \beta x_1).$$

Das logistische Modell ist daher ein exponentielles Modell für das odds. Es wird auch log-lineares odds Modell genannt, da $\ln O(x_1) = \alpha + \beta x_1$.

e. Andere Modelle:

Selten, aber gerade in der Strahlenepidemiologie immer wieder, werden andere Modelle betrachtet, nämlich ein lineares Modell für das odds, also

i.
$$O(x_1) = \alpha + \beta x_1.$$

In diesem Modell muss β als der odds in Subpopulation mit dem Unterschied 1 in x interpretiert werden.

(9) Zur Vermeidung von Missverständnissen ist zu empfehlen, nicht nur den „Namen“ eines Modells anzugeben, sondern auch den Parameter der modelliert wird. Also, zu sagen, wir modellieren ein logistisches Modell für R(x), ein exponentielles Modell für odds(x), oder lineares Modell für ln(odds). Die Zusammenhänge werden in den Tabellen 1 und 2 noch einmal dargestellt.

(10) Für die Inzidenzzeit (siehe oben) kann nun folgende Betrachtung durchgeführt werden: T ist eine positive Zufallsvariable, daher ist ln(T) definiert.

a.
$$\ln E(T | X = x) = \alpha + \beta D,$$

Die Parameter α und β sind hier natürlich anders zu interpretieren. α ist die durchschnittliche ln(Inzidenzzeit) in der Population für die X=0 ist, oder $\exp(\alpha)$ ist die durchschnittliche Inzidenzzeit der "nicht exponierten" und β ist die Differenz der durchschnittlichen ln(Inzidenzzeit), wenn die Populationen mit X=x+1 und X=x verglichen werden. Zu beachten ist, dass im Modell üblicherweise $-\beta$ betrachtet wird, um eine einheitliche Interpretation zu ermöglichen, $\beta > 0$ bedeutet, dass X ein Risikofaktor ist, (also exponierte Personen eine kürzere Überlebenszeit haben). Zur Modellierung der Überlebenszeit wird häufig das so genannte Cox-Modell genutzt. Auf diese Modelle soll hier nicht mehr eingegangen werden, da für unsere Fragestellung keine Überlebenszeiten vorliegen.

Tabelle 1: Ableitungen zum Vergleich von Risiken, mit R_1 = Risiko für Subpopulation $X=x+1$ und R_0 = Risiko für Subpopulation mit $X=x$)

Type	Ausdruck	Wertebereich	„kein Effekt“
Risikodifferenz	$R_1 - R_0$	-1, 1	0
Relatives Risiko	R_1 / R_0	0, unendlich	1
ln RR	$\ln R_1 - \ln R_0$	- unendlich; + unendlich	0
odds ratio (OR)	$R_1 / (1 - R_1) / R_0 / (1 - R_0)$	0, unendlich	1
ln OR	$\ln(R_1 / (1 - R_1)) - \ln(R_0 / (1 - R_0))$	- unendlich; + unendlich	0

Tabelle 2: Zusammenhang der oben beschriebenen Modelle

Parameter			
R= Risiko	$R = a + bx$ Lineares Modell	$R = \exp(a + bx)$ Exponentielles Modell	$R = \text{expit}(a + bx)$ Logistisches Modell
ln (R)		$\ln(R) = a + bx$ log-lineares Risikomodell	
logit (R)			logit (R) = $a + bx$ (1) Logit-linear Risikomodell, (2) Logit Modell
O= odds = $(R / (1 - R))$			$O = \exp(a + bx)$ (1) Exponentielles Modell für odds, (2) logit lineares Risikomodell, (3) logit Modell
ln (O)			$\ln(O(x)) = a + bx$ (1) Log-lineares odds Modell
Vergleich von Subpopulation mit $X=x+1$ bzw $X=x$	b ist Differenz der Risiken	b ist $\ln(\text{RR})$ $\exp(b)$ ist RR	$\exp(b)$ ist das odds ratio

2.2 Literaturübersicht: Methodische Probleme ökologischer Studien

2.2.1 Ökologische Studien: Definition und Kritik

2.2.1.1 Definition

In einer ökologischen Studie werden Daten auf Gruppenebene erhoben und analysiert. Exposition, Erkrankung und Kovariablen stellen Eigenschaften von Gruppen von Individuen dar und nicht Eigenschaften der Individuen selber.

In der Regel werden ökologische Studien durchgeführt, weil Daten auf individueller Ebene fehlen. In diesem Sinne bezeichnet man ökologische Studien auch als unvollständiges Studiendesign (Kleinbaum et al. 1982). In der Epidemiologie werden ökologische Studien seit etwa 150 Jahren zu verschiedenen Fragestellungen durchgeführt und haben zum Teil erhebliche gesellschaftliche und politische Konsequenzen.

Der Name „ökologische Studie“ geht auf frühe epidemiologische Studien Mitte des 19. Jahrhunderts zurück. In einer der ersten epidemiologischen Studien überhaupt, untersuchte John Snow den Zusammenhang zwischen der Anzahl Erkrankter oder Verstorbener pro Londoner Mietshaus mit strukturellen Variablen, z.B. dem Trinkwasseranschluss der Häuser (Snow 1854, zit. nach Gordis 2000). Später bürgerte sich für dieses Studiendesign, bei dem die Beobachtungseinheit Gruppen von Individuen waren, abgeleitet vom griechischen „Oikos“ =Haus, der Begriff ökologische Studien ein. In der Epidemiologie bezeichnet der

Begriff „Ökologische Studien“ ein Studiendesign. Es besteht kein Zusammenhang mit der biologischen Forschungsrichtung Ökologie, der Lehre von der Abhängigkeit eines Lebewesens von den biotischen und abiotischen Umweltfaktoren. Analog zu dem Begriff ökologische Studien (ecologic studies) werden die Begriffe aggregierte Studien (aggregated studies) oder Studien auf Gruppenebene (studies on group level) verwendet (Greenland 2001).

2.2.1.2 Kritik an ökologischen Studien

Der Zusammenhang zwischen Risikofaktoren und ihren gesundheitlichen Folgen wirkt sich auf unterschiedlichen Ebenen aus.

Auf physiologischer Ebene, auf der sich Reaktionen körpereigener biochemischer Moleküle mit in den Körper eingedrungenen chemischen Schadstoffen oder energiereicher Strahlung ereignen. Auf individueller Ebene, auf der durch Exposition gegenüber einer individuellen Dosis eines Risikofaktors eine individuelle gesundheitliche Zielvariable verursacht wird. Auf Gruppenebene, auf der ein für alle Gruppenmitglieder gleicher Risikofaktor bei einem Teil der Individuen gesundheitliche Folgen verursacht und z.B. eine bestimmte Rate von Neuerkrankungen hervorruft.

Diese Ebenen werden im Folgenden am Beispiel von Schilddrüsenkrebs dargestellt: Ein unkontrolliert wachsender Tumor ist die Folge einer einzigen, entarteten Schilddrüsenzelle, deren Teilungstätigkeit sich der körperlichen Kontrolle entzieht. Die Ursache dafür ist wiederum ein komplexes Zusammenspiel einer Mutation an bestimmten Stellen der DNA und eines Versagens der zelleigenen Reparaturmechanismen. Diese Ebene bezeichnen Epidemiologen als physiologische Ebene. Das Risiko einer einzelnen Follikelzelle der Schilddrüse zu entarten ist eine Funktion ihres Grundrisikos und ihrer Exposition, z.B. gegenüber ionisierender Strahlung, sowie weiteren kanzerogenen Stoffen. Das Grundrisiko wird durch die genetische Ausstattung der Zelle bedingt, die ererbte oder neu eingetretene Mutationen enthalten kann. Weitere Einflussfaktoren sind z.B. der Zustand im Mitosezyklus, der Stoffwechszustand (anabolisch/metabolisch für Thyreoglobulin), frühere DNA-Schädigungen, die den Reparaturmechanismus beeinflussen und das Zellalter. Viele weitere Einflussfaktoren sind denkbar. Diese Risikofaktoren sind für die Zellen einer menschlichen Schilddrüse nicht zu jedem Zeitpunkt gleich. Die Zellen sind auch nicht gleichermaßen gegenüber einem kanzerogenen Faktor von außen exponiert. Eine individuelle Messung der Risikofaktoren auf Ebene der Zellen wäre nur unter erheblichem labortechnischem Aufwand möglich.

Auf dieser physiologischen Ebene kann der Einfluss von Risikofaktoren auf menschliche Erkrankungen, z.B. auf Krebserkrankungen, aus ethischen Gründen meistens nur in Tierexperimenten oder an Zellkulturen simuliert werden. Die Ergebnisse werden auf das komplexe System des menschlichen Organismus übertragen. Inwieweit eine solche Übertragung den physiologischen Vorgängen im menschlichen Körper entspricht, bleibt fraglich (Cohen 1994, Greenland 2001).

In epidemiologischen Studien auf individueller Ebene, z.B. in Fall-Kontroll- oder Kohortenstudien, wird die Auswirkung gesundheitlicher Risikofaktoren auf das Individuum untersucht. Die individuellen Daten zur Exposition und zu weiteren Risikofaktoren kann man als teilweise recht grobe Aggregationen dieser Faktoren auf Zellebene betrachten. Ob die Zelle oder die Zellen, die bei einem kranken Menschen tatsächlich kanzerogen entartet sind, genau der gemessenen individuellen Exposition des Menschen ausgesetzt waren, ist

unwahrscheinlich, weil es in der Schilddrüse Areale mit unterschiedlicher Aktivität gibt. Auch die Varianz der Exposition aller Zellen einer Schilddrüse ist nicht bekannt.

Obwohl der Schluss von der individuellen Ebene auf die physiologische Ebene ein Schluss von einem einfachen System auf ein höchst komplexes ist, werden individuelle Risikoschätzer häufig auf die physiologische Ebene übertragen. Die individuelle epidemiologische Studie an einer repräsentativen Bevölkerungsstichprobe gilt in der Epidemiologie als Goldstandard (Cohen 1994). Möglicherweise kommt es von dem Schluss des individuellen Risikos in einer populationsbezogenen Untersuchung auf das physiologische Risiko einer einzelnen Zelle eines Zellverbandes ebenso zu „ökologischen“ Verzerrungen, wie man sie bei der Übertragung von Risikoschätzern auf Gruppenebene auf das Risiko auf individueller Ebene beschrieben hat.

Inwieweit Studien auf Gruppenebene einen Beitrag zur Risikoforschung leisten können, wird in der epidemiologischen Literatur intensiv diskutiert. Diese Diskussion wurde von der Beobachtung ausgelöst, dass die Ergebnisse ökologischer Studien zum Teil von denen individueller Studien zur gleichen Fragestellung abweichen. In Einzelfällen kommt es vor, dass der Einfluss der gleichen Exposition auf eine Erkrankung auf ökologischer Ebene ein protektiver Faktor zu sein scheint, auf individueller Ebene aber ein Risikofaktor, oder umgekehrt. Beispiele sind Cohens ökologische Analysen zu Radon und Lungenkrebs (Cohen 1990), oder Durkheims Beobachtungen zu Suizidrisiko und Konfessionszugehörigkeit (Durkheim 1951). Diese Beispiele werden im weiteren Verlauf dieser Arbeit beschrieben. Im Mittelpunkt der Diskussion steht die Frage, ob in diesem Fall der individuelle Schätzer grundsätzlich als valider gelten muss, oder ob unter bestimmten Umständen auch der ökologische Schätzer die wahren Zusammenhänge mit hinreichender Genauigkeit beschreiben kann.

2.2.2 Aufbau des Review „Ökologische Studien“

Eine Medline-Recherche mit dem Suchbegriff „ecologic/aggregate studies“ zeigt von 1990 bis 2002 ca. 70 theoretische Veröffentlichungen zum diesem Thema. Dazu kommt eine weitaus größere Zahl angewandter ökologischer Studien, deren Ergebnisse veröffentlicht wurden. Eine Übersicht der Themen dieser angewandten Studien der letzten zwei Jahre ist in Abschnitt 0. dargestellt. In Abschnitt 2.2.3 werden die gebräuchlichsten mathematischen Modelle in ökologischen Studien und die zugrunde liegenden Modelle auf individueller Ebene erläutert.

Die Diskussion um die Validität ökologischer Studien in der Epidemiologie wurde 1990 mit einer Veröffentlichung von Cohen über den Zusammenhang zwischen Radon und Lungenkrebs erneut aufgenommen (Cohen 1990). In der folgenden Diskussion verteidigte Cohen das ökologische Studiendesign und versuchte Szenarien zu definieren, in denen ökologische Studien mindestens ebenso valide sind wie individuelle Studien (Cohen 1994, 1995). Im Abschnitt 2.2.4 wird das Ergebnis von Cohens Studie im Vergleich zu den Ergebnissen individueller Studien zu Radon und Lungenkrebs, sowie Cohens Argumente für die Validität ökologischer Studien kurz dargestellt. Dieses Beispiel spielt für unsere weiteren Betrachtungen eine Rolle, weil es sich hier auch um die Untersuchung von einer quantitativen (stetigen) Expositionsvariable handelt, während in vielen anderen ökologischen Studien (nur) ein binärer Risikofaktor (und dessen Prävalenz) betrachtet wird.

Eine Reihe anderer Autoren widersprach Cohens Ansicht und belegte die Einschränkungen von Studien auf Gruppenebene mit theoretischen Beispielen und Simulationsstudien (Walter 1991, Greenland und Robins 1994, Piantadosi 1994, Schwartz 1994, Stidley und Samet 1994, Sheppard und Prentice 1995, Sheppard et al. 1996). 1998 veröffentlichte Morgenstern einen Reviewartikel über ökologische Studien, der eine Übersicht über die theoretische Zuordnung verschiedener Studiendesigns gab und die Kritikpunkte an ökologischen Studien übersichtlich darstellte. Die Inhalte dieses Beitrags, und neuere Ergänzungen werden in Abschnitt 2.2.6 zusammengefasst dargestellt.

Einige methodische Probleme, die die Validität ökologischer Studien einschränken können, werden bei Morgenstern (1998) nur kurz umrissen, oder wurden erst in den folgenden Jahren veröffentlicht. Dazu gehört die Konstruktvalidität von Variablen auf Gruppenebene, das Messfehlerproblem, die Varianz der Exposition sowie der Migrations-Bias. Diese Themen werden in Abschnitt 2.2.6 beschrieben.

Auch wenn Gründe für den ökologischen Bias in theoretischen Arbeiten beschrieben werden, bedeutet das nicht, dass sie in der angewandten epidemiologischen Forschung Bedeutung haben. Diese Frage kann nur anhand von Studien geklärt werden, die denselben Datensatz sowohl auf individueller Ebene als auch auf Gruppenebene auswerten. Zwei Studien zum Thema Radon und Lungenkrebs mit parallelen Analysen auf zwei Ebenen wurden in den letzten Jahren in Schweden und England durchgeführt und geben einen Einblick in die praktische Bedeutung des ökologischen Bias. Sie werden in Abschnitt 2.2.7 dargestellt. In der Zusammenfassung, Abschnitt 2.2.8, erfolgt ein kurzer Überblick über den gegenwärtigen Stand der Diskussion über die Validität von ökologischen Studien in der Epidemiologie.

Neuere ökologische Studien im Überblick

Trotz der bekannten Einschränkungen bei der Interpretation ökologischer Studien werden diese aus praktischen Gründen weiterhin häufig durchgeführt. In den Jahren 2001 und 2002 finden sich in einer Medline-Recherche mit dem Suchbegriff „Ecologic studies/study“ 16 englischsprachige Artikel, in denen Ergebnisse rein ökologischer Studien im epidemiologischen Bereich vorgestellt werden. Die Aufstellung ist mit Sicherheit nicht vollständig, weil viele ökologische Studien veröffentlicht werden, ohne dass die aggregierte Daten- und Analyseebene explizit genannt wird. In allen Untersuchungen ist die Zielvariable eine Rate (Mortalitäts-, Inzidenz-, Impf-, Resistenzrate) in einer abgegrenzten Bevölkerungsgruppe. Die Exposition wird in fünf der 16 Studien als kontextuelle Variable erfasst, in sieben weiteren wird die mittlere Exposition aus Umweltdaten geschlossen und in vier der Studien handelt es sich um aggregierte Werte individueller Messungen. Tabelle 3 gibt einen Überblick der Titel und ihrer Expositions- und Zielvariablen.

Tabelle 3: Medline-Recherche Stichwort „ecologic studies“, angewandte ökologische Studien 2001/2002

Autor	Titlel	Ort	Fragestellung
1. Studien mit kontextuellen Expositionsvariablen			
Wiegersma PA, Hofman A, Zielhuis GA	Evaluation of community-wide interventions: the ecologic case-referent study design.	Eur J Epidemiol 2001;17(6):551-7	Langzeitevaluation von Präventionsprogrammen im Jugendalter
Roetzheim RG, Gonzalez EC,	Primary care physician supply and colorectal cancer.	J Fam Pract 2001 Dec;50(12):1027-31	Einfluss ärztlicher Aufklärungsprogramme auf

Ramirez A, Campbell R, van Durme DJ.			die Dünndarmkrebsrate
Diez Roux AV	Investigating neighborhood and area effects on health.	Am J Public Health 2001 Nov;91(11):1783-9 Related Articles, Links	Einfluss der Nachbarschaft auf gesundheitliche Outcomes.
Garau J.	Have guidelines had an impact on antimicrobial resistance?	Semin Respir Infect 2001 Sep;16(3):210-4	Einfluss spezieller gesetzlicher Vorschriften auf die Antibiotikaresistenzrate von pathogenen Keimen.
Young FW, Lyson TA	Structural pluralism and all-cause mortality.	Am J Public Health 2001 Jan;91(1):136-8	Einfluss struktureller Pluralität auf die allgemeine Mortalitätsrate
2. Studien mit Umweltfaktoren als Exposition			
Steiner GG.	Cancer incidence rates and environmental factors: an ecological study.	J Environ Pathol Toxicol Oncol 2002;21(3):205-12	Einfluss verschiedener Umweltfaktoren (Breitengrad, Temperatur, Wasser) auf die Krebsinzidenz
Koifman S, Koifman RJ, Meyer A.	Human reproductive system disturbances and pesticide exposure in Brazil.	Cad Saude Publica 2002 Mar-Apr;18(2):435-45	Zusammenhang zwischen Pestizidverseuchung und Fortpflanzungsrate
Grant WB.	An estimate of premature cancer mortality in the U.S. due to inadequate doses of solar ultraviolet-B radiation.	Cancer 2002 Mar 15;94(6):1867-75	Einfluss von UV-B-Strahlen auf die Kinderkrebsrate
Reynolds P, von Behren J, Gunier RB, Goldberg DE, Hertz A, Harnly ME.	Childhood cancer and agricultural pesticide use: an ecologic study in California.	Environ Health Perspect 2002 Mar;110(3):319-24	Einfluss landwirtschaftlicher Pestizide auf die Kinderkrebsrate
Hopenhayn C, Stump ML, Browning SR.	Regional assessment of atrazine exposure and incidence of breast and ovarian cancers in Kentucky.	Arch Environ Contam Toxicol 2002 Jan;42(1):127-36	Einfluss von Herbiziden auf die Mortalitätsrate von Brust- und Eierstockkrebs
Grant WB.	An ecologic study of dietary and solar ultraviolet-B links to breast carcinoma mortality rates.	Cancer 2002 Jan 1;94(1):272-81	Einfluss von Ernährung und UV-B-Strahlen auf die Brustkrebsmortalität
Aldrich TE, Andrews KW, Liboff AR.	Brain cancer risk and electromagnetic fields (EMFs): assessing the geomagnetic component.	Arch Environ Health 2001 Jul-Aug;56(4):314-9	Einfluss von elektromagnetischen Feldern auf Hirntumore
3. Studien mit aggregierten Expositionsvariablen			
Hull BP, McIntyre PB, Sayer GP.	Factors associated with low uptake of measles and pertussis vaccines--an ecologic study based on the Australian Childhood Immunisation Register.	Aust N Z J Public Health 2001 Oct;25(5):405-10	Einfluss von Rasse und sozioökonomischen Faktoren auf die Impfquote bei Kindern
Koushik A, King WD, McLaughlin JR.	An ecologic study of childhood leukemia and population mixing in Ontario, Canada.	Cancer Causes Control 2001 Aug;12(6):483-90	Einfluss der Bevölkerungsvermischung auf kindliche Leukämien

Benigni R, Giaimo R, Matranga D, Giuliani A.	The potential impacts of climate variability and change on health impacts of extreme weather events in the United States.	Environ Health Perspect 2001 Jul;109(7):705-9	Einfluss der Lebensbedingungen für Frauen auf den Frauenanteil der Krebsmortalitätsrate
Smerhovsky Z, Landa K, Rossner P, Brabec M, Zudova Z, Holan N, Pokorna Z, Mareckova J, Hurychova D	Risk of cancer in an occupationally exposed cohort with increased level of chromosomal aberrations.	Environ Health Perspect 2001 Jan;109(1):41-5	Zusammenhang zwischen der Häufigkeit bestimmter Chromosomenaberrationen und der Krebsinzidenz in einer exponierten Population

Die abhängige Variable bei 16 der 18 genannten ökologischen Studien ist eine Rate, der eine binäre individuelle Variabel (krank/gestorben: ja/nein) zugrunde liegt. Die unabhängige Variable ist bei den kontextuellen und aggregierten Variablen ebenfalls der Anteil einer binominal oder nominal skalierten Variablen in der Gruppe. Bei den Umweltfaktoren, die nicht auf individueller Ebene gemessen wurden, obwohl eine individuelle Variation der Exposition vorliegt, handelt es sich bei allen Studien um Variablen mit stetigem Messniveau. In der theoretischen Literatur über ökologische Studien ist die Situation einer auf individueller Ebene binären Einflussvariablen auf eine ebenfalls binäre Zielvariable am häufigsten untersucht worden. Deutlich geringer ist das Wissen über das Verhalten stetiger Variablen in ökologischen Studien.

2.2.3 Mathematische Modelle in ökologischen Studien

In analytischen Studien der Krebs Epidemiologie wird auf individueller Ebene der Effekt einer Exposition auf eine binäre Zielvariable (krank:ja/nein) geschätzt. Dabei wird das Verhältnis der Erkrankungsraten innerhalb der exponierten und der nichtexponierten Unterpopulationen gebildet. Regressionsmodelle, besonders die logistische Regression werden eingesetzt, um zusätzliche Einflussfaktoren in das Modell mit aufzunehmen, oder wenn es sich um stetige Expositionsvariablen handelt. In einer ökologischen Studie ist dieses Vorgehen nicht möglich, weil die gemeinsame Verteilung von Exposition und Krankheit innerhalb der Gruppen nicht bekannt ist. Aus der binären Zielvariablen (krank: ja/nein) wird die stetige Variable „Erkrankungsrate in einem Cluster“. Diese kann dann in einem Regressionsmodell modelliert werden.

Das häufigste Modell bei der Analyse ökologischer Daten ist das lineare Modell: Auf individueller Ebene wird die Wahrscheinlichkeit einer Person zu erkranken in Abhängigkeit von einer binären Expositionsvariablen mit folgendem Modell beschrieben:

$$Y_{ik} = \begin{cases} 0 & \text{nicht krank} \\ 1 & \text{krank} \end{cases} \quad (2.1)$$

X_{ik} beschreibt die Exposition (binär oder anders) für die i-te Person des k-ten Clusters.

Modell 1: lineares Modell

Modell 1 auf individueller Ebene:

$$P(Y_{ik} = 1 | X_i = x_i) = \alpha + \beta x_i + \varepsilon \quad (2.2)$$

Modell 1 auf ökologischer Ebene:

$$P_k = B_0 + B_1 P_k^x + \varepsilon \quad (2.3)$$

wobei P_k der Anteil Erkrankter im k-ten Cluster ist und P_k^x Anteil exponierter Personen in diesem Cluster (bzw. die mittlere Exposition bei einer stetigen Variable) ist. Zu beachten ist, dass entsprechende Restriktionen (Wertebereich) für die Parameter α , β , B_0 und B_1 notwendig sind, damit $0 \leq P \leq 1$ gilt.

Im Fall einer binären Expositionsvariablen ist das relative Risiko (RR) auf individueller Ebene das Risiko zu Erkranken der Exponierten geteilt durch das Risiko zu erkranken der Nichtexponierten. Auf ökologischer Ebene entspricht diesem das Verhältnis der Erkrankungsraten in einem Cluster, in dem alle Personen exponiert waren und in einem Cluster, in dem niemand exponiert war. In ökologischen Studien mit linearen Risikomodellen lässt sich das RR angeben als:

$$RR = 1 + \frac{B_1}{B_0} \quad (2.4)$$

Das RR ist in diesem Fall sowohl vom Grundrisiko als auch vom expositionsspezifischen Risiko abhängig.

Modell 2: log-lineares Modell

Modell 2 auf individueller Ebene:

$$\ln P(Y_{ik} = 1 | X_i = x_i) = \tilde{\alpha} + \tilde{\beta} x_i + \varepsilon \quad (2.5)$$

Modell 2 auf ökologischer Ebene:

$$\ln P_k = \tilde{B}_0 + \tilde{B}_1 P_k^x + \varepsilon \quad (2.6)$$

wobei P_k der Anteil Erkrankter im k -ten Cluster ist und P_k^x Anteil exponierter Personen in diesem Cluster (bzw. die mittlere Exposition bei einer stetigen Variablen). Das Relative Risiko lässt sich berechnen als:

$$RR = \exp(\tilde{B}_1) \quad (2.7)$$

In diesem Fall ist das RR nicht vom Grundrisiko der Erkrankung abhängig.

Modell 3: logistisches Modell

In der Epidemiologie wird, um den Effekt einer binären oder stetigen Exposition auf eine binäre Zielvariable zu schätzen, häufig die Logistische Regression angewendet.

Modell 3 auf individueller Ebene:

$$\begin{aligned} \text{logit}P(Y_{ik} = 1 | X_i = x_i) &= \tilde{\alpha} + \tilde{\beta}x_i + \varepsilon \\ \Leftrightarrow P(Y_{ik} = 1 | X_i = x_i) &= \frac{\exp(\tilde{\alpha} + \tilde{\beta}x_i + \varepsilon)}{1 + \exp(\tilde{\alpha} + \tilde{\beta}x_i + \varepsilon)} \end{aligned} \quad (2.8)$$

Modell 3 auf ökologischer Ebene:

$$\text{logit}P_k = \tilde{B}_0 + \tilde{B}_1 P_k^x + \varepsilon \quad (2.9)$$

wobei P_k der Anteil Erkrankter im k -ten Cluster ist und P_k^x Anteil exponierter Personen in diesem Cluster (bzw. die mittlere Exposition bei einer stetigen Variablen). Im Falle eines logistischen Modells ist das RR weniger handhabbar:

$$RR = \frac{\exp \tilde{B}_1 \times (1 + \exp \tilde{B}_1)}{1 + \exp(\tilde{B}_0 + \tilde{B}_1)} \quad (2.10)$$

Das Odds-Ratio (OR) lässt sich dagegen einfach darstellen als:

$$OR = \exp \tilde{B}_1 \quad (2.11)$$

Bei seltenen Erkrankungen, zu denen die Krebserkrankungen zählen, kann das OR als Annäherung an das RR betrachtet werden.

Bisher haben wir eine vereinfachte Beschreibung der drei wichtigsten Modelle dargestellt. Es fehlt die Einbeziehung der Confounder und Interaktionen in diese Modellgleichungen. Dieser Teil wird im Verlauf der Simulationsstudien noch weiter bearbeitet. Bemerkenswert ist aber, dass häufig in ökologischen Studien das lineare Modell benutzt wird (und in der Literatur die Rechtfertigung des Studiendesigns darauf beruht), während beim Vorliegen individueller Daten sowohl aus statistisch-mathematischen Erwägungen als auch wegen der biologischen Plausibilität dem logistischen Modell der Vorzug gegeben wird.

2.2.4 Ein prominentes Beispiel für einen Bias in ökologischer Studien: Cohens ökologische Analyse zu Radon und Lungenkrebs

Ökologische Studien werden in der deskriptiven und analytischen Epidemiologie weiterhin häufig durchgeführt. Die neuere Kritik an Risikoschätzern aus ökologischen Studien wurde von einer Untersuchung zum Risiko von Radon im Niedrigdosisbereich auf die Inzidenz von Lungenkrebs angestoßen (Cohen 1990). Die Studie beruht auf Mortalitätsraten und aggregierte Umweltmessungen von Radon in Wohnräumen für 1601 Regionen in den USA. Aufgrund des überraschenden Ergebnisses und der nachfolgenden Kritik wurden die Daten noch einmal weitergehend analysiert (Cohen 1995). Cohen geht davon aus, dass die Wahrscheinlichkeit an Lungenkrebs zu erkranken linear mit der individuellen Exposition gegenüber Radon ansteigt (lineares Modell ohne Schwellenwert). Sowohl mit als auch ohne Korrektur für die mittlere Raucherrate wurde eine negative Beziehung zwischen der Lungenkrebsrate und der mittleren Radonexposition der Regionen beobachtet. Das heißt, dass das Risiko zu erkranken proportional zur mittleren Radonkonzentration der Region fiel. Dieses Ergebnis steht in auffälligem Gegensatz zu den Ergebnissen individueller Untersuchungen im Niedrig- und Hochdosisbereich (BEIR VI). Abbildung 1 zeigt Cohens ökologische Regression im Vergleich mit den Ergebnissen individueller Fall-Kontroll- und Kohortenstudien, sowie mit den von hochexponierten Minenarbeitern extrapolierten Risiken.

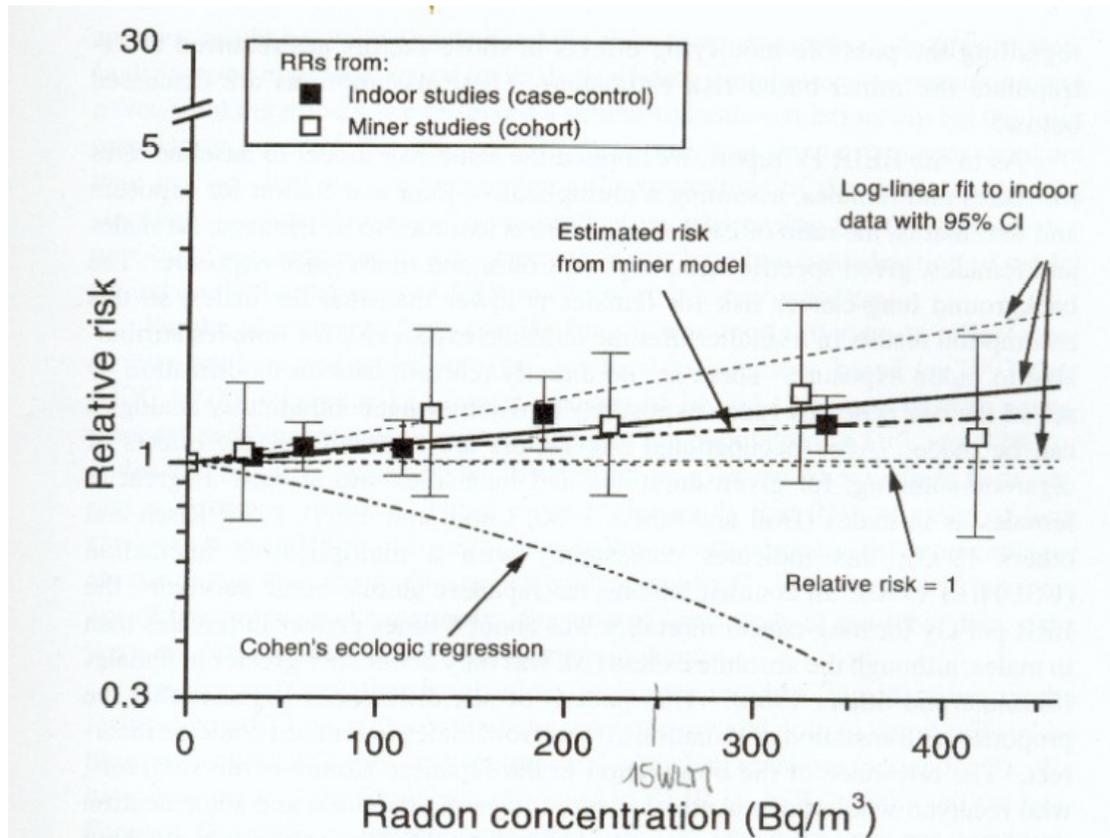


Abbildung 1: Zusammenfassung des relativen Risikos (RR) einer Metaanalyse für Radon in Wohnräumen, sowie gepoolte Analyse von Studien an Minenarbeitern, transformiert für den Dosisbereich <JHM-3, sowie Cohens ökologische Regression. (aus Beir IV 1999, Abb. 3-2, S.89)

Der Unterschied zwischen den individuellen Risikoschätzern und der ökologischen Studie ist sehr auffällig. Confounding auf Gruppenebene durch über 54 sozioökonomische Faktoren, sowie durch geographische Faktoren, Höhe und Klima, kann nur einen kleinen Teil dieser Abweichung erklären. Bekannte Korrelationen zwischen Rauchen und Radon, verbunden mit dem Urbanisierungsgrad der Region, wurden in die Berechnung mit einbezogen und zeigten keine Effekte. Der Autor schließt einen ökologischen Bias als Erklärung seines Ergebnisses aus. Er betrachtet sein Ergebnis und die Diskrepanz zu den Risikoschätzern auf individueller Ebene als Beleg für das Versagen des traditionell angewandten Modells einer linearen Dosis-Wirkungsbeziehung ohne Schwellenwert für die Beziehung zwischen Radon und Lungenkrebs (Cohen 1995).

Cohen (1994) argumentiert, dass der Einfluss eines Risikofaktors auf individueller Ebene in der Epidemiologie im Mittelpunkt des Interesses steht und dass ein nachgewiesenes Risiko auf individueller Ebene unkritisch auf die physiologische Ebene übertragen wird. Cohen zählt, ausgehend von seiner ökologischen Studie über Radon und Lungenkrebs, die Vorteile ökologischer Studien gegenüber individuellen Studien auf:

Aufgrund staatlicher Register sind Sekundärdaten für sehr viele, politisch, sozial oder regional begrenzte Bevölkerungsgruppen vorhanden. In einem ökologischen Studiendesign kann die Anzahl der Gruppen (in seinem Fall 1729), verglichen mit vielen individuellen Studien, groß sein.

Die statistische Unsicherheit in der Mortalitätsrate einer Region ist viel kleiner als die Unsicherheit, ob eine bestimmte Person mit einer gegebenen Exposition an Lungenkrebs sterben wird oder nicht.

Gute Daten sind als Sekundärdaten für viele Confounding-Faktoren vorhanden, weil diese für andere Zusammenhänge registriert werden. In individuellen Studien können dagegen nur wenige Confounder überprüft werden.

Die Frage der Repräsentativität der Studienstichprobe, die in individuellen Studien immer gestellt werden muss, ist für einen großen Teil der ökologischen Studien nicht relevant, weil dem Risikoschätzer die gesamte Population zugrunde liegt.

Cohen weist auf die Gefahr des ökologischen Trugschlusses hin, nachdem die mittlere Exposition einer Gruppe von Menschen nicht generell ihr mittleres Risiko für eine bestimmte Krankheit determiniert. (Ecologic fallacy: The average exposure to a group of people does not, in general, determine their average risk.) Nach Cohen (1994) gilt der ökologische Trugschluss dann nicht, wenn der Zusammenhang zwischen Exposition und Erkrankungswahrscheinlichkeit linear ist. Die Annahme ist aber nicht hinreichend, weil zusätzliche Annahmen über Confounder auch im linearen Modell notwendig sind. Weitere Bedingungen über die Unverzerrtheit des Schätzers auf Gruppenebene werden von Morgenstern (1998) zusammengestellt (siehe unten). Bisher gibt es aber keine systematische methodische Untersuchung über notwendige und hinreichende Bedingungen für die Unverzerrtheit des ökologischen Schätzers. Einige Aspekte werden wir weiter untersuchen.

Cohens Studie löste erneut eine öffentlich geführte Diskussion über die Vorteile und Risiken ökologischer Studien aus. In einer Reihe von 1991 bis 1996 veröffentlichten Artikeln stellen die Autoren ihre Kritikpunkte an ökologischen Studien dar und belegen sie mit Ergebnissen von Simulationsstudien oder analytischen Beweisführungen (Walter 1991, Brenner et al. 1992, Greenland und Robins 1994, Piantadosi 1994, Schwartz 1994, Stidley und Samet 1994, Susser 1994, Sheppard und Prentice 1995, Sheppard et al. 1996).

Lubin bezieht sich in zwei Arbeiten (Lubin 1998, 2002) direkt auf die cohensche Studie und hat mathematisch exakt das Potential für den ökologischen Bias aufgezeigt.

In der allgemeinen Beurteilung ökologischer Studien haben sich vier Positionen herausgebildet:

Cohen (1994) hält ökologische Studien unter bestimmten Bedingungen auch für Schätzer auf individueller Ebene für besser und valider als individuelle Studien.

Greenland und Robin (1994) und viele andere Autoren vertreten die Ansicht: Ökologische Studien sind nützlich, mit ihren Ergebnissen muss aber vorsichtig und kritisch umgegangen werden.

Für Schwartz (1994) sind ökologische Studien ausschließlich geeignet, um strukturelle, kontextuelle und soziologische Effekte zu untersuchen.

Piantadosi (1994) hält ökologische Studien für ungeeignet, um Risikofaktoren zu untersuchen. Ihre Aufgabe sieht er ausschließlich in der Generierung von Hypothesen, die dann in individuellen Studien weiter untersucht werden müssen.

Diese und weitere Veröffentlichungen demonstrieren auch die unterschiedlichen Ebenen der Diskussion um ökologische Studien. Greenland und Robin sowie Piantadosi diskutieren auf einer mathematischen Ebene, während Cohen und Schwartz die hinter individuellen und ökologischen Variablen stehenden Konstrukte qualitativ darstellen.

2.2.5 Methodische Aspekte in ökologischen Studien

Der Fokus der Diskussion liegt auf der Rolle möglicher Confounder in individuellen und ökologischen Studien. Wir werden im Folgenden die wichtigsten Aspekte dieser Diskussion in Anlehnung an eine Übersichtsarbeit von Morgenstern (1998) zusammenfassen.

2.2.5.1 Konzepte in ökologischen Studien

Ökologische Studien nutzen in verschiedenen Bereichen unterschiedliche Konzepte. Unterschiede bestehen auf der Variablebene (level of measurement), der Analyseebene (level of analysis) und der Interpretationsebene (level of inference).

Variablebene

Im Gegensatz zu individuellen Variablen, die Eigenschaften von Individuen darstellen, sind ökologische Variablen die Eigenschaften von Gruppen, Organisationen oder Plätzen. Drei Typen ökologischer Variablen lassen sich unterscheiden:

Aggregierte Messungen sind:

a) Zusammenfassungen individueller Messwerte innerhalb einer Gruppe.

Im Idealfall handelt es sich um den Durchschnittswert individueller Messwerte aller Gruppenmitglieder. In der praktischen Anwendung liegen meistens nur Stichproben von individuellen Messwerten vor, aus denen der mittlere Wert der Gruppe geschätzt wird. Beispiele sind der durchschnittliche Blutdruck oder der Anteil Raucher einer Schulklasse. Hier wird unter aggregierten Messungen sowohl Durchschnittswerte aus stetigen individuellen Variablen, als auch Raten von binären Variablen verstanden. Zwischen diesen beiden Typen aggregierter Messungen besteht aber ein Unterschied. Bei der Aggregation stetiger oder ordinalskaliert Variablen zu Mittelwerten geht die Information über die Variation dieser Messwerte innerhalb des Clusters verloren. Bei der Zusammenfassung individueller binärer Variablen zu Raten findet kein Informationsverlust statt. In unseren Studien nach Tschernobyl ist die Zielvariable die Inzidenzrate einer auf individueller Ebene binären Variablen (Schilddrüsenkrebs: ja/nein).

b) Messungen von Umweltfaktoren

sind physikalische Eigenschaften der Umgebung, der die Mitglieder einer Gruppe ausgesetzt sind. Beispiele sind UV-Index, Staubbelastung oder Radonbelastung. Für diese Variablen existieren zwar Werte auf individuellem Level, die aber nicht gemessen wurden. Die individuelle Variation hängt sowohl von der regionalen Variabilität als auch den Lebensgewohnheiten der Individuen ab. Die Expositionsvariable „Strahlendosis“ unserer Studien zur Schilddrüse ist ein Umweltfaktor, weil sie anhand eines radioökologischen Modells aus radioaktivem Cäsium berechnet wird. Für Cluster, für die individuelle Messungen vorliegen, erfolgt dann allerdings eine Anpassung der errechneten Werte an die gemessenen. Die Exposition ist damit eine Mischung aus einem Umweltfaktor und einer aggregierten Variablen.

c) Strukturelle Messungen

sind Eigenschaften der Gruppe, für die es keine analogen Werte auf individueller Ebene gibt. Beispiele sind Populationsdichte, Existenz eines Früherkennungsprogramms oder der Typ des Gesundheitssystems. Für unsere Studien zum Schilddrüsenkrebs nach Tschernobyl ist das Vorhandensein eines Screeningprogramms eine strukturelle Variable.

Analyseebene

Für eine statistische Analyse werden die Variablen auf eine gemeinsame Ebene gebracht. Dabei unterscheidet man

a) Analyse auf individueller Ebene

Alle Variablen liegen auf individueller Ebene vor, die gemeinsame Verteilung von Prediktor- und Zielvariablen, sowie möglichen Störgrößen ist bekannt.

b) Vollständig ökologischen Analyse

Alle Variablen liegen nur auf Gruppenebene vor, die gemeinsame Verteilung von Prediktor- und Outcomevariable, sowie möglichen Störgrößen ist nicht bekannt.

c) Teilweise ökologische Analyse

Bei einer Analyse von drei oder mehr Variablen ist die gemeinsame Verteilung innerhalb der Gruppen von einigen Variablen bekannt. Z.B. kennt man die individuellen Daten von Alter, Geschlecht und Krankheitsstatus, aber nicht die individuellen Messwerte einer Exposition.

d) Multilevelanalyse

Die Multilevelanalyse berücksichtigt explizit Daten sowohl auf individueller Ebene als auch auf Gruppenebene. Z.B. werden innerhalb mehrerer Studienregionen individuelle Daten erhoben und deren Risikoschätzer später auf Gruppenebene erneut ausgewertet. In diesem Review wird auf die Multilevelanalyse nicht näher eingegangen. Obwohl diese Methode sehr viel versprechend ist, kann sie ebenso wie Fall-Kontroll- und Kohortenstudien nicht angewandt werden, wenn individuelle Daten fehlen. Eine Übersicht über den Stand der Multilevelanalyse gibt Greenland (2002).

Interpretationsebene

Auch die Interpretation eines Studienergebnisses, z.B. die Deutung eines relativen Risikos kann auf unterschiedlichen Ebenen erfolgen.

a) Biologische Interpretation

Die Deutung eines Ergebnisses auf individueller Ebene wird biologische Interpretation genannt. In der Epidemiologie werden in der Regel Auswirkungen von Expositionen auf das individuelle Risiko geschätzt. Etwas anderes ist die biologische Interpretation eines kontextuellen Faktors zu sehen. Z. B. kann der mittlere sozioökonomische Status einer Wohngegend unabhängig vom individuellen sozioökonomischen Status einer Person ein Einflussfaktor für eine Infektionskrankheit sein.

b) Ökologische Interpretation

Ökologische Interpretation ist die Interpretation eines Ergebnisses auf Gruppenebene, z.B. kann der Einfluss der Klinikdichte einer Region auf die altersspezifische Herz-

Kreislaufmortalitätsrate geschätzt werden. Das Risiko ist dann eine Aussage über die Gruppe und lässt sich nicht als individuelles Risiko deuten.

c) Cross-level inference

Häufig werden Daten zwar auf Gruppenebene erhoben und analysiert, die Interpretation wird aber auf individueller Ebene durchgeführt. In der Untersuchung von Cohen (1995) wurde der Einfluss der mittleren Radonkonzentration einer Region auf die Lungenkrebsrate der Population untersucht. Alle Daten lagen auf Gruppenebene vor, aus dem Ergebnis schloss der Autor aber, dass Radon ein protektiver Faktor für das individuelle Lungenkrebsrisiko sei. Eine solche Übertragung der Ergebnisse von einer Ebene zur anderen ist Gegenstand der meisten Kritiken an ökologischen Studien.

2.2.5.2 Vorteile ökologischer Studien

Trotz ihrer methodischen Schwächen werden ökologische Studien aus pragmatischen Gründen häufig durchgeführt. Dafür gibt es verschiedene Gründe:

Sie sind günstig und schnell durchzuführen.

Ökologische Studien können häufig mit Sekundärdaten durchgeführt werden, z.B. können Informationen aus Krebsregistern mit Informationen statistischer Erhebungen an denselben Bevölkerungsgruppen in Verbindung gebracht werden.

Individuelle Daten fehlen.

Besonders im Bereich der Umweltepidemiologie ist es oft nicht möglich, individuelle Expositionen zu erheben. Luftschadstoffe können z.B. leicht über regionale Messstationen gemessen werden, eine individuelle Erhebung der eingeatmeten Schadstoffe ist äußerst aufwendig. Trotzdem kann sich die individuelle Exposition von den regionalen Messwerten deutlich unterscheiden.

Die Varianz innerhalb der Gruppen ist klein, zwischen verschiedenen Gruppen groß.

Viele Umweltfaktoren, z.B. UV-Index, Strahlendosis, mittlere Temperatur, aber auch Lebens- und Ernährungsgewohnheit unterscheiden sich innerhalb einer Bevölkerung nur wenig, zwischen verschiedenen, räumlich getrennten Populationen aber stark. In diesem Fall ist die Power individueller Studien innerhalb einer Population eingeschränkt, weil für den Nachweis signifikanter kleiner Risikounterschiede bedingt durch geringe Expositionsunterschiede größere Studienpopulationen benötigt werden, als für den Nachweis großer Risikounterschiede.

Fokus des Interesses sind Effekte auf Gruppenebene.

Ökologische Studien bieten sich an, wenn kontextuelle Einflüsse auf Erkrankungsraten untersucht werden sollen, z.B. für die Evaluation neuer Versorgungsstrukturen, oder der Einführung eines Früherkennungsprogramms.

Analyse und Präsentation ökologischer Studien sind einfach.

Besonders bei großen Surveys werden ökologische Analysen zur Kontrolle und Präsentation erster Ergebnisse eingesetzt.

2.2.5.3 Studiendesign

Das Design einer ökologischen Studie lässt sich in zwei Dimensionen klassifizieren:

Exploratives - versus analytisches Studiendesign

In vielen ökologischen Studien geht es darum, Unterschiede in Krankheitsraten zwischen verschiedenen Gruppen darzustellen. Eine spezifische Exposition wird nicht gemessen. In diesem Fall spricht man von einem explorativen Studiendesign. Wenn der Einfluss einer oder mehrerer Expositionen auf eine Outcomevariable untersucht werden soll, spricht man von einem analytischen Design.

Studiendesigns nach Einteilung der Gruppen

Multiple-Group-Design (Mehrgruppenvergleich, häufig räumliche Korrelation)

In einem multiple-group-design werden mehrere Gruppen von Personen zum selben Zeitpunkt verglichen. Die Gruppen werden meistens geographisch definiert. Es werden die Populationen von Ortschaften, Kreisen oder Ländern miteinander verglichen. Die Mitglieder der Gruppen können aber auch im selben Raum leben, sich aber anhand von anderen Merkmalen, z.B. Geschlecht, Alter oder Rasse voneinander unterscheiden

Time-Trend Design (zeitliche Korrelation)

In einem time-trend-design werden Populationen derselben Region zu unterschiedlichen Zeitpunkten verglichen. Z.B. wird der Grad der Umweltbelastung zu unterschiedlichen Zeitpunkten mit Veränderungen von Krankheitsraten in Verbindung gebracht.

Mixed Designs

Das Mixed Design stellt eine Kombination aus Multiple-Group und Time-Trend Design dar. Zeitliche Trends von Expositionen und Krankheitsraten werden parallel in mehreren geographischen Gruppen verfolgt, die Auswertung kann sowohl zwischen den Gruppen als auch zeitlich erfolgen. Auf diesem Prinzip beruht die Gesundheitsberichtserstattung (z.B. Krebsatlas, Analysen mit Daten des Krebsregisters).

Effekt-Schätzer in ökologischen Studien

Es wurde bereits dargestellt, mit welchen Modellen Effekt-Schätzer in individuellen und in ökologischen Studien durchgeführt werden. Ebenso wie in individuellen Studien können auch in ökologischen Studien weitere Einflussfaktoren für die Krankheit den Risikoschätzer verzerren, wenn sie mit der Exposition korreliert sind. Zur Kontrolle von Confoundern werden in ökologischen Analysen zwei Methoden benutzt.

Confounder auf Gruppenebene können als additive oder multiplikative Kovariablen in das Regressions-Modell mit aufgenommen werden. Damit kann, unter bestimmten Annahmen, ein für den Confounder adjustiertes relatives Risiko geschätzt werden.

Die zweite Methode zur Confounderkontrolle in einer ökologischen Analyse besteht darin, die Erkrankungsraten im Vorhinein für die Confounder zu standardisieren und die Regression mit den standardisierten Raten als Zielvariable durchzuführen. Dafür müssen aber die gemeinsamen individuellen Verteilungen der Erkrankungsraten und Confounderraten, sowie die gemeinsamen Verteilungen der Expositionsvariablen und der Confoundervariablen bekannt sein. Dies ist in ökologischen Studien selten gegeben.

2.2.5.4 Methodische Probleme in ökologischen Studien

Der ökologische Bias

Ein ökologischer Bias tritt auf, wenn ein Risikoschätzer in einer ökologischen Analyse dem korrespondierenden Risikoschätzer auf individueller Ebene nicht entspricht. Ursache für den ökologischen Bias ist die Abhängigkeit der Verteilungen der Expositions- und der Kovariablen. Die gemeinsame Verteilung der Variablen ist in ökologischen Studien nicht bekannt und die aggregierten Daten lassen darüber keine Aussage zu. Im englischen Sprachraum hat sich die Bezeichnung „ecologic fallacy“ eingebürgert, im Deutschen spricht man von dem ökologischen Trugschluss.

Ein hypothetisches Beispiel von Szklo & Nieto (2000) verdeutlicht das Problem.

Abbildung 2: Schematische Darstellung einer hypothetischen Studie, in der ein Ecologic Bias auftritt. Jedes Kästchen stellt ein hypothetisches Individuum dar, wobei Kästchen mit fettem Rahmen Individuen mit inzidenten Verkehrsunfällen repräsentieren. Die Ziffern in den Kästchen stellen das Jahreseinkommen der jeweiligen Individuen dar. Quelle: Szklo & Nieto (2000).

Population A: Durchschnittseinkommen US \$ 23.940; Inzidenz der Verkehrsunfälle 4/7=57%						
10,5	34,5	28,5	12,2	45,6	17,5	19,8
Population B: Durchschnittseinkommen US \$ 22.430; Inzidenz der Verkehrsunfälle 3/7=43%						
12,5	32,5	24,3	10,0	14,3	38,0	26,4
Population C: Durchschnittseinkommen US \$ 21.410; Inzidenz der Verkehrsunfälle 2/7=29%						
28,7	30,2	13,5	23,5	10,8	22,7	20,5

In Abbildung 2 wird der Zusammenhang zwischen (Jahres-)einkommen und Verkehrsunfällen in drei Gruppen mit jeweils sieben Individuen dargestellt. In einer rein ökologischen Analyse besteht eine positive Beziehung zwischen Durchschnittseinkommen und der Unfallrate einer Gruppe. Population A weist das höchste Durchschnittseinkommen und die höchste Inzidenz der Verkehrsunfälle auf, gefolgt von Population B und Population C. Bei der Betrachtung der gemeinsamen Verteilung von Einkommen und Verkehrsunfällen auf individueller Ebene wird aber deutlich, dass ausschließlich Individuen mit einem Jahreseinkommen unter US \$ 20.000 Unfälle haben. Fasst man alle Individuen der drei Populationen zusammen, so ergibt sich für diejenigen mit Verkehrsunfall ein Durchschnittseinkommen von US \$ 13.230, für diejenigen ohne Verkehrsunfall von US \$ 32.310. Auf individueller Ebene besteht also eine negative Beziehung zwischen Jahreseinkommen und dem Unfallrisiko. Eine Erklärung für dieses Phänomen bietet eine unterschiedliche Konstruktvalidität der Exposition auf den beiden Analyseebenen (siehe Abschnitt 2.2.6).

Für das Auftreten des ökologischen Bias gibt es unterschiedliche Gründe:

Bias innerhalb der Gruppe: Verzerrungen der Beziehung zwischen Exposition und Erkrankung innerhalb der Gruppen können zu einer Verzerrung des ökologischen Schätzers führen. Solche Verzerrungen können wie in individuellen Studien durch Confounding, Selektionsmethoden oder Missklassifikation entstehen.

Confounding auf Gruppenebene: Die Grundinzidenzrate einer Erkrankung in der nichtexponierten Bevölkerung kann sich zwischen den Gruppen der ökologischen Analyse unterscheiden. Wenn diese unterschiedlichen Grunderkrankungsraten mit der Höhe der Exposition korrelieren, kommt es zu einer Verzerrung des ökologischen Schätzers, auch wenn auf individueller Ebene kein Confounding vorliegt. Dies passiert besonders dann, wenn eine kontextuelle Confoundervariable die Erkrankungsrate beeinflusst, oder wenn eine Confoundervariable mit individueller Ausprägung sich von Cluster zu Cluster „sprunghaft“ verändert.

Effekt-Modifikation auf Gruppenebene: Zwischen den Gruppen einer ökologischen Studie können Variablen variieren, die die Auswirkung der Exposition auf die Erkrankung verändern. Wenn solche Variablen mit der Höhe der Exposition der Gruppen korreliert sind, kommt es ebenfalls zu verzerrten Schätzern.

Confounding und Effektmodifikation auf Gruppenebene bezeichnet man als Cross-Level Bias. Für das Auftreten gibt es drei Gründe:

- a) Es gibt Risikofaktoren für die Erkrankung, die zwischen den Gruppen unterschiedlich verteilt sind.
- b) Die Exposition wirkt sich auf Gruppenebene anders auf die Erkrankung aus als auf individueller Ebene. Im obigen Beispiel könnte sich das Durchschnittseinkommen einer Gruppe von Menschen in einem Wohngebiet auf die Verkehrsdichte auswirken. Geht man davon aus, dass Personen mit niedrigerem Einkommen weniger sichere Autos fahren, könnten genau diese Personen in Wohngebieten mit höherem mittleren Einkommen und daher höherer Verkehrsdichte gefährdeter sein als in Gegenden mit niedrigerem mittleren Einkommen.
- c) Die Erkrankungswahrscheinlichkeit einer Person ist abhängig von der Erkrankungsrate innerhalb der Gruppe. Diese Bedingung gilt für Infektionserkrankungen.

Liegen in einer Untersuchung ausschließlich aggregierte Daten vor, kann man diesen nicht entnehmen, ob ein Cross-Level-Bias vorliegt. Wie gut sich ein Modell an die ökologischen Daten anpassen lässt, gibt keine Auskunft darüber, ob die Schätzer verzerrt sind oder nicht.

Probleme der Confounder-Kontrolle

Risikoschätzer in ökologischen Studien können durch Confounder verzerrt sein. Die Bedingungen, die einen zusätzlichen Risikofaktor zu einem Confounder machen, unterscheiden sich zwischen den Ebenen. In der individuellen Analyse ist ein Confounder sowohl mit der Erkrankungswahrscheinlichkeit als auch mit der Exposition korreliert. In einer ökologischen Studie können Risikofaktoren, die auf individueller Ebene keine Beziehung zur Exposition aufweisen, zu Confoundern werden, wenn sie auf Gruppenebene mit der mittleren Exposition korrelieren. Ein Confounder auf individueller Ebene, der keine Korrelation zu den aggregierten Werten der Exposition aufweist, bewirkt keinen ökologischen Bias.

Confounder-Kontrolle in ökologischen Analysen ist komplexer als in individuellen Analysen. Weil die gemeinsame Verteilung von Exposition und Confounder auf individueller Ebene nicht bekannt ist, aber trotzdem zu einer Verzerrung des ökologischen Schätzers führen kann, führt ein Adjustieren für den Confounder nicht zwangsläufig zu einem geringeren Bias.

Greenland und Coautoren haben untersucht, unter welchen Bedingungen in ökologischen Studien kein Bias auftritt (Greenland und Morgenstern 1989, Greenland 1992, Greenland und Robins 1994). Die Ergebnisse lassen sich auf die folgenden Punkte zusammenfassen:

- a) Innerhalb der Gruppen besteht kein Confounding der individuellen Auswirkung einer Exposition durch nicht gemessene zusätzliche Risikofaktoren, Selektionsmethoden oder Missklassifikationen.

b) Auch auf ökologischer Ebene besteht kein unkontrolliertes Confounding zwischen den aggregierten Werten der Exposition und weiteren Risikofaktoren der Erkrankung auf Gruppenebene.

c) Die unterschiedlichen Risikofaktoren der Erkrankung sind additiv.

d) Die Zugehörigkeit zu einer bestimmten Gruppe hat keinen Einfluss auf den Effekt der einzelnen Prediktorvariablen auf individueller Ebene.

Wenn alle vier Bedingungen gleichzeitig erfüllt sind, tritt kein ökologischer Bias auf. Sind sie teilweise oder alle nicht erfüllt, besteht Gefahr für das Auftreten eines ökologischen Bias.

Ein weiterer Grund für einen ökologischen Bias ist Missklassifizierung der Confoundervariablen. Dieses Problem kann auch auf individueller Ebene auftreten, ist aber auf Gruppenebene schwerer zu kontrollieren, weil in der Regel nur der augenblickliche Zustand der Variablen bekannt ist. Der Einfluss vieler Kovariablen auf die Erkrankungsrate liegt aber weit zurück. Z.B. sind nicht die momentanen Rauchgewohnheiten einer Person für ihr Lungenkrebsrisiko relevant, sondern ihre Rauchgewohnheiten in der Vergangenheit. Während diese in individuellen Studien mittels Fragebogen retrospektiv ermittelt werden können, muss man sich in ökologischen Studien meistens auf Sekundärdaten beziehen, die für die Vergangenheit häufig nicht in ausreichender Qualität vorliegen.

2.2.5.5 Weitere Probleme ökologischer Studien

Adäquate Daten fehlen

Ökologische Studien basieren häufig auf Sekundärdaten. Informationen über die individuelle Krankheitsgeschichte fehlen meistens. Häufig sind diese Daten unvollständig, ungenau und nicht reliabel. Die Daten unterschiedlicher Register sind nicht direkt vergleichbar. Für viele Krankheiten liegen zwar Mortalitätsraten, aber keine Inzidenzraten vor. Mortalitätsraten sind aber nicht nur von den Risikofaktoren für die Erkrankung, sondern auch von den Einflussfaktoren für ihren Verlauf abhängig, die ihrerseits als Confounder den Risikoschätzer verzerren können. Häufige Fehlerquellen sind:

Erfassung der Krankheiten

Es können Probleme der Krebsregistrierung entstehen (Doppelmeldung, verschiedenes Meldeverhalten der Regionen, u.U. abhängig von der Exposition). Die Art der Registrierung kann sich über die Zeit verändert haben (ICD-Änderungen, Änderungen in der Klinik-Landschaft). Personen sind aus dem Erfassungsbereich fortgezogen. Weniger schwere Krankheiten werden mit geringerer Wahrscheinlichkeit gemeldet als schwere Fälle, so dass die Inzidenz- und Mortalitätsraten unterschiedlicher Qualität sind. Während diese nichtdifferenzielle Missklassifikation in individuellen Studien erlaubt, die Verzerrung des Schätzers noch abzuschätzen, geht dies in ökologischen Studien nicht mehr.

Erfassung der Population durch Censusdaten

Personen mit zweitem Wohnsitz, z.B. StudentInnen und Armeeangehörige sind meist im Census des ersten Wohnsitzes erfasst, wenn sie erkranken werden sie aber oft am zweiten Wohnsitz gemeldet. Censuserfassung und geographische Ortsgrenzen haben sich mit der Zeit geändert.

Erfassung der Exposition

Die Erfassung von Umweltdaten als Grunddaten für alle Individuen ist mit vielen Fehlerquellen behaftet. Dabei treten zwei getrennte Fehlerquellen auf: Zum einen wird der Umweltfaktor selber ungenau erfasst. Dabei spielen räumliche und zeitliche Veränderungen der Messqualität eine Rolle. Zum anderen bleibt offen, inwieweit der Umweltfaktor die individuellen Expositionen widerspiegelt. Z.B. sind Personen, die Flaschenwasser trinken von der Trinkwasserverschmutzung ihrer Region nicht betroffen.

Zeitliche Beziehung zwischen Exposition und Erkrankung

Während in Kohortenstudien sichergestellt ist, dass die Exposition einer Person vor ihrer Erkrankung lag, ist dies für ökologische Studien nicht nachweisbar. Wenn sich der Krankheitsstatus auf die Expositionswahrscheinlichkeit auswirkt, z.B. weil erkrankte Personen eine belastete Wohngegend verlassen, kann es in ökologischen Studien zu einer Missklassifikation der Exposition und möglicher weiterer Einflussvariablen kommen. Gerade in der Krebs epidemiologie, dem wichtigsten Feld der Strahlen epidemiologie, ist mit Latenzzeiten von 10 Jahren und mehr zu rechnen. Deswegen muss auf relativ alte Umweltmessungen zurückgegriffen werden, deren Qualität im Nachhinein oft nicht festzustellen ist.

Kolinearität

Auf Gruppenebene sind die Risikofaktoren für eine Erkrankung oft stärker korreliert als auf individueller Ebene. Deswegen ist es letztlich schwer einzuschätzen, welcher Faktor die Unterschiede in den Erkrankungsraten zwischen verschiedenen Gruppen bedingt. Dies Problem ist am größten, wenn in eine ökologische Studie wenige große, heterogene Regionen eingehen.

Migration zwischen Gruppen

Migration zwischen den Regionen einer ökologischen Studie führt zu Missklassifikation der Exposition, weil Personen in einer anderen Region erkranken, als sie exponiert waren. In ökologischen Studien ist dieses Problem mit der fehlenden Nachzeitigkeit von Exposition und Erkrankung verknüpft. Migration ist vor allem dann ein Problem, wenn lange Latenzzeiten zwischen Exposition und Erkrankung liegen. Durch Migration zwischen den Bezirken der Studie waren Menschen möglicherweise verschieden hoher Exposition ausgesetzt, ohne dass man dies nachvollziehen könnte. Migration ist häufig abhängig vom Gesundheitszustand, gesunde Menschen migrieren häufiger aus ökonomischen Gründen. Andererseits migrieren Menschen mit bestimmten Erkrankungen, z.B. Allergien, häufiger in Gegenden mit günstigem Klimaprofil. Es ist daher sinnvoll, relativ stabile Populationen für ökologische Studien auszuwählen. Allerdings ist die Übertragbarkeit der Ergebnisse dadurch eingeschränkt. Auf den Migrations-Bias in ökologischen Studien wird im folgenden Abschnitt näher eingegangen.

2.2.6 Weitere Diskussionspunkte zu ökologische Studien

Seit Morgensterns Review über ökologische Studien (1998) sind eine Reihe weiterer Artikel erschienen, die sich mit der Theorie ökologischer Studien und dem ökologischen Bias auseinandersetzen. Im Folgenden werden einzelne Kritikpunkte an ökologischen Studien herausgegriffen, die bei Morgenstern noch nicht vertiefend behandelt wurden, bzw. in der neueren Literatur weiterentwickelt wurden

Der ökologische Trugschluss als Validitätsproblem

Im Mittelpunkt der Diskussion über Schwächen ökologischer Studien stehen Fragen über den Einfluss von Confoundern und Effektmodifikatoren auf individueller und Gruppenebene. Während dieser Punkt bei Morgenstern ausgiebig behandelt wird, ist den Unterschieden in der Exposition zwischen individueller und ökologischer Ebene nur wenig Raum gewidmet. Im Folgenden sollen deswegen einige Kritikpunkte, die die Erfassung der Expositionsvariablen in ökologischen Studien betreffen, zusammengefasst werden.

Konstrukt-Validität

Auch bei Abwesenheit von Confoundern können Unterschiede zwischen den Risikoschätzern individueller und ökologischer Studien bestehen bleiben. Diese gehen auf unterschiedliche Konstruktvaliditäten der Expositionsvariablen zurück (Schwartz 1994). Der Begriff Konstrukt wird in der empirischen Sozialforschung verwendet. Ein Konstrukt bezeichnet eine Eigenschaft von Individuen, die keine direkt messbare Entsprechung hat. Konstrukte werden als Indexe mehrerer korrelierter Surrogatvariablen bestimmt. Z.B. wird in der Psychologie das Konstrukt Depressivität mithilfe von Fragen zur Ängstlichkeit, Antriebslosigkeit und Traurigkeit bestimmt.

Dieselbe Expositionsvariable kann auf ökologischer Ebene eine andere Konstruktvalidität aufweisen als auf individueller Ebene. Ein Beispiel dafür war die mögliche Erklärung für den ökologischen Trugschluss in dem von Szklo & Nieto (2000) konstruierten Beispiels einer ökologischen Studie (Abbildung 2).

Das *individuelle* Einkommen ist eine Teilvariable des Konstruktes „individueller Sozioökonomischer Status“, und korreliert als solche möglicherweise mit nicht erhobenen Variablen, z.B. dem Sicherheitsstandard des Autos und/oder dem Fahrstil. „Arme“ sind in diesem Beispiel grundsätzlich gefährdeter als „Reiche“.

Das *mittlere* Einkommen einer Wohngegend ist eine Teilvariable des Konstruktes „Sozioökonomischer Status der Wohngegend“ und korreliert möglicherweise mit dem (nicht erhobenen) Verkehrsaufkommen. In „reichen“ Wohngebenden fahren mehr Autos und das Risiko für Unfälle ist höher. Dieser Unterschied wird aber nur bei den individuell gefährdeteren „Armen“ sichtbar.

Für das tatsächliche Risiko der Einzelperson besteht eine Wechselwirkung zwischen der Variable auf individueller Ebene und derselben Variable auf Gruppenebene.

Ein weiteres bekanntes Beispiel ist das Suizidrisiko von Katholiken und Protestanten preußischer Provinzen zwischen 1883 und 1890 (Durkheim 1951).

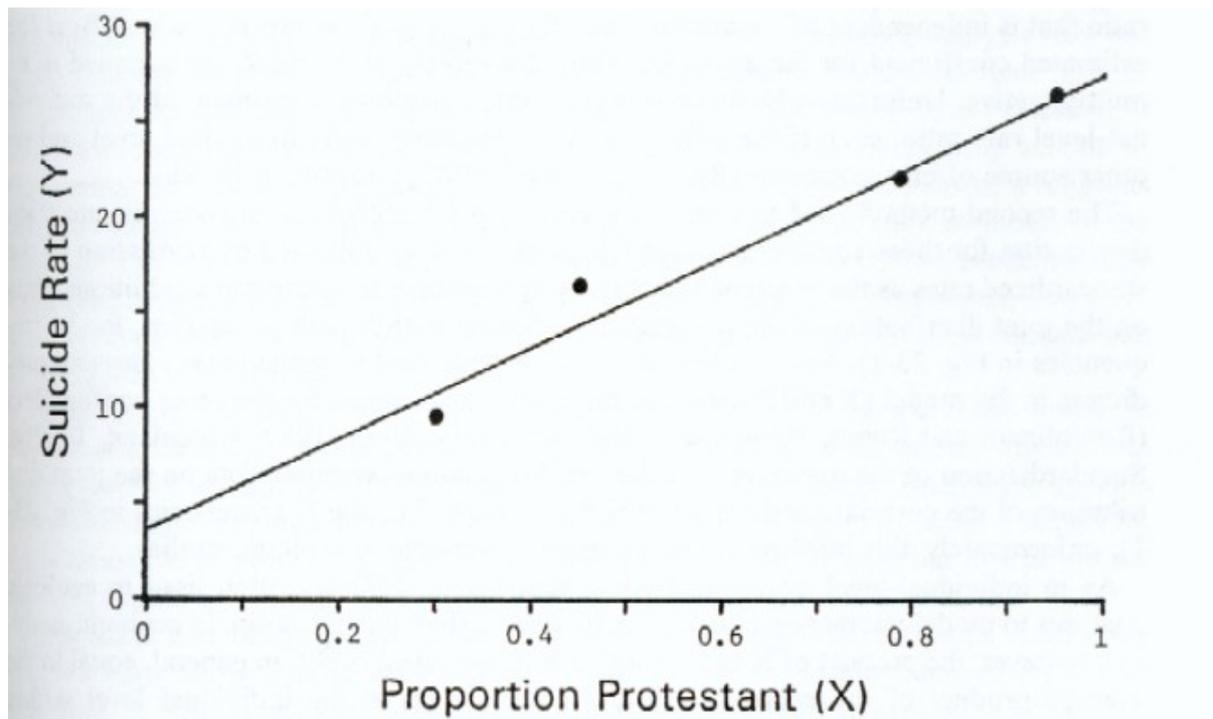


Abbildung 3: Suizidrate (Y, pro 10^5 Personenjahre) in Abhängigkeit von dem Protestantenanteil in preußischen Provinzen, 1883-1890. Aus: Durkheim E. *Suicide: a study in Sociology*. New York, Free Press, 1951, nach Morgenstern 1997).

Auf Gruppenebene besteht ein positiver linearer Zusammenhang zwischen dem Anteil an Protestanten und der Suizidrate der Region. Auf individueller Ebene waren es aber Katholiken, die die Suizide begingen, und zwar umso häufiger, je höher der Protestantenanteil in ihrer Region war. Auch hier ist eine Wechselwirkung zwischen der Exposition „Konfession“ auf individueller Ebene und der Exposition „Anteil der Menschen mit der eigenen Konfession an der Bevölkerung“ denkbar: Je vereinzelter die Katholiken in der Diaspora leben, desto größer ist ihr Suizidrisiko. Die individuelle Konfession ist Teil eines anderen Konstruktes als der Anteil der beiden Konfessionen auf Gruppenebene.

In der Diskussion über ökologische Studien wird diese Aufspaltung der Expositionsvariablen in zwei unterschiedliche Konstrukte auf individueller und ökologischer Ebene häufig übersehen. Der ökologische Bias kann in diesem Fall auch bei Abwesenheit oder vollständiger Kontrolle von Confoundern auftreten. Gleichzeitig kann eine Veränderung der Konstruktvalidität beim Übergang zwischen den Ebenen auch bei Confoundervariablen auftreten. Gelöst werden kann das Problem der Konstruktvalidität auf unterschiedlichen Ebenen nur durch die gleichzeitige Erfassung der Variablen auf individueller und auf Gruppenebene im Rahmen einer Multilevelanalyse (Greenland 2002).

In der neueren Literatur über den Einfluss sozioökonomischer Faktoren versucht man, diese unterschiedlichen Konstrukte derselben Expositions- oder Confoundervariable in einem Multileveldesign auch in die Analyse von Risikofaktoren auf individueller Ebene mit einzubeziehen (Diez-Roux 2000).

Zufällige Messfehler in ökologischen Studien

Zufällige Messfehler erzeugen in ökologischen Studien andere Verzerrungen als in individuellen Studien und können zur Entstehung eines ökologischen Bias beitragen. Brenner

et al. (1992) untersuchten den Effekt von nichtdifferenziellen Messfehlern in ökologischen Studien, in denen die Exposition auf individueller Ebene binär war und sich entsprechend auf Gruppenebene als Rate der exponierten Personen an der Gesamtpopulation darstellte. Ein nichtdifferenzieller Messfehler bedeutet, dass eine Fehlklassifikation des Expositionsstatus auf individueller Ebene jede Person der Population gleichermaßen betreffen kann und nicht mit dem Krankheitsstatus assoziiert ist.

Brenner zeigte, dass sich die Auswirkung nichtdifferenzieller Messfehler auf den Risikoschätzer einer binären Expositionsvariablen auf eine ebenfalls binäre Erkrankungsvariable im linearen Modell zwischen individuellen und ökologischen Studien deutlich unterscheidet. In ökologischen Studien bewirkt eine nichtdifferenzielle Missklassifikation grundsätzlich einen Bias, der zu einer Überschätzung des tatsächlichen Risikos führt. Damit unterscheidet sich der Bias in ökologischen Studien von dem in individuellen Studien, der grundsätzlich eine Abweichung in Richtung 0 bewirkt (Flegal et al. 1986). Dieser Unterschied trat auch in ökologischen Studien mit einem log-linearen Modell auf, die Verzerrungen sind aber weniger ausgeprägt als im linearen Risikomodell.

Brenners Analysen über die Auswirkung von Messfehler bezogen sich auf rein-ökologische Studien, in denen sowohl Exposition als auch Zielvariable als aggregierte Daten vorliegen. Daneben kommen Messfehler auch in teilökologischen Studien vor, in denen nur die Exposition auf Gruppenebene erfasst wird, Ziel- und Kovariablen aber auf individueller Ebene vorliegen. Diese Studien werden teilweise als individuelle Studien mit messfehlerbehafteten Expositionsdaten betrachtet (Künzli und Tager 1997, Webster 2000). Dagegen wies Webster (2002) für teilökologische Kohortenstudien, Björk und Strömberg (2002) für teilökologische Fall-Kontrollstudien nach, dass auch in diesen Studientypen unter den von Brenner (1992) formulierten Bedingungen nichtsystematische Messfehler zu einer Überschätzung des RR führen. Bias durch diese Messfehler können damit auch in teilökologischen Studien, in denen nur die Exposition auf Gruppenebene erfasst wird, auftreten.

Systematische (nicht-zufällige) Messfehler in ökologischen Studien

In ökologischen Studien liegen statt individuellen Messwerten für die Exposition aggregierte Daten vor. Exposition auf Gruppenlevel liegt nicht nur in rein-ökologischen Studien vor, sondern auch in so genannten partiell-ökologischen Fall-Kontroll- und Kohortenstudien, in denen der Erkrankungsstatus und weitere Kovariablen individuell erfasst werden, die Exposition aber aus der Gruppenzugehörigkeit der Person geschlossen wird. Zu diesem Typ gehören Studien, die eine Job-Exposure-Matrix benutzen, um die Exposition eines bestimmten Arbeitsplatzes abzuschätzen. Wenn Expositionen auf Gruppenebene genutzt werden, die aus großen Stichproben mit mehr als 100 individuellen Messungen aggregiert wurden, ist der Bias des Risikoschätzers, verursacht durch nichtdifferenzielle Messfehler auf individueller Ebene generell vernachlässigbar (Prentice und Sheppard 1995). Dagegen können systematische Messfehler auf Gruppenebene größere Bias verursachen, als auf individueller Ebene. Björk und Strömberg (2002) untersuchten den Effekt von systematischen Messfehlern in teilökologischen Fall-Kontroll-Studien, in denen Krankheitsstatus, Gruppenzugehörigkeit und Kovariablen auf individueller Ebene vorlagen, die binäre Exposition aber auf ökologischer Ebene. Unter Annahme eines linearen Risikomodells wurden in Simulationsstudien und anhand empirischer Daten die Auswirkung von systematischen Messfehlern auf das Odds-Ratio und das 95%-CI beobachtet.

Systematische Messfehler der Exposition erzeugten in der teilökologischen Studie Abweichungen der Steigung und des Intercepts der Modellgeraden, die den Abweichungen in rein individuellen Fall-Kontroll-Studien mit entsprechenden Messfehlern entsprachen. Für beide Studientypen konnten Verzerrungen in beide Richtungen entstehen. Messfehler, die auf ökologischer Ebene zu einer systematischen Verzerrung aller Gruppenmittelwerte führten, erzeugten größere Bias, die sowohl ein größeres als auch ein kleineres Risiko vortäuschten. Messfehler, die zu einer zufälligen Fluktuation der aggregierten Werte führten, erzeugten dagegen kleinere Verzerrungen hin zu einem geringeren Risikoschätzer.

Auch der weiter unten besprochene Migration-Bias kann als Spezialform der Missklassifikation betrachtet werden. Seine Auswirkungen auf ökologische Risikoschätzer muss deswegen weiter untersucht werden.

Varianz der Exposition

Sheppard und Prentice (1995) zeigten anhand einer Simulationsstudie die Bedeutung der Varianz der Exposition innerhalb und zwischen den Gruppen für den Risikoschätzer von individuellen und ökologischen Studien. Die Simulationen lehnten sich an eine ökologische Studie zu mittlerem Fettverzehr und Brustkrebs in 21 Ländern an (Prentice und Sheppard 1990). Dem Modell liegt eine exponentielle Beziehung zwischen Exposition und Zielvariablen zugrunde.

Das Verhältnis Varianz der Kovariablen zwischen den Gruppen zu der Varianz der Kovariablen innerhalb der Gruppen hat einen großen Einfluss auf die Effektivität ökologischer Schätzer im Vergleich zu individuellen Schätzern innerhalb der Gruppen. Im extremen Fall, wenn die Intragruppen-Variabilität 0 ist, kann keine sinnvolle Risikoschätzung innerhalb einer Gruppe durchgeführt werden. Ist die Intergruppen-Variabilität 0, kann kein Schätzer mit aggregierten Daten entstehen. Die Simulationen wurden mit unterschiedlichen Verhältnissen zwischen der Varianz innerhalb der Gruppen und der Varianz zwischen den Gruppen durchgeführt.

Tabelle 4 Mittelwerte aus 5000 Simulationen der Risiko-Schätzer und Standardabweichungen in individuellen und ökologischen Analysen. Der wahre Betawert betrug $2,937 \times 10^{-3}$. Die Ergebnisse basieren auf 21 Gruppen mit 10.000 Personen. (nach Sheppard und Prentice 1995, Tab.1, verändert)

Analyse	Cofaktor-Varianz-Verhältnis	Risikoschätzer $\beta \times 10^{-3}$	Standardabweichung $\beta \times 10^{-3}$
Mittlerer individueller Schätzer innerhalb der Gruppen	0,28	2,933	0,183
	0,55	2,933	0,291
	1,25	2,919	0,505
	2,65	2,925	0,804
	11,1	2,947	1,74
Ökologische Analyse	0,28	2,933	0,426
	0,55	2,027	0,510
	1,25	2,929	0,613
	2,65	2,936	0,675
	11,1	2,943	0,717

Die Risikoschätzer sowohl der individuellen als auch der ökologischen Analyse zeigen wenige Verzerrungen, alle liegen innerhalb von einem Prozent des wahren Wertes. Die Standardabweichungen der Risikoschätzer sind in der individuellen Analyse innerhalb der Gruppen kleiner als die Standardabweichungen der ökologischen Analysen, solange das Verhältnis der Varianz der Exposition zwischen den Gruppen zu der Varianz innerhalb der

Gruppen klein ist. Bei steigendem Verhältnis sind die Abweichungen der ökologischen Analyse kleiner als die der individuellen Analyse.

Wenn die Variabilität der Exposition innerhalb der Gruppen groß ist im Verhältnis zu der Variabilität zwischen den Gruppen, ist die individuelle Analyse innerhalb der Gruppen effektiver als die ökologische Analyse, deren Risikoschätzer deutlich größere Konfidenzintervalle aufweisen. Ist aber die Variabilität zwischen den Gruppen groß im Verhältnis zu der Variabilität innerhalb der Gruppen, dann sind ökologische Analysen effizienter als individuelle Analysen innerhalb der Gruppen.

Dieses Ergebnis der Simulationsstudie wird in ökologischen Studien häufig zitiert und als Motivation für das ökologische Design angegeben (z.B. Largade und Pershagen 1999, Darby et al. 2001). Es wurde aber nur ein Modell ohne Confounder und Messfehler untersucht. Inwieweit sich das Ergebnis auf andere Modelle und Bedingungen übertragen lässt, ist bisher nicht untersucht worden.

Migration-Bias

In seiner Veröffentlichung untersucht Tong (2000) eingehend mögliche Verzerrungen, die durch die Migration von Personen zwischen den Regionen entsteht, die als Gruppen in eine ökologische Analyse eingehen. „Migration bias occurs when there is differential factor-related migration within, into or out of each study population group.“ Migration-Bias ist auch ein Problem individueller Studiendesigns, in ökologischen Studien wird er aber leicht übersehen, weil diese in der Regel Daten nur zu einem bestimmten Zeitpunkt erheben.

Der Migration Bias beeinflusst ökologische Studien besonders bei langen Latenzzeiten zwischen Exposition und Erkrankung. Wenn Personen in einer Region exponiert waren und später in einer anderen Studienregion erkranken, wird ihnen eine falsche Exposition zugeordnet. Die Wahrscheinlichkeit einer Migration ist in der Regel abhängig von Alter, Geschlecht und anderen sozialen Faktoren, aber auch von der Gruppengröße. Ein Migration-Bias in einer ökologischen Studie ist dann zu erwarten, wenn Personen, die später erkranken, mit einer größeren Wahrscheinlichkeit migrieren als Personen, die nicht erkranken werden. Schon ein kleiner Anteil eines solchen differenzierten Migrations-Bias kann ein signifikantes Risiko in einer ökologischen Studie überdecken oder ein nicht vorhandenes Risiko vortäuschen. Der Migration Bias tritt unabhängig von weiteren Confoundern oder Effektmodifikationen auf Gruppenebene auf.

Zur Vermeidung des Migration-Bias führen folgende Strategien:

Die Auswahl einer stabilen Population. Insgesamt ereignet sich Migration seltener in ländlichen Gebieten als in städtischen, seltener in kleinen Städten als in Großstädten und seltener in Entwicklungsländern als in entwickelten Ländern. Auch Geschlecht, Alter und sozialer Status können mit Migration in Verbindung stehen.

Die Auswahl großer Gruppen (z.B. Länder statt Bundesländer) führt zu geringerer Migration zwischen den Gruppen, kann aber auf der anderen Seite durch größere Variabilität der Exposition und der Kovariablen kontraproduktiv sein.

Die Auswahl kleiner Zeiteinheiten. Das Sammeln aller zugänglichen Informationen über Häufigkeit von Migration in unterschiedlichen Regionen und Zeitabschnitten. Auch aus begrenzten Informationen kann auf die verschiedenen Regionen hochgerechnet werden.

Die Untersuchung von Faktoren, die mit der Migration korrelieren. Handelt es sich dabei um Risikofaktoren für die betreffende Erkrankung, steigt die Gefahr eines Migrations-Bias.

Der Migrations-Bias tritt auch in individuellen Studien auf, in ökologischen Studien hat er aber eine besondere Bedeutung, weil häufig große Populationen betrachtet werden, in denen Ex- und Emigration eine wichtige Rolle spielen. Besonders bei langandauernden Expositionen und langen Latenzzeiten sollten möglichst viele Informationen über die Migration zwischen den Gruppen der ökologischen Studie, sowie in die - und aus der gesamten Studienpopulation eingeholt werden und als Kofaktoren in die Analysen mit eingehen.

2.2.7 Vergleichende Analysen auf individueller und ökologischer Ebene

In den bisher vorgestellten Untersuchungen wurde eine Reihe isolierter Gründe für das Auftreten eines ökologischen Bias theoretisch analysiert. In wieweit diese Gründe in realistischen Untersuchungen tatsächlich vorliegen und in welchem Maße sie die Risikoschätzer beeinflussen, ist in den meisten Fällen nicht bekannt (Webster 2002). In realistischen Studien liegen die Einflussfaktoren für den Risikoschätzer nicht getrennt vor, sondern in Wechselwirkung mit einer Reihe weiterer Faktoren, die zum Teil nicht erfasst worden sind. Untersuchungen großer Datensätze individueller Studien, die zusätzlich auf Gruppenebene ausgewertet werden können, bieten eine Möglichkeit, das Auftreten ökologischer Bias auch praktisch zu beobachten und die Gründe dieser Verzerrungen zu untersuchen. Im Folgenden werden zwei neuere Untersuchungen zu Radon und Lungenkrebs dargestellt, in denen parallele Auswertungen sowohl auf individueller Ebene als auch auf Gruppenebene durchgeführt wurden.

Einleitung zu beiden Studien

Arbeiter im Uran-Bergbau, die hohen Dosen radioaktiven Radons ausgesetzt sind, haben ein erhöhtes Risiko, an Lungenkrebs zu erkranken. Radon wurde deswegen als menschliches Karzinogen klassifiziert. Die Exposition durch Radon in der Innenraumluft ist in vielen Ländern die größte Quelle von natürlichen ionisierenden Strahlen. Überträgt man die an Bergarbeitern gewonnenen Risikoschätzer zu Radon und Lungenkrebs auf die Innenluftbelastung, scheint Radon nach Rauchen der zweitwichtigste Risikofaktor für Lungenkrebs zu sein. Nach diesen Berechnungen wäre Radon in Großbritannien die Ursache für 2000 Todesfälle im Jahr (Darby et al 2001). Nachdem eine ökologische Studie von Cohen et al. (1990, 1995) eine streng negative Beziehung zwischen der durchschnittlichen Radonexposition und der Inzidenz für Lungenkrebs erbracht hatte, obwohl individuelle Untersuchungen sowohl im Niedrigdosisbereich als auch bei hoch exponierten Minenarbeitern eine positive Beziehung zeigten, stand die Beziehung zwischen Radon und Lungenkrebs im Mittelpunkt der Diskussion über den ökologischen Bias. In den letzten Jahren sind zwei Studien durchgeführt worden, in denen die Daten großer Fall-Kontroll-Studien zusätzlich zu der individuellen Analyse auch auf Gruppenebene ausgewertet wurden. Dadurch ist ein direkter Vergleich der Ergebnisse auf individueller - und auf Gruppenebene unter realistischen Bedingungen möglich.

Die beiden Studien wurden in Schweden von Lagarde und Pershagen, in Süd-West-England von Darby, Deo und Doll durchgeführt und 2001 veröffentlicht. Im Folgenden werden Methoden und Ergebnisse der beiden Studien zusammengefasst.

Material und Methoden der schwedischen Studie

Für die Fall-Kontrollstudie wurden zwischen 1980 und 1984 1360 Lungenkrebsfälle und 2847 Bevölkerungskontrollen erhoben. Das Alter der Fälle und Kontrollen lag zwischen 35 und 74 Jahren. Alle Studienteilnehmer lebten seit 1947 in Schweden und waren in einer von 109 Gemeinden gemeldet. Die Gemeinden repräsentierten Gegenden mit hoher oder niedriger Radonkonzentration in Wohnräumen. Jedem Fall wurden zwei altersgematchte Kontrollen zugeordnet. Die Radonkonzentration in den Wohnungen von Fällen und Kontrollen wurde während der Heizperiode in jeweils zwei Räumen von insgesamt 8992 Wohnungen gemessen. Informationen über den Raucherstatus und weitere Risikofaktoren für Lungenkrebs wurden durch Fragebogenerhebungen von den Studienteilnehmern oder nahen Verwandten erhoben.

Die individuelle mittlere Radonkonzentration wurde als gewichtete Summe aus den Messwerten der Wohnungen berechnet. Berücksichtigt wurde die relative Wohndauer im Verhältnis zur Gesamtbeobachtungszeit einer Person mit erreichbaren Messwerten. Für die ökologische Analyse wurden die so errechneten individuellen Expositionsdaten der Kontrollen aggregiert, um die mittlere Radonkonzentration einer Gemeinde zu erhalten. Als Heimatgemeinde eines Studienteilnehmers wurde die Gemeinde definiert, in der er innerhalb des Zeitfensters von 1947 bis 1984 die längste Zeit gemeldet war. Für die anderen Risikofaktoren Stadt/Land, Geschlecht und Alter wurde keine Aggregation vorgenommen. Der Raucherstatus wurde auf individueller Ebene über einen zeitgewichteten Index bestimmt, auf Gruppenebene wurden die geschlechtsspezifischen Mittelwerte berechnet. In einigen Analysen wurde der Breitengrad als weiterer Kofaktor berücksichtigt. Dabei wurde der Breitengrad der größten Siedlung der Gemeinde gewählt.

Es wurde eine konditionale, lineare Regression durchgeführt und das zusätzliche relative Risiko für verschiedene Ebenen der Stratifizierung berechnet. Likelihood-basierte 95%-Konfidenzintervalle wurden zur Überprüfung der Signifikanz herangezogen.

Die Ergebnisse der ökologischen Analyse können verzerrt sein, weil das Modell nur für Situationen angepasst ist, in denen die Radon-Exposition und der Raucherstatus innerhalb der Gruppen unabhängig ist oder eine additive Beziehung vorliegt. Um herauszufinden, inwieweit die unterschiedlichen Risikoschätzer der beiden Analyse-Ebenen auf Wechselwirkungen und Korrelationen von Radon und Rauchen innerhalb der Gruppen zurückgeht, wurde folgendes Modell (nach Lubin 1998) an die Daten angepasst:

$$\begin{aligned} R_k &= \alpha \sum \theta_{ik} n_k^{-1} (1 + \beta \sum \theta_{ik} w_{ik} n_k^{-1} \div \sum \theta_{ik} n_k^{-1}) \\ &= \alpha \theta_k \{1 + \beta \text{avg}(\theta_{ik} w_{ik}) \div \theta_k\} \end{aligned}$$

wobei α die Baseline rate der Erkrankung nach Adjustierung für die Kovariablen darstellt, θ_{ik} das relative Risiko durch Rauchen, bezogen auf den Raucherstatus der i-ten Person der k-ten Gemeinde. w_{ik} ist die individuelle Radonexposition, θ_k der mittlere Raucherstatus, n_k die

Ergebnis der schwedischen Studie

Für die Auswertung auf individueller Ebene wurde nur die Gruppe der männlichen Raucher herangezogen, die die meisten Fälle enthielt. Innerhalb der Gemeinden waren die mittleren Radonkonzentrationen der Fälle immer höher als die der Kontrollen, was auf eine positive Beziehung zwischen Radonexposition und dem Risiko, an Lungenkrebs zu erkranken hindeutet. Im Gegensatz dazu nahm die Lungenkrebsrate der gesamten Gemeinde mit zunehmender mittlerer Radonkonzentration ab. Die mittlere Radonkonzentration in Gemeinden mit kleinen Odds-Ratios auf individueller Ebene war höher als in Gemeinden mit

größeren Odds-Ratios. Es bestand also eine negative Beziehung zwischen dem Odds-Ratio einer Gemeinde und ihrer mittleren Radon-Konzentration, obwohl innerhalb der Gruppen auf individueller Ebene eine positive Beziehung bestand.

Tabelle 5: Vergleich des individuell und ökologisch geschätzten zusätzlichen relativen Risikos (ERR) bei 100 Bqm-3 und 95%-Confidenzintervalle (95%-CI) für unterschiedliche Adjustierung für die Covariablen in der schwedischen Fall-Kontroll-Studie zu Radon und Lungenkrebs 1980-1984. (nach Lagarde und Pershagen 1999, verändert)

Einbezogene Covariablen, zusätzlich zu Alter, Geschlecht, Stadt/Land und Beruf	Individuelle Analyse		Analyse auf Gruppenebene	
	ERR	95%-CI	ERR	95% CI
Keine	0,07	-0,01 bis 0,15	-0,02	-0,21 bis 0,17
Rauchen (aggregiert)	0,07	-0,01 bis 0,15	-0,02	-0,21 bis 0,17
Rauchen (individuell)	0,08	0,00 bis 0,002	0,00	-0,21 bis 0,21
Breitengrad	0,09	0,01 bis 0,20	0,05	-0,19 bis 0,29
Breitengrad und Rauchen (aggregiert)	0,09	0,01 bis 0,120	0,05	-0,20 bis 0,29
Breitengrad und Rauchen (individuell)	0,11	0,02 bis 0,24	0,14	-0,17 bis 0,45

Die ökologische Analyse führte auch bei unterschiedlicher Adjustierung für Kovariablen zu negativen Beziehungen zwischen Radonkonzentration und Lungenkrebs. Nur wenn die Kovariable Breitengrad in die Analyse mit einbezogen wurde, näherten sich die ökologischen Risikoschätzer denen auf individueller Ebene an.

Der individuelle Risikoschätzer für alle Personen betrug, bei Adjustierung für Alter, Geschlecht, Stadt/Land und Beruf, aber nicht für Gemeinde, 0,08 (95%-CI 0,00 bis 0,17), zeigte also eine positive Beziehung zwischen Radonkonzentration und Lungenkrebs, liegt aber etwas niedriger als der mittlere Schätzer innerhalb der Gruppen mit 0,11 (0,01 bis 0,27).

Material und Methoden der englischen Studie

In die Fall-Kontroll-Studie wurden innerhalb von 4 Jahren 982 Lungenkrebs-Fälle eingeschlossen. Die Teilnehmer waren jünger als 75 Jahre alt und hatten mindestens 20 Jahre im Studiengebiet gelebt. Sie wohnten in Cornwall oder Devon, wo sie einem von 16 Regierungsbezirken zugeordnet werden konnten. Den Fällen wurden 1486 Bevölkerungskontrollen gleichen Alters und Geschlechts zugeordnet, die zufällig aus den Registern der betreffenden Bezirke ausgewählt wurden.

Individuelle Exposition: Für Fälle und Kontrollen wurde die Radonkonzentration in den Innenräumen aller Wohnungen der letzten 35 Jahren gemessen. 15% des Wohnraumes wurde nicht erfasst und aus Mittelwerten der Region geschätzt. Als Exposition wurde die Radonkonzentration der Räume gewichtet nach den darin verbrachten Lebensjahren verwendet.

Aggregierte Exposition: Mittelwerte der Kontrollen wurden für jeden Regierungsbezirk ermittelt. Es wurde eine log-Normalverteilung vorausgesetzt: $\exp(x+0,5s^2)$ mit x Mittelwert und s^2 Varianz des Logarithmus der individuellen Daten.

Mögliche Confounder wie Rauchen, Alter, Geschlecht, sozialer Status und Wohnort wurden mittels Fragebogen erfasst. Die Confounder gingen auf individuellem Level in die Berechnung ein, nur für Rauchen wurde ein Durchschnittswert für die Region aus den Daten der Bevölkerungskontrolle erstellt.

Berechnung:

Das relative Risiko wurde als Maximum-Likelihood-Schätzer des Odds-Ratios mittels logistischer Regression adjustiert für Alter, Geschlecht und andere Faktoren ermittelt. Die Berechnungen wurden getrennt für Cornwall und Devon durchgeführt.

Ergebnisse der englischen Studie

Es gab keinen signifikanten Unterschied des relativen Risikos für Lungenkrebs zwischen den 6 Bezirken in Cornwall, aber das relative Risiko korrelierte mit der mittlerer Radonkonzentration (0,71, $p=0,11$)

In Devon gab es signifikante Unterschiede zwischen den relativen Risiken der 10 Distrikte, aber keine Korrelation zwischen den relativen Risiken und den mittleren Radonkonzentrationen (-0,35, $p=0,31$).

Die Streuung der individuellen Expositionswerte war deutlich größer als die Streuung der Mittelwerte zwischen den Distrikten.

Tabelle 6: Relatives Risiko (RR) und 95%- Confidenzintervalle (95%-CI) für Lungenkrebs bei Radonkonzentration in Innenräumen von 100 Bqm⁻³ vergleichen mit 0 Bqm⁻³ bei ökologischer und individueller Regression unter Berücksichtigung unterschiedlicher Confounder. Daten der englischen Fall-Kontroll Studie für Radon in Innenräumen. (Darby et al. 2001, Tab. 2, verändert)

Adjustiert für	Ökologische Analyse		Individuelle Analyse	
	RR	95%-CI	RR	95%-CI
Cornwall				
Keine Variablen	1,46	0,89 bis 2,38	1,08	0,95 bis 1,23
Rauchen (individuell)	1,44	0,88 bis 2,36	1,08	0,95 bis 1,23
Rauchen (ökologisch)	1,44	0,82 bis 2,51	1,14	0,98 bis 1,23
Devon				
Keine Variablen	0,36	0,16 bis 0,82	0,83	0,66 bis 1,04
Rauchen (ökologisch)	0,45	0,17 bis 1,17	0,85	0,67 bis 1,06
Rauchen (individuell)	0,56	1,21 bis 1,48	0,97	0,75 bis 1,24
Stadt-Land	0,93	0,34 bis 2,55	0,88	0,7 bis 1,11
Stadt-Land und Rauchen (ökologisch)	1,56	0,48 bis 5,05	0,89	0,71 bis 1,12
Stadt-Land und Rauchen (individuell)	1,35	0,42 bis 4,37	1,01	0,79 bis 1,31

Die Unterschiede zwischen Devon und Cornwall in der Beziehung zwischen Lungenkrebsrate und mittlerer Radonkonzentration in den Bezirken zeigte sich auch in den Ergebnissen der ökologischen Regression. Nach Adjustierung für Alter und Geschlecht war in Cornwall ein Anstieg der mittleren Radon-Konzentration von 0 auf 100 Bq /m³ mit einem relativen Risiko von 1,46 (95%CI: 0,89 bis 2,38) verbunden. In Devon führte derselbe Anstieg der Exposition zu einem relativen Risiko von 0,36 (95%-CI: 0,16 – 0,82).

Innerhalb der Länder, war der Punktschätzer des relativen Risikos gleich, unabhängig davon, ob Rauchen auf individueller Ebene oder auf aggregierter Ebene verwendet wurde. Das 95%-CI war allerdings auf individueller Ebene größer als auf aggregierter Ebene. Adjustierung für

soziale Schicht führte zu keiner Veränderung des RR in Cornwall und erhöhte das RR in Devon etwas. Auch nach Adjustierung für Rauchen und soziale Schicht blieb das RR in Devon unter 1, in Cornwall über 1.

Um herauszufinden, inwieweit Rauchen als Confounder einen ökologischen Bias darstellt, wurde der Zusammenhang zwischen den täglichen Zigaretten und der Radonkonzentration auf individuellem Level errechnet. Ein Problem würde auftreten, wenn für Distrikte mit hohen mittleren Radonkonzentrationen eine andere Beziehung bestände als für weniger belastete Distrikte. Für alle Personen gemeinsam betrug die Korrelation zwischen der Radonkonzentration und dem Raucherstatus $-0,02$. Für die einzelnen Distrikte lag der Korrelationskoeffizient für Cornwall immer nahe 0, so dass kein Bias durch Confounding auf individueller Ebene zu erwarten ist. In Devon lagen die Korrelationen dreier Distrikte höher: 0,32, 0,2, 0,18, eine Verzerrung des Schätzers ist wahrscheinlich.

Städtische oder ländliche Wohnlage ist normalerweise nicht als Risikofaktor für Lungenkrebs bekannt. Mit der Urbanisierung sind möglicherweise die Raucher-Vergangenheit der Studienteilnehmer, die Belastung am Arbeitsplatz oder die Luftverschmutzung korreliert. In dieser Analyse erwies sich der Faktor als Confounder auf Gruppenebene, nicht aber auf individueller Ebene. Innerhalb der wenigen städtischen Bezirke, nicht aber in den ländlichen Bezirken Devons bestand eine positive lineare Beziehung zwischen der mittleren Radonkonzentration und der distriktspezifischen Korrelation zwischen Radon und Raucherstatus.

Städtische oder ländliche Wohnlage war für sich allein ein Risikofaktor für Lungenkrebs sowohl auf individueller als auch auf ökologischer Ebene, auch nach Adjustierung für den Faktor Rauchen (RR=1,33 (CI 1,06 – 1,67)), bzw. 1,30 (1,04 – 1,61). Das RR war in den drei Distrikten mit städtischen Ballungsgebieten in Devon deutlich höher als in den ländlichen Regionen. Wurden diese drei Distrikte bei der aggregierten Berechnung ausgeschlossen, lag das RR für Devon gesamt bei 0,25. Bei Adjustierung für Stadt-Land und Rauchen ergab sich für Devon ein RR, das dem für Cornwall gut entspricht. In den städtischen Gebieten war die Distrikt-spezifischen Radon-Rauchen-Korrelation und die mittlere Radon-Konzentration linear.

Diskussion beider Studien

Ökologische Studien sind dann sinnvoll, wenn die Verteilung der Expositions-Variable Clusters innerhalb bestimmter Gebiete bildet und die Varianz zwischen den Gebieten sehr viel größer ist als innerhalb der Gebiete (Sheppard et al. 1996). Häufig, wie in beiden vorliegenden Studien, ist die Varianz zwischen den Gruppen aber deutlich kleiner als innerhalb der Gruppen. In ökologischen Studien mit multifaktoriellen Beziehungen ist große Vorsicht bei der Interpretation nötig, besonders wenn Effektmodifikationen mit Kovariablen vorliegen. Wenn auf individueller Ebene Korrelationen zwischen Exposition und weiteren Risikofaktoren der Erkrankung bestehen, und die Kofaktoren nicht gleichmäßig über die Gruppen verteilt sind, kann ein ökologischer Bias erzeugt werden. Besonders wenn der relativ geringe Effekt einer Exposition auf eine Erkrankung untersucht wird, und gleichzeitig ein starker weiterer Risikofaktor vorliegt, hängt der Effekt der Exposition auf die Erkrankung innerhalb der ökologischen Einheiten deutlich stärker von der gemeinsamen Verteilung von Exposition und Kofaktor ab, als von der Höhe der Exposition selber. Adjustierung auf Gruppenebene verändert den ökologischen Bias nicht, weil der Grund der Verzerrung, die unterschiedliche gemeinsame Verteilung von Exposition und Confounder innerhalb der Gruppen nicht berührt wird.

In der schwedischen Studie tritt dieses Problem nicht auf, weil keine starke Beziehung zwischen der Korrelation zwischen Radonkonzentration und Raucherstatus innerhalb jeder Gruppe, und den mittleren Radonkonzentrationen der Gruppen besteht. Die Tatsache, dass eine Adjustierung für den Breitengrad die Diskrepanz zwischen individuellem und ökologischem Schätzer deutlich verkleinert spricht dafür, dass mit der geographischen Verteilung der Gruppen weitere unkontrollierte Risikofaktoren verbunden sind. Viele Umweltfaktoren, die mögliche Risiken für Erkrankungen darstellen, sind geographisch nicht zufällig verteilt. In ökologischen Studien mit räumlich getrennten Gruppen ist das Problem unkontrollierter Confounder auf Gruppenebene groß.

In der englischen Studie bringt die ökologische Analyse der Bezirke von Cornwall einen Risikoschätzer, der dem der individuellen Analyse weitgehend entspricht, und dessen Konfidenzintervall den individuellen Schätzer einschließt. Eine ökologische Auswertung über die Bezirke in Devon erbringt dagegen eine signifikant negative Beziehung zwischen Radon und Lungenkrebs, das Konfidenzintervall schließt den individuellen Schätzer aus. Erst nach Adjustierung für den Urbanisierungsgrad der Distrikte entspricht das Relative Risiko weitgehend dem individuellen Schätzer. In Devon besteht für die städtischen Bezirke ein negativer Zusammenhang zwischen der mittleren Radonexposition und der Korrelation zwischen Rauchen und Radon. Die gemeinsame Verteilung der Exposition und eines wichtigen Confounders unterscheiden sich zwischen den Gruppen. Es besteht damit eine Effektmodifikation auf Gruppenebene.

In beiden Studien war Rauchen als wichtigster Confounder für die Beziehung zwischen Radon und Lungenkrebs bekannt, aber auch die Kontrolle dieses Confounders auf mehreren Ebenen, wie sie für ökologische Studien vorgeschlagen wurde (Greenland und Robin 1994), konnte den ökologischen Bias nicht beseitigen.

Sowohl die schwedische als auch ein Teil der englischen ökologischen Analyse des Einflusses von Innenraumradon auf das Lungenkrebsrisiko zeigen negative Beziehungen, obwohl individuelle Analysen in beiden Fällen positive Beziehungen hervorbringen. Adjustierung für einen Faktor, der normalerweise nicht als Risikofaktor für die Erkrankung auf individueller Ebene bekannt ist, führt zu einer Angleichung der ökologischen Ergebnisse an die individuellen Risikoschätzer. In der schwedischen Studie ist dieser Faktor der Breitengrad der Gemeinde, in der englischen Studie der Urbanisierungsgrad. Der Grund für den Einfluss der beiden Faktoren Breitengrad und Urbanisierung ist noch nicht verstanden. Es kann sich dabei um weitere, mit der Exposition korrelierte Risikofaktoren für die Erkrankung handeln, oder aber um unterschiedliche Beziehungen der mittleren Exposition und der gemeinsamen Verteilungen der Variablen innerhalb der Distrikte. Weitere Verzerrungen können durch Messfehler auftreten.

Der Breitengrad ist kein bekannter Risikofaktor für Lungenkrebs. Mit der Urbanisierung sind möglicherweise die Raucher-Vergangenheit der Studienteilnehmer, die Belastung am Arbeitsplatz oder die Luftverschmutzung korreliert, die mit den angewandten Messinstrumenten nicht ausreichend erfasst worden sind. Warum diese Faktoren auch mit der Radonkonzentration korrelieren, ist nicht direkt einsichtig.

Wenn es bei der epidemiologischen Untersuchung eines Risikofaktors für eine Krankheit Unterschiede zwischen den Schätzer ökologischer und individueller Studien gibt, müssen nicht die individuellen Schätzer dem wahren Wert ausnahmslos besser entsprechen als die ökologischen Schätzer (Greenland und Robins 1994). Auch individuelle Analysen können Gegenstand unterschiedlicher Verzerrungen sein. Nur selten werden individuelle Daten in

Kohorten- oder Fall-Kontroll-Studien auch auf Gruppenebene ausgewertet. Wären die beiden vorliegenden Untersuchungen nur als ökologische Studien durchgeführt worden, hätte nur die Analyse im Bezirk Cornwall ein glaubwürdiges Ergebnis hervorgebracht. Die ökologischen Analysen in Devon und in der schwedischen Studie weichen von den individuellen Ergebnissen extrem ab und wären als Grundlage gesundheitswissenschaftlicher Entscheidungen nicht geeignet. Andererseits zeigen die Untersuchungen, dass ökologische Analysen zusätzliche Einflussfaktoren auf Gruppenebene identifizieren können, die unter bestimmten Umständen auch individuelle Studien beeinflussen können.

2.2.8 Zusammenfassung

Ökologische Studien haben eine Reihe von Vorteilen, vor allem sind sie schnell und preiswert durchzuführen. In vielen Fällen sind sie aufgrund der Datenlage das einzige mögliche Studiendesign. Auf der anderen Seite bestehen große Probleme in der Interpretation ökologischer Analysen, besonders wenn von Risiken auf der Gruppenebene auf individuelle Risiken geschlossen wird.

Einige mögliche Ursachen für den ökologischen Bias sind theoretisch intensiv untersucht worden. Im Mittelpunkt stehen die Auswirkungen von Messfehlern und Confounding auf individueller Ebene und auf Gruppenebene. In der Regel wurden diese Auswirkungen in linearen oder log-linearen Modellen beobachtet. Die meisten Autoren beziehen sich dabei mehr oder weniger konkret auf die Probleme bei der Risikoschätzung von Radon auf Lungenkrebs. Weniger gut ist bekannt, ob die in diesem Beispiel auftretenden Probleme, verursacht durch einen einflussreichen weiteren Risikofaktor für die Erkrankung, auch auf andere epidemiologische Fragestellungen übertragen werden können. Noch weitgehend unbekannt sind z.B. der Einfluss von Screeningprogrammen oder mögliche Verzerrungen bei nicht linearen oder log-linearen Zusammenhängen zwischen Exposition und Erkrankung.

In der Praxis werden ökologische Studien weiterhin für viele Fragestellungen durchgeführt und häufig wird der gewonnene Schätzer unkritisch auf das individuelle Risiko durch die Exposition übertragen.

Um einen möglichst wenig verzerrten Schätzer zu erhalten ist es wichtig, so viele Variablen wie möglich auf individueller Ebene zu erfassen. Auch wenn die Expositionsvariable ausschließlich auf Gruppenebene vorliegt, ist es nachteilig, die anderen Variablen ebenfalls nur auf dieser Ebene zu erheben. Die Genauigkeit eines Effektschätzers aus einer ökologischen Studie kann oft durch zusätzliche Informationen über die gemeinsame Verteilung innerhalb der Gruppen von Kovariablen, von Exposition und Kovariablen und von Erkrankungsstatus und Kovariablen verbessert werden.

Umstritten bleibt, inwieweit valide Schlussfolgerungen über den Effekt einer Exposition auf eine Erkrankung aus rein ökologischen Studien gewonnen werden können. Die Validität von Schlussfolgerungen kann von einer genauen Klärung und kritischen Untersuchung der Vorannahmen für diese Schlussfolgerung nur gewinnen.

Ökologische Studien haben sich in theoretischen und angewandten Studien als äußerst sensibel gegenüber Bias durch Confounding und Messfehler erwiesen. Auch wenn individuelle Studien ähnliche Probleme haben, dürfen diese nicht übersehen oder gleichgesetzt werden.

Die Möglichkeit von Bias beweist andererseits nicht die tatsächliche Anwesenheit von Bias in einer konkreten Untersuchung. Eine Diskrepanz zwischen Schätzern auf individuellem -und

auf Gruppenlevel bedeutet noch nicht allein, dass die ökologischen Schätzer einen größeren Bias aufweisen. Gründe dafür sind:

Die beiden Schätzer-Typen werden von teilweise gleichen, teilweise anderen Verzerrungsquellen beeinflusst. Es ist daher möglich, dass die Schätzer auf individueller Ebene stärker verzerrt sind.

Die Effekte, die in den beiden Studientypen gemessen werden überlappen zwar, sie können aber doch unterschiedlich sein. Ökologische Studien haben eine kontextuelle Komponente, die in individuellen Studien fehlt, weil sie nur auf eine begrenzte Population bezogen sind. Tatsächlich kann die kontextbezogene Komponente der ökologischen Studien als Schlüssel ihrer größten Stärke und Schwäche bezeichnet werden, denn sie ist besonders wichtig, auch wenn sie extrem anfällig für Bias ist (Greenland 2001).

Multilevelstudien bieten die Möglichkeit, innerhalb einer Analyse individuelle und ökologische Komponenten zu vereinen. Solange es aber zu einer bestimmten, epidemiologischen Fragestellung keine guten Multilevelstudien gibt, werden ökologische Studien weiter für Kontroversen sorgen und den Anstoß für neue, aussagekräftigere Studien bilden.

2.3 Ökologische Studien zur Untersuchung des Krebsrisikos durch ionisierende Strahlung

Ionisierende Strahlen gehören zu den bestuntersuchten menschlichen Kanzerogenen. Im Hochdosisbereich sind ionisierende Strahlen Risikofaktoren für eine Reihe von Tumorformen, vor allem für Leukämie, Schilddrüsenkrebs, Brustkrebs und Lungenkrebs (Boice et al. 1997). Im Niedrigdosisbereich ist das Abschätzen des Risikopotentials schwieriger. Von ionisierenden Strahlen im Niedrigdosisbereich sind aber weite Teile der Bevölkerung betroffen, deswegen sind Risiken in diesem Bereich weiterhin von großem gesundheitswissenschaftlichem Interesse. Ökologische Studien waren in den achtziger Jahren des letzten Jahrhunderts ein häufiges Studiendesign in diesem Forschungsbereich. Inzwischen liegen aber zu vielen Forschungsfragen aussagekräftige Studien mit individuellen Daten vor und die Kritik an ökologischen Studien vergrößerte sich. Deswegen wurden in den neunziger Jahren deutlich weniger ökologische Studien zum Krebsrisiko durch ionisierende Strahlung durchgeführt als in den Jahren zuvor. Die theoretische Diskussion über die Validität von Risikoschätzern aus ökologischen Studien hat dazu geführt, dass ökologische Studien in Reviews selten zitiert werden, ohne auf ihre methodischen Einschränkungen hinzuweisen.

In anderen Bereichen der Umweltepidemiologie ist das ökologische Studiendesign nach wie vor häufig, auch zur Abschätzung quantitativer Fragestellungen (siehe auch Tab.1). Die folgende Zusammenstellung ökologischer Studien in der Strahlenepidemiologie führt ökologische Studien in der Strahlenepidemiologie seit 1990 auf. Einen Schwerpunkt bilden Studien zu Kinderkrebs in der Umgebung von Atomkraftwerken.

2.3.1 Methode der Literatursuche

Zum Auffinden ökologischer Studien in der Epidemiologie ionisierender Strahlung seit 1990 wurden Recherchen in der Literaturdatenbank Medline nach englischsprachigen Artikeln durchgeführt. Weitere Artikel wurden aus den Literaturverzeichnissen der dort gefundenen Studien gewonnen. Die theoretische Literatur zu den speziellen Problemen des ökologischen

Studiendesigns findet man unter den Suchbegriffen „ecologic(al) study(ies)“, „aggregated studies“, „studies on group level“, (Morgenstern 1998, Greenland 2001). Der Suchbegriff „ecologic(al) study(ies)“ in Verbindung mit „ionising radiation“ erzielte aber nur zwei relevante Ergebnisse zum Thema Krebs nach Tschernobyl (Tukiendorf 2001, Kritidis und Florou 2001) und ein Ergebnis zum Thema Krebs in der Umgebung von Atomkraftwerken (Laurier et al. 2000). Im Gegensatz zu epidemiologischen Fragestellungen zu anderen Umwelt-Risikofaktoren, einschließlich zu nichtionisierenden Strahlen (siehe auch Tabelle 1), wurden in der Epidemiologie ionisierender Strahlen ökologische Studien fast nur zur Beantwortung qualitativer Fragestellungen angewandt. Die Autoren ordnen das Studiendesign selber in den Bereich der deskriptiven bzw. hypothesengenerierenden Epidemiologie ein. Eine Medline-Recherche mit den Suchbegriffen: „clusterstudy“, „incidence-study“ und „geographical studies“, ergänzt durch die Suchfunktion „ähnliche Themen suchen“ erzielte die in Tabelle 1 bis 5 zusammengefassten Ergebnisse.

2.3.2 Strahlenquellen

Menschen sind in ihrer natürlichen Umgebung generell gegenüber verschiedenen Quellen ionisierender Strahlung exponiert. Die wichtigste Quelle sind natürliche Hintergrundstrahlen, zu denen die kosmischen und terrestrischen Strahlen gehören, sowie Strahlung beim Zerfall von radioaktiven Stoffen, die mit der Nahrung in den Körper gelangen oder eingeatmet werden. Die auf menschliches Handeln zurückzuführenden Strahlenquellen sind Bestrahlung im medizinischen Bereich zu diagnostischen und therapeutischen Zwecken, berufliche Exposition, Exposition gegenüber radioaktivem Fallout nach Atombombenversuchen und Reaktorunfällen, sowie Exposition in der Umgebung von atomaren Anlagen. Die verschiedenen Strahlenquellen senden die physikalisch unterschiedlichen α , β und γ - Strahlen aus. Auf die Strahlenarten wird im Folgenden nicht weiter eingegangen, weil sie für die Wahl des Studiendesign keine Rolle spielen.

Experimentelle Studien sind im Humanbereich aus ethischen Gründen fast nie möglich, deswegen werden epidemiologische Beobachtungsstudien an Populationen durchgeführt, die aus unterschiedlichen Gründen gegenüber ionisierenden Strahlen exponiert waren. Neben Fall-Kontroll Studien und Kohortenstudien werden zu bestimmten Fragestellungen der Strahlenepidemiologie häufig ökologische Studien durchgeführt.

2.3.3 Ausgewählte Studien

Epidemiologische Forschung zu gesundheitlichen Folgen ionisierender Strahlung wird in den folgenden Bereichen durchgeführt:

- Medizinische Anwendungen
- Berufliche Exposition
- Folgen der Atombombenabwürfe und -versuche
- Krebs in der Umgebung von Kernkraftwerken
- Folgen von Reaktorunfällen
- Radon in Innenräumen
- Natürliche Hintergrundstrahlung

In Studien im Bereich der medizinischen Exposition und der beruflichen Exposition liegen sowohl Expositions- als auch Erkrankungsdaten auf individueller Ebene vor. Ein Teil der Studien zur beruflichen Exposition kann nach der Terminologie von Morgenstern (1998) als

teilökologische Studien betrachtet werden, weil die Exposition häufig aus dem Arbeitsplatz anhand einer *job-exposure-matrix* auf Gruppenebene geschätzt wird. Zu den Folgen der Atombombenabwürfe auf Hiroshima und Nagasaki und der amerikanischen Atombombentests (Marshall Island, Utah und Nevada) wurden große Kohortenstudien durchgeführt und die Exposition auf individueller Ebene geschätzt.

1) Radon in Innenräumen

In ökologischen Studien zu Radon in Innenräumen und Lungenkrebs erwies sich Radon als ein protektiver Faktor (Cohen 1990, Cohen 1995). Dieses Ergebnis, das im klaren Gegensatz zu Studien auf individueller Ebene steht, kann durch mögliche Korrelationen mit der Confounder-Variablen „Rauchen“ erklärt werden. Der entsprechende mathematische Mechanismus, der einen ökologischen Bias verursacht, wurde in den Arbeiten von Lubin (1998, 2002) erläutert. Die besonderen Probleme ökologischer Studien bei dieser Fragestellung wurden bereits in Abschnitt 2.2.4 am Beispiel von Cohens ökologischer Regression dargestellt. Die Studien von Darby et al. (2001) in England und Lagarde und Pershagen (2001) in Schweden zeigen Gründe für den Bias in ökologischen Studien im Vergleich zu Studien auf individueller Ebene mit realen Daten auf (Abschnitt 2.2.7).

2) Krebs in der Umgebung von Atomkraftwerken

1983 wurde von einem Leukämie-Cluster bei Kindern in der Umgebung des Kernkraft-Reaktors Sellafield in Schottland berichtet. Seitdem wurden das Krebsrisiko und insbesondere das Leukämierisiko von Kindern und Jugendlichen in der Umgebung von kerntechnischen Anlagen in vielen Ländern untersucht und kontrovers diskutiert. Eine große Zahl von Studien in diesem Bereich wurde ausschließlich auf Gruppenebene durchgeführt, weil die gemeinsame Verteilung von Exposition und Krankheit auf individueller Ebene nicht vorlag.

Tabelle 7: Ökologische Studien zu regionalen Clustern in der Umgebung von kerntechnischen Anlagen

Quelle	Region	Studienperiode	Exposition	Risikomaß	Ergebnis (RR ¹ , SIR ² , AR ³)	Erkrankung und Altersgruppe
Draper et al. (1993)	England Sellafield Seascale	1984-1990	2 Distrikte in Cumbria, ein Distrikt in Copeland	Inzidenz	<15 Jahre: SIR: 14,4 (s.) ⁴ 15-24 Jahre: SIR: 24 (s.)	Leukämie und Non-Hodgkin's Lymphome (< 25 Jahre)
Kinlen (1993)	LJ. England Seascale	1951-1991	Einwohner Seascales	Inzidenz	In Seascale geboren: Leukämie : SIR: 8,4 (s.); Non-Hodgkin's L. SIR: 20,2 (s.) Zugezogen: Leukämie: SIR: 7,2 (s.); Non-Hodgkin's-Lymphome: SIR: 16,5 (s.)	Leukämie und Non-Hodgkin's Lymphome (<25 Jahre)
Black et al. (1994)	Schottland, Dounreay	1968-1991	25 km Umkreis 1) bezogen auf Nachbargebiete 2) bezogen auf ganz Schottland	Inzidenz	<25 Jahre: 25 km: SIR: 2,3 (s.) 15 km: SIR: 3,0 (s.) <4 Jahre, (<25 km): 1) SIR: 1.04 (n. s.) ⁵ 2) SIR: 2.97 (s.)	Leukämie und Non-Hodgkin's Lymphome (< 25 Jahre) (<4 Jahre)
Laurier (2000)	D. Frankreich, La Hague	1978-1996	≥ 1 Jahr < 10 km	Inzidenz	AR = 0.002	Leukämie: (< 25 Jahre)
Rommens et al. (2000)	Frankreich, La Hague	1978-1996	<25 km	Inzidenz	AR = 0,002	Leukämie (< 25 Jahre)
Guizard et al. 2001	Frankreich La Hague	1978-1998	<10km <35 km	Inzidenz	35 km-Radius: SIR: 1.23 [KI=0.73-1.41] 0-10 km: 2,17 [KI=0.71-5,07]; 5-9 Jahre, Leukämie: SIR = 6.38 [KI=1.32-18.6]	Krebs, Leukämie (< 25 Jahre)
Hattchouel et al. (1995)	Frankreich La Hague	1968-1989	16 km Radius	Mortalität	SIR: 0,4 [KI: 0,1-7,3]	Leukämie (< 25 Jahre)
Hill & Laplanche (1992)	Frankreich La Hague	1968-1987	10 km Radius	Mortalität	SIR= 0,8 [KI:0,4-1,4]	Leukämie (< 25 Jahre)
Viel et al. (1995)	Frankreich, La Hague	1978-1992	< 10 km < 35km	Inzidenz	< 10 km: SIR: 2,8 [KI: 0.8-7.3] < 35 km: SIR: 1,1 [KI: 0.7-1.6]	Leukämie (< 25 Jahre)

Bouges et al. (1999)	Frankreich, Marcoule	1985-1995	< 35 km Wiederauf- bereitungs-anlage	Inzidenz	SIR = 1.06 [KI = .72-1.31]	Leukämie, Lymphome, Schilddrüsenneoplasmen (< 15 Jahre)
Dieckmann H. (1992)	Deutschland Krümmel	1990-1991	1 Siedlung	Inzidenz	SIR: 7/< 1	Leukämie und Non- Hodgkin's Lymphome (< 25 Jahre)
Grosche et al. (1999)	Vergleich Deutschland Krümmel – Vergleich USA	1990-1996	< 10 km Radius Kernkraft (Tritium)	Inzidenz Deutschland- USA	SIR: 3,3 [KI:1.6-6.0]	Leukämie (< 15 Jahre)
Hoffmann W. (1997)	Deutschland Krümmel	1990-1991 1990-1995	5 km Radius	Inzidenz	1990-1991: SIR=1.18 [KI: .49-2.83] 1990-1995: SIR=0.46 [KI: .21-1.03]	Leukämie (< 15 Jahre)

¹ RR: Relatives Risiko² SIR: Standardisierte Inzidenzrate³ AR: Absolutes Risiko⁴ (s.): signifikant, $p < 0.05$ ⁵ (n. s.): nicht signifikant, $p > 0.05$

Tabelle 8: Ökologische Studien in der Umgebung mehrerer kerntechnischer Anlagen (multi-site studies)

Quelle	Region	Studienperiode	Anzahl nuklearer Anlagen (N)	Exposition	Risikomaß	Ergebnis (RR; SIR; SMR ¹)	Erkrankung und Altersgruppe
Barton C. (2001)	West-Berkshire, U.K.	1972-1996	N = 6	< 10 km zur Atomkraftanlage	Inzidenz	SIR= 1.30 (s.) ²	Leukämie(< 5 Jahre)
Bithell et al. (1994)	England und Wales	1966-1987	N= 23	25 km Umkreis	Inzidenz	SIR= 1.10 (s. bei 2 Anlagen)	Leukämie und Non-Hodgkin's Lymphome (< 15 Jahre)
Clarke et al. (1991)	Ontario, Kanada	1964-1986	N = 6	Umgebung Kernkraftanlage	Inzidenz	SIR: 1.06 (n.s.) ³	Leukämie (< 15 Jahre)
Goldsmith JR. (1992)	UK	1971-1980	N = 21	Umgebung von 6 Nuklearanlagen	Inzidenz	SIR: 1.23 (n.s.)	alle Krebslokalisationen (< 10 Jahre)
Hattchouel et al. (1995)	Frankreich	1968-1993	N = 13	4 Zonen: < 5 km; 5-10 km; 10-13 km; 13-16 km	Mortalität	SMR = 0.80 [KI: 0.62-1.01])	Leukämie (< 25 Jahre)
Iwasaki et al. (1995)	Japan	1973-1987	alle Anlagen Japans	Umgebung Kernkraftanlage	Mortalität	SMR: 1.06 (n.s.)	Leukämie (< 25 Jahren)
Kaatsch et al. (1998)	Deutschland	1991-1995	N = 20	< 15 km	Inzidenz	< 5 Jahre + < 5 km: RR = 1.39 (KI: 0.69-2.57) alle Fälle (N = 2390): RR = .099 [KI: 0.91-1.07]	alle Krebslokalisationen (< 15 Jahre)
Körblein Hoffmann (1999)	Westdeutschl and Reanalyse der Daten: Michaelis et al. (1998)	1980-1995	N = 15	< 5 km	Inzidenz	Krebs (<15 Jahre): 1,22 (s.) Leukämie (<15 Jahre): 1,34 (n.s.) Krebs (<5 Jahre): 1,53 (s.) Leukämie (<5 Jahre): 1,76 (s.)	Alle Krebslokalisationen Leukämie (<15 Jahre) (<5 Jahre)
Lopez-Abente et al. (1999)	Spanien,	1975-1993	N = 7 Kernkraftanlagen + 5 Kernbrennstoffanlagen	< 30 km	Inzidenz	erhöhtes Risiko einer leukämisch bedingten Erkrankung bei 2 der 7 Anlagen RR = 1.3 [KI: 1.03-1.64] bzw. RR = 1.68 [KI: 0.92-3.08]	Leukämie, Lymphome, Myelome (alle Altersgruppen)
Lopez-Abente et al. (2001)	Spanien	1975-1993	N = 8	< 30 km	Inzidenz	Lungenkrebs RR=1.12, [KI: 1.02-1.25]; Nierenkrebs RR= 1.37, [KI: 1.07-1.76] in	Alle Krebslokalisationen (alle Altersgruppen)

					der Umgebung von Uran-verarbeitenden Anlagen, nicht bei anderen Kernkraftanlagen		
Mangano et al. (2002)	U.S.A.	1987- 2002	N = 8	< 64 km; downwind	Inzidenz, Mortalität	2 Jahre nach Schließung: Mortalität um 17,4% gesunken (s.); Inzidenz um 25% gesunken (s.)	Alle Krebslokalisationen (< 5 Jahre)
McLaughin et al. (1993)	Ontario (Kanada)	1950-1987	N = 5 N1 = 54	< 25 km	Inzidenz	SIR: 1.17 [KI: 0.88-1.53]	Leukämie (< 15 Jahre)
Michaelis (1998)	J. Deutschland	1980-1995	N = 20 Westdeutschland	< 15 km < 5 km	Inzidenz	RR= 0.99 [KI: 0.91-1.07] RR= 1.0 [KI: 0.87-1.16] RR= 1.49 [KI: 0.98-2.2]	Alle Krebslokalisationen (<15 Jahre) , Leukämie (< 15 Jahre) Leukämie (< 5 Jahre)
Sharp et al. (1996)	Schottland	1968-1993	N = 7	< 25 km	Inzidenz	bei 3 Anlagen RR= 1.99 (s.); 1.02 (n.s.); 1.08 (n.s.)	Leukämie und Non-Hodgkin's Lymphome (< 15 Jahre);
Waller et al. (1995)	Schweden	1980-1990	N =20	<25 km	Inzidenz	Keine erhöhte Inzidenz	Leukämie (< 15 Jahre)

¹ SMR: standardisierte Mortalitätsrate

² (s.): signifikant: $p < .05$

³ (n.s.): nicht signifikant: $p > .05$

In Tabelle 7 und in Tabelle 8 werden ökologische Studien zu diesem Thema seit 1990 zusammengefasst. Dabei wurde unterschieden zwischen Studien, die ein mögliches regionales Cluster in der Umgebung einer bestimmten Kernkraftanlage (one-site-studies) und Studien, die das Krebsrisiko in der Umgebung mehrerer Anlage eines Landes (multi site studies) untersuchten. Insgesamt wurden 13 Veröffentlichungen zu lokalen Studien und 15 Veröffentlichungen zu überregionalen Studien dargestellt.

Studiendesigns

Erkrankung:

In den Studien wurden entweder alle Krebslokalisationen und/oder speziell Leukämien und z.T. Non-Hodgkin's Lymphome untersucht. Von den insgesamt 28 Studien untersuchten 5 Studien die Altersgruppe Kinder unter 4 bzw. 5 Jahre, eine Studie untersuchte die Altersgruppe <10 Jahre, 12 Studien untersuchten die Altersgruppe unter 15 Jahren und 12 Studien Kinder und junge Erwachsene unter 24 bzw. 25 Jahren (Mehrfachnennungen möglich). Die beiden spanischen Studien (Lopez-Abente et al. 1999, 2001) untersuchten das Krebsrisiko aller Altersgruppen. Die Erkrankungen lagen auf Gruppenebene als Inzidenzrate oder Mortalitätsrate vor. Die Informationen wurden den lokalen oder nationalen Krebsregistern bzw. Sterberegistern entnommen.

Exposition:

In der Einteilung von Morgenstern (1998) werden diese Studien den rein ökologischen *multiple-group-designs* zugeordnet. Es werden verschiedene geographische Einheiten zum selben Zeitpunkt verglichen. Die Studie von Jablon et al. (1991) und Mangano et al. (2002) beziehen zusätzlich eine zeitliche Komponente mit ein, in dem sie Phasen vor der Inbetriebnahme bzw. nach der Schließung einer Anlage als Kontrollzeiträume benutzten. Die Frage, ob es sich bei diesem Design um rein deskriptive Studien, in denen nur Erkrankungsraten verglichen werden oder um analytische Studien, in denen eine Beziehung zwischen Exposition und Erkrankung untersucht wird, ist schwer abzuschätzen und wird weiter unten diskutiert.

Risikomaß:

Als Zusammenhangsmaß wird in den meisten Studien das Standardisierte Inzidenzratio (SIR) oder Mortalitätsratio (SMR) berechnet, wobei die beobachtete Inzidenz in den „exponierten“ Clustern auf die erwartete Inzidenz dieser Cluster, berechnet aus den Daten der Gesamtpopulation oder einer Vergleichspopulation, bezogen wird. In einigen Studien wird das SIR als Relatives Risiko der „exponierten“ Cluster im Bezug auf „nichtexponierte“ Vergleichspopulationen interpretiert (Cook-Mozaffari et al. 1989). Körblein (2001) benutzt in einer Neuauswertung von Daten des deutschen Kinderkrebsregisters zusätzlich den Quotienten aus den SIR in der Umgebung eines Kernkraftwerkes und dem SIR einer Kontrollgruppe. Dieses bezeichnet er fälschlicherweise als Relatives Risiko.

Ergebnisse:

Studien zu Krebs in der Umgebung von Atomkraftwerken wurden in Großbritannien, Frankreich, Deutschland, Schweden, Spanien, Kanada und den USA durchgeführt. Grundsätzlich kann zwischen lokalen Studien und überregionalen Studien unterschieden werden.

Die kleineren lokalen Studien untersuchen ein Cluster (Häufung von Erkrankungsfällen) um eine einzelne atomare Anlage. Diese sind in Tabelle 7 dargestellt. Die Studien sind nicht direkt vergleichbar, weil Unterschiede in der Zielgröße, Inzidenz, Mortalität, den Altersgruppen und Krankheiten bestehen. Die kerntechnischen Anlagen und der Radius der

„exponierten“ Cluster unterscheiden sich ebenfalls. Bisher konnten die Leukämie-Cluster um das deutsche Atomkraftwerk Krümel, die britischen Wiederaufbereitungsanlagen Sellafield und Dounreay und die Leukämie-Cluster in der Nähe der Atomwaffenfabriken Aldermaston und Burghfield (Roman et al. 1987, Bithell et al. 1994) signifikant über einen längeren Zeitraum nachgewiesen werden. Andere Cluster, z.B. La Hague in Frankreich, wurden beobachtet, weitere Auswertungen der Daten bestätigten das Ergebnis nicht (Laurier et al. 2002). Die lokalen Cluster beinhalten einige wenige Fälle. Entsprechend groß ist der Einfluss von zufälligen Schwankungen der Erkrankungsrate. Schon kleine Veränderungen im Zeitintervall und in der geographischen Auswahl der Studienpopulation können das Ergebnis verändern.

In einer Reihe überregionaler Studien wurde das Leukämierisiko um mehrere kerntechnische Anlagen gemeinsam untersucht (Tabelle 8). Auf diese Weise standen mehr Fälle zu Verfügung und die Aussagen können mit größerer statistischer Genauigkeit gemacht werden. Die meisten Studien wurden in Großbritannien durchgeführt, aber auch in den USA, Kanada, Deutschland, Frankreich, Schweden und Spanien. Obwohl sich auch diese Studien in vielen Punkten ihres Studiendesigns unterscheiden, zeigen sie übereinstimmend keinen Anstieg der Kinderkrebsinzidenzen in der Nähe der kerntechnischen Anlagen. Eine Ausnahme stellt möglicherweise die jüngste Altersgruppe (<5 Jahre) in unmittelbarer Nähe der Anlagen (5km-Radius) dar. Für diese Altersgruppe fanden Barton (2001) in West-Berkshire, Großbritannien, und Körblein und Hoffmann (1999) ein signifikant erhöhtes Risiko, an Leukämie zu erkranken im Vergleich zu den Kontrollregionen.

Diskussion

Die Studien zu Kinderkrebs in der Umgebung von kerntechnischen Anlagen werden von den Autoren als Inzidenz-Studien oder Cluster-Studien bezeichnet und damit der deskriptiven Epidemiologie zugerechnet. Mit dieser Terminologie schirmen die Autoren ihre Studien gegen die generelle Kritik an „analytischen ökologischen Studien“ (nach der Def. von Morgenstern 1998) ab. Sie erheben keinen Anspruch, eine Beziehung zwischen einer Strahlenexposition und den Inzidenzen zu messen, sondern untersuchen lediglich das Vorhandensein lokaler Cluster. Die Studien unterscheiden sich aber von anderen deskriptiven Darstellungen der geographischen Verteilung von Inzidenzen, z.B. den Krebsatlanten, und haben eine analytischen Komponente. Nicht die räumliche Entfernung zur Anlage, sondern die damit möglicherweise verbundene radioaktive Belastung ist die eigentlich interessierende und biologisch plausible Exposition. Die Entfernung der geographischen Cluster zur Anlage kann als Surrogatvariable dieser Exposition betrachtet werden.

Für die ökologischen Studien zu Kinderkrebs in der Umgebung von kerntechnischen Anlagen gelten die bereits in Abschnitt 2.2.6 und 2.2.6 ausführlich beschriebenen Einschränkungen. Auch wenn auf den ersten Blick eine strukturelle Expositionsvariable vorliegt (Wohnort innerhalb einer geographischen Einheit), wurde die Einteilung dieser geographischen Einheiten von ihrer möglichen Beziehung zu der eigentlich interessierenden Exposition gegenüber radioaktiven Strahlen bestimmt. Die individuelle Exposition gegenüber ionisierenden Strahlen in der Umgebung von Atomkraftwerken ist durch Messgeräte kaum nachweisbar. Es ist auch unbekannt, auf welchen Wegen die Kinder vornehmlich exponiert werden, möglicherweise ist die Exposition über die Nahrungskette (Fisch und Meerestiere) ausschlaggebend. Die Variabilität der wahren Exposition innerhalb der Zonen ist unbekannt.

Über Confounder und Verzerrungsquellen liegen in den Studien in der Regel keine Informationen vor. In einigen Studien wurden Bevölkerungsdichte und städtisch/ländliche Region als Auswahlkriterien der Vergleichspopulation herangezogen. Mögliche Confounder,

die sich sowohl auf individueller als auch auf Gruppenebene auswirken können, sind ionisierende Strahlung aus anderen Quellen (Medizin, Fallout, Hintergrundstrahlung), Bevölkerungsvermischung (Kinlen et al. 1995), genetische Strahlenschäden in den Keimzellen der Eltern (Gardner et al. 1990). Dazu kommen weitere mögliche Einflussgrößen wie genetische Risikofaktoren, Rauchen der Eltern, Ernährung und weitere Schadstoffbelastungen.

Ein *cross-level-bias* tritt auf, wenn Risikoschätzer auf Gruppenebene auf Zusammenhänge auf individueller Ebene übertragen werden. In den hier untersuchten ökologischen Studien zu Kinderkrebs in der Umgebung von kerntechnischen Anlagen wird auf die Übertragung der Ergebnisse auf die individuelle Ebene verzichtet. Radioaktive Strahlung wird als eine mögliche Ursache für die in einigen Studien nachgewiesenen Leukämie-Cluster diskutiert. Die Festlegung der geographischen Cluster legt diesen Schluss allerdings nahe.

Kinderkrebs in der Umgebung von kerntechnischen Anlagen ist ein Thema, das auf großes Interesse und Besorgnis in der Öffentlichkeit stößt. Wenn positive Zusammenhänge zwischen Kernkraftwerken und Kinderkrebs veröffentlicht werden, gehen Einschränkungen im Studiendesign und den verwendeten Methoden in den Massenmedien verloren. In Deutschland erregten die Arbeiten von Körblein zu Kinderkrebs um deutsche und speziell bayrische Atomkraftwerke großes Aufsehen (Körblein und Hoffmann 1999, Körblein 2001, Michaelis et al. 2001, Kinderkrebsregister 2001). Körblein hatte Daten des Deutschen Kinderkrebsregisters erneut ausgewertet. Neben anderen Schwächen war die Untersuchung nicht hypothesengeleitet durchgeführt, sondern es wurden unterschiedliche zeitliche Intervalle und räumliche Zusammenstellungen „ausprobiert“, bis signifikante Ergebnisse erreicht wurden. Die wissenschaftliche Fragwürdigkeit dieses Vorgehens, in Verbindung mit anderen Einschränkungen des ökologischen Studiendesign war der Öffentlichkeit nur schwer zu kommunizieren.

3) Hintergrundstrahlung

Für die meisten Menschen stellt die natürliche Hintergrundstrahlung die größte Quelle radioaktiver Gesamtstrahlung dar. Die Auswirkung auf die Gesundheit betrifft die gesamte Bevölkerung, die in Gebieten mit erhöhter Strahlenbelastung wohnt. Die Exposition gegenüber atmosphärischer und terrestrischer Strahlung kann praktisch nicht individuell bestimmt werden. In Regionen mit hoher Radonkonzentration in Innenräumen kann dieser Effekt nur schwer von den Auswirkungen der Gamma-Strahlung getrennt werden. In den achtziger Jahren des letzten Jahrhunderts wurden eine Reihe ökologischer Studien durchgeführt (IARC 2000), in denen die mittlere gemessene Strahlung in einem Gebiet als Exposition auf Gruppenebene zu den Krebsinzidenzen der Gebiete in Beziehung gesetzt wurde. In den neunziger Jahren wurden nur noch wenige ökologische Studien zu diesem Thema durchgeführt.

Tabelle 9: Ökologische Studien zu Krebs bei natürlicher Strahlung

Quelle	Region	Studienperiode	Exposition	Risikomaß	Ergebnis (RR/ARR ¹ /AR ² /ERR ³)	Erkrankung
Gilman und Knox (1998)	Groß-britanien	1953-1964	Hintergrundstrahlung (Gamma)	Mortalität	Kein Zusammenhang zwischen Mortalität und Gamma-Strahlung, wenn adjustiert für Radon-Konzentration	alle Krebslokalisationen (Leukämie und solid tumor getrennt) (<15 Jahre)
Little MP. (2002)	Japan	1958-1987	Hintergrundstrahlung attributabel zu Strahlung aus Atombombe	Inzidenz	ARR: 17,3-32% des Schilddrüsenkrebs auf Hintergrundstrahlung (0,5 mSv/Jahr) zurückführbar	Schilddrüsenkrebs (alle Altersgruppen)
Muirhead et al. (1991)	459 ländliche Bezirke Britannien	1969-1983	Hintergrundstrahlung (Gamma und Radon)	Inzidenz	Kein Zusammenhang zwischen Hintergrundstrahlung (im Innen- und Außenraum) und lokalen Inzidenzraten	Leukämie (<15 Jahre)
Nair et al. (1999)	Indien Kerala	keine Angabe	Hintergrundstrahlung (Gamma)	Inzidenz	Kein Zusammenhang zwischen erhöhter natürlicher Strahlung und Krebsinzidenz	alle Krebslokalisationen (alle Altersgruppen)
Sun et al. (2000)	China Yangjiang	1979-1995	Hintergrundstrahlung	Inzidenz	ERR= -0.11 [KI: -0.67-0.69] pro Sievert	alle Krebslokalisationen alle Altersgruppen
Tao et al. (1999, 2000)	China Yangjiang	1979-1995	Hintergrundstrahlung	Inzidenz	1979-1995: RR = 0.99 [KI: 0.87-1.14] 1987-1995: RR = 0.96 [KI: 0.80-1.15] ERR/Sievert = -0.10 [-0.67, 0.69]	alle Krebslokalisationen (alle Altersgruppen)
Wei & Sugahara (2002)	China	1979-1995	Hintergrundstrahlung	Mortalität	ERR = 0.01 [KI: 0.67-0.69]/Sv	alle Krebslokalisationen (alle Altersgruppen)

¹ ARR: attributables Risiko

² AR: absolutes Risiko

³ ERR: excess relative risk

In Tabelle 9 sind 7 Veröffentlichungen zu natürlicher Hintergrundstrahlung dargestellt, die sich auf fünf Datensätze aus Japan, Großbritannien, Indien und China beziehen. Während sich in den Studien in Großbritannien, Japan und Indien kein Zusammenhang zwischen der Krebsinzidenz und der Hintergrundstrahlung zeigte, wurde in einem Untersuchungsgebiet in China ein signifikant erniedrigtes zusätzliches Relatives Risiko (ERR) bzw. zusätzliches Absolutes Risiko gefunden.

Auch diese Studien weisen die Einschränkungen ökologischer Studien auf. Häufig war es nicht möglich, für Confounder wie Rauchen, sozioökonomischer Status der Region, genetische Faktoren und andere externe Belastungen zu kontrollieren. Auch regionale und zeitliche Unterschiede in der Krebsregistrierung können sich auswirken. Individuelle Variabilität in der Exposition können ebenfalls nicht kontrolliert werden. Der Schluss, dass erhöhte natürliche Strahlung kein Risiko für Krebserkrankungen auf individueller Ebene darstellt, kann allein aufgrund dieser Studien nicht gezogen werden. Studien auf individueller Ebene kommen allerdings zu ähnlichen Ergebnissen (IARC 2000).

4) Radioaktiver Fallout nach Atomwaffentests

Bei mehreren Atomwaffentests wurden Teile der Bevölkerung gegenüber radioaktiver Strahlung exponiert. Dazu gehörten die Bewohner der Marshall Inseln, Utha und Nevada. Ökologische Studien in diesem Bereich wurden vor allem in den 70ziger und 80ziger Jahren durchgeführt (siehe Boice et al. 1997).

Tabelle 10: Ökologische Studien zu Krebs nach nuklearem Fallout

Quelle	Region	Studien periode	Exposition	Risikomaß	Ergebnis (RR/AR/ERR)	Erkrankung und Altersgruppe
Darby et al. (1992)	Skandinavien	1950-1960	Fallout von nuklearer Waffen	Tests Inzidenz	< 15 Jahre: RR= 1.07 [KI: 1.00-1.14]; < 5 Jahre: RR = 1.11 [KI: 1.00-1.24]	Leukämie (< 15 Jahre) (<5 Jahre)
Gilbert et al. (1998)	USA	1973-1994	Fallout der Nevada Nuklearbombentests (50-er Jahre)	Mortalität Inzidenz	Mortalität: < 1 Jahr bei Exposition: ERR/Gy = 10,6 [KI : -1.1-29] Geburtenkohorte 1950- 1959 ERR/Gy = 12.0 [KI: 2.8-31]; Inzidenz: <1 Jahr bei Exposition: ERR/Gy:2,4 [KI : -0.5-5.6] sonst keine erhöhten ERR; Morbiditäts-oder Inzidenzraten	Schilddrüsenkrebs (<1 Jahr bei Exposition) (<5 Jahre bei Exposition) (<15 Jahre bei Exposition) (Geburtenkohorte 1950-1959)
Lund & Galanti (1999)	Skandinavien (Norwegen, Schweden)	1947-1970	Atomarer Fallout von Atombombentests	Inzidenz	7-14 Jahre: RR = 1.7 [KI: 1.0-3.0] 20-24 Jahre: RR = 0,9 [KI: 0.7-1.2]	Schilddrüsenkrebs (< 25 Jahre)
Zaridze DG (1994)	Kasachstan	1981-1990	<200 km: Strahlentests 1) in der Luft vor 1963; 2) Undergroundtests 1965; 3) Reservoir („Atomsee“)	Inzidenz < 200 km gegenüber < 400 km Entfernung zum Testgelände	Leukämie;: 1) RR = 1.76 [KI : 1.19-2.59]; 2) RR = 1.70 [KI :1.15-2.52]; 3): RR = 1.72 [KI : 1.2-2.47]. Alle Krebslokalisationen: 1) RR = 2.02 [KI :1.59-2.56]; 2) RR = 1.75 [KI :1.37-2.21]; 3):RR = 1.77 [KI :1.41-2.22]	Alle Krebslokalisationen, Leukämie (< 15 Jahre)

In Tabelle 10 sind vier neuere Studien zu diesem Thema zusammengestellt. Zwei Studien in Skandinavien zeigen ein erhöhtes Risiko für Leukämie und Schilddrüsenkrebs bei Kindern, die zeitlich mit Veränderungen von atomarem Fallout einhergehen. Eine weitere Studie untersucht Zusammenhänge zwischen Krebs und radioaktivem Fallout von Atombombentests in Kasachstan. Es war ein signifikant erhöhtes Leukämierisiko für Kinder im Umkreis von 200 km vom Testgelände, bezogen auf Kinder im Umkreis von 400 km, festzustellen. In den USA untersuchten Gilbert et al. (1998) die Mortalität an Schilddrüsenkrebs in der gesamten USA und die Schilddrüsenkrebsinzidenz in ausgesuchten Gebieten, für die entsprechende Daten vorlagen. Sie fanden erhöhte Mortalitäts- und Inzidenzraten für Kinder, die bei der Exposition jünger als ein Jahr waren, aber keine Zusammenhänge zwischen Inzidenz und Exposition für ältere Kinder. Die Mortalität war nur für die Geburtskohorte 1950-1959 erhöht. Auch bei diesen Ergebnissen kann nicht zwischen dem Einfluss der Exposition und anderen Einflussvariablen, die sich über die Zeit verändern, unterschieden werden. Insbesondere Migration (Tong 2000) kann die Ergebnisse durch die langen Beobachtungszeiträume verzerren.

5) Tschernobyl

Seit dem Reaktorunfall in Tschernobyl ist zumindest in Europa der Einfluss von radioaktivem Fallout aus den unterschiedlichen Quellen nicht mehr zu trennen. Eine Reihe von ökologischen Studien wurde in unterschiedlichen europäischen Ländern durchgeführt, um mögliche Zusammenhänge zwischen radioaktiver Exposition durch den Tschernobyl-Unfall und Krebsinzidenzen unterschiedlicher Lokalisationen und Altersstufen zu untersuchen.

Tabelle 11: Ökologische Studien zu Kinderleukämie nach Tschernobyl

Quelle	Region	Studienperiode	Risikomaß	Ergebnis (RR, SIR)	Erkrankung und Altersgruppe (< 15 Jahre)
Studien außerhalb der direkt betroffenen Gebiete Weißrussland, Ukraine und Russland					
Parkin et al. (1992, 1993, 1996)	Europa	1980-1991 1980-1988 1980-1985 1988-1989 1990-1995	Inzidenz	Signifikante Zunahme der Inzidenz; Kein Zusammenhang zwischen Leukämieraten und Exposition nach Tschernobyl	Kinderleukämie, Lymphome
Auvinen (1994)	Finland	1976-1992	Inzidenz	Keine erhöhte Inzidenz	Kinderleukämie
Gunay (1996)	Türkei	1986-1995	Inzidenz	Gehäuftes Auftreten von Fällen in einigen Kliniken	Kinderleukämie
Hjalmarsson (1994)	Schweden	1980-1992	Inzidenz	Keine erhöhte Inzidenz bei allen Kindern und Kindern 0-5 Jahre	Kinderleukämie
Michaelis (1997)	Deutschland	1986-1987	Inzidenz	Zunahme der Inzidenz; keine Zunahme der Inzidenz, die auf die Kontamination zurückgeführt werden kann	Kinderleukämie
Steiner (1998)	Deutschland	1986-1987	Inzidenz	Zunahme der Inzidenz; keine Zunahme der Inzidenz nach Kontaminationshöhe	Kinderleukämie
Steiner (1998)	West-Deutschland	> 1986	Inzidenz	Erhöhte Inzidenz bei Kindern, die nach Tschernobyl geboren wurden	Kinderleukämie
Petridou (1994)	Griechenland	1980-1991	Inzidenz	Keine erhöhten Inzidenzen nach Tschernobyl	Kinderleukämie
Petridou (1996)	Griechenland	1980-1996	Inzidenz	In utero-Exposition 2.60 [KI: 1.4-5.1]; Exposition zwischen 12-47 Monaten: keine erhöhten Inzidenzen	Kinderleukämie nach in utero Exposition
Petridou (1996)	Griechenland	1980-1990	Inzidenz bei in Utero Exponierten	RR = 2.6 [KI: 1.4-5.1]	Kinderleukämie
Tondel	Schweden	1978-1992	Inzidenz	0-19 Jahren: Zunahme der	Neoplasmen bei Kindern (und

(1996)				Gehirntumore ohne deutlichen Kontaminationseinfluss von Tschernobyl	Erwachsenen)
Tukiendorf A.(2001)	Polen	1987-1996	Inzidenz	In Abhängigkeit von der Dosis und Geschlecht: keine signifikanten Zusammenhänge zwischen Kontamination und Inzidenz	Leukämie (alle Altersgruppen)
Studien in den direkt betroffenen Gebieten Weißrussland, Ukraine und Russland					
Ivanov (1993)	et al. Belarus	1979-1985	Inzidenz	Keine erhöhte Inzidenz	Kinderleukämie
Ivanov (1996)	et al. Belarus	1982-1994	Inzidenz	Keine allgemein erhöhte Inzidenz; keine erhöhte Inzidenz, die auf die Kontamination zurückgeführt werden kann	Kinderleukämie
Ivanov (1997)	et al. Russland	1981-1995	Inzidenz	Keine erhöhte Inzidenz	Kinderleukämie
Ivanov (1998)	et al. Belarus	1982-1994	Inzidenz (in Utero Exponierte)	Keine erhöhte Inzidenz	Kinderleukämie
Noshchenko (2001)	et al. Ukraine	1986-1996	Inzidenz	Zunahme der Inzidenz in kontaminierten Gebieten	Kinderleukämie
Prisyazhiuk (1991)	et al. Ukraine	1981-1990	Inzidenz	Keine erhöhte Inzidenz	Kinderleukämie
Prisyazhiuk (1995)	et al. Ukraine	1980-1993	Inzidenz	Erhöhte Inzidenz, aber nicht mehr nach 1986	Kinderleukämie

Tabelle 12: Ökologische Studien zu Schilddrüsenkrebs bei Kindern und Erwachsenen nach Tschernobyl

Quelle	Region	Studienperiode	Risikomaß	Ergebnis (RR/SIR/ERR/EAR ¹)	Erkrankung und Altersgruppe (< 15 Jahre)
Studien außerhalb der direkt betroffenen Gebiete Weißrussland, Ukraine und Russland					
Cotterill et al. (2001)	Nord-England: Distrikte	5 1987-1997 bezogen auf 1968-1986	Inzidenz	0-15 Jahre: RR: 2,7 [KI:0.6-12.1] 15-24 Jahre: RR: 2,2 [KI:1.3-3.6]	Schilddrüsenkrebs (<15 Jahre) (15-24 Jahre)
Ivanov et al. (1997)	Estland	1981-1995	Inzidenz	Keine erhöhte Inzidenz	Schilddrüsenkrebs, (Erwachsene)
Studien in den direkt betroffenen Gebieten Weißrussland, Ukraine und Russland					
Bleuer et al. (1997)	Belarus	1990-1994			Schilddrüsenkrebs (0-14 Jahre)
Buglova et al. (1996)	Weißrussland exponierte Gebiete	1990-1992	Inzidenz	Bei Jungen höhere Inzidenzraten als erwartet bei unterschiedlich hoher Exposition.	Schilddrüsenkrebs
Heidenreich et al. (1999)	Weißrussland, exponierte Gebiete	1986-1995	Inzidenz	Risiko von Kindern <10 Jahre bei Exposition ist 3 mal höher als von älteren Kindern Risiko der Mädchen ist 1,5 mal höher als das der Jungen	Schilddrüsenkrebs: (Alterskohorte < 19 Jahre bei Unfall)
Ivanov et al. (1999)	Russland	1982-1995	Inzidenz	0-4 Jahre bei Exposition/1991-1995: SIR= 8.0 (s.) gegenüber > 15 Jahren	Schilddrüsenkrebs (0-4 Jahre bei Exposition)
Ivanov et al. (1999)	Russland	1982-1995	Inzidenz	1991-1995 höhere Inzidenzen als 1986-1990	Schilddrüsenkrebs (Erwachsene)
Ivanov et al. (2003)	Bryansk (Russland)	1986-1998	Inzidenz	1986-1990: SIR: 1.27 [KI: 0.92-1.73]; 1991-1998 Männer: SIR: 1.45 [KI: 1.20-1.73]; 1991-1998 Frauen: 1.94 [KI: 1.82-2.10]; ERR/Gy : 0.7	Schilddrüsenkrebs (15-69 Jahre)
Jacob et al. (1999, 2000)	exponierte Gebiete Weißrussland Russland	1991- 1995	Inzidenz	EAR/10 ⁴ PY *Gy: 2.1 (KI: 1.00 – 4.50)	Schilddrüsenkrebsinzidenz Kinder, (4 Monate bis 15 Jahre bei Exposition)

Mettler et al. Ukraine (1992)	1990	Inzidenz	Keine erhöhten Inzidenzen	Schilddrüsenknoten (Erwachsene)
Prisyazhiuk et al. Ukraine (1995)	1980-1993	Inzidenz	Keine erhöhte Inzidenz	Schilddrüsenkrebs (Erwachsene)
Shibata et al. Belarus (2001)	1998-2000		OR Gruppe 3) gegenüber 2): 1.21 [KI: 0.9-3.10]; gegenüber 1): 1.1 [KI: 0.3-1.76]	Schilddrüsenkrebs (< 7 Jahre in 3 Gruppen: 1): geboren 1.1.1987-31.12.1989; 2):27.4.1986-31.12.1986; 3) 1.1.1983-26.4.1986
Sobolev et al. Ukraine (1997)	1991-1995	Inzidenz	1): erhöhte Inzidenzen; 2): erhöhte Inzidenz nur in stark kontaminierten Gebieten	Schilddrüsenkrebs bei Kindern 1) < 9 Jahre 2) < 17 Jahre
Tronko et al. Ukraine (1999)	1981-1997	Inzidenz	1990-1997: erhöhte Inzidenz; besonderer Anstieg bei Kindern 0-5 Jahre im Jahre 1986; Gebiete mit Dosis > 0.5 Gy größte Inzidenz	Schilddrüsenkrebs bei Kindern (< 15 Jahre) und Jugendlichen (15-18 Jahren)

¹ EAR: excess absolute risk

In Tabelle 11 und in Tabelle 12 sind insgesamt 33 Veröffentlichungen zu ökologischen Studien nach dem Tschernobylunfall dargestellt. Überwiegend wurden Schilddrüsenkrebs und Leukämie bei Kindern untersucht, den beiden Tumorarten, die nach den Atombombenabwürfen auf Hiroshima und Nagasaki zuerst anstiegen.

Studiendesign

Diese Studien sind überwiegend als rein deskriptive Studien durchgeführt worden. In Ländern mit Krebsregistern wurde der zeitliche Verlauf der Inzidenzraten vor und nach 1986 beobachtet. Einige Studien, besonders in den betroffenen Ländern Weißrussland, Ukraine und Russland, wurden die Erkrankungsraten hochexponierter und weniger exponierter Gebiete verglichen. Die Studien von Jacob et al. (1999, 2000) nehmen eine quantitative Risikoschätzung vor. In der Einteilung von Morgenstern (1998) sind diese Studien dem ökologisch deskriptiven *time-trend-design* oder *multiple-group-design* zuzuordnen.

Exposition:

Die Exposition gegenüber radioaktivem Fallout nach dem Tschernobyl-Unfall ist zeitlich festgelegt. Die gesamte Bevölkerung eines Landes wurde gleichzeitig überwiegend in den ersten Tagen und Wochen nach dem Unfall exponiert. Deswegen ist z.B. für Kinderkrebs eine einfache Einteilung in exponierte Geburtskohorten, die in einem gewissen Zeitraum vor dem Unfall geboren waren und nichtexponierte Geburtskohorten, die nach dem Unfall geboren wurden, möglich. Auch geographisch ist in den meisten europäischen Ländern eine sehr grobe Einteilung in stärker und weniger stark kontaminierte Gebiete möglich. Die Exposition ist aber nur auf Gruppenebene bekannt. Ihr liegt eine unbekannt Variation auf individueller Ebene zugrunde.

Erkrankung:

Die Inzidenzen der Erkrankungen wurden bei allen Studien aus regionalen oder nationalen Krebsregistern entnommen. Die Schätzung der erwarteten Fälle wird bei den verschiedenen Studien unterschiedlich durchgeführt. Die Studien in den Nord- und Westeuropäischen Ländern greifen auf die Daten ihrer Krebsregister vor dem Unfall zurück. Für die betroffenen Länder in Osteuropa wurden Krebsregister aber erst in den Jahren nach dem Unfall eingerichtet. Diese Studien benutzen entweder die Inzidenz der Zeit direkt nach dem Unfall (1986 bis 1990) als Basis zur Berechnung der erwarteten Fälle oder schätzen diese aus den Inzidenzen anderer Europäischer Länder. Besonders für die niedrigen und stark schwankenden Schilddrüsenkrebsinzidenzen bei Kindern kann nur eine grobe Schätzung gemacht werden.

Ergebnisse:

Ein Anstieg der Schilddrüsenkrebsinzidenzen unter exponierten Kindern und Jugendlichen in den betroffenen Gebieten in Weißrussland, der Ukraine und Russland zeigt sich durchgehend in allen Untersuchungen. Außerhalb dieser Gebiete wurde in England ebenfalls ein signifikanter Anstieg der Schilddrüsenkrebsinzidenzen bei Jugendlichen nachgewiesen, es fehlt aber ein Zusammenhang zwischen der Höhe der radioaktiven Kontamination und den Inzidenzen in einzelnen Gebieten.

Die Inzidenz von Kinderleukämie wurde außerhalb der betroffenen Gebiete von Parkin (1992, 1993, 1996) für ganz Europa analysiert. Zusätzlich wurden Studien in Finnland, Schweden, Deutschland, der Türkei, Griechenland und Polen durchgeführt. Obwohl die Inzidenzen im Beobachtungszeitraum in einigen Bereichen anstiegen, konnte kein Zusammenhang mit der Höhe der Kontamination durch den Tschernobyl-Unfall hergestellt werden. Eine Ausnahme stellen die Ergebnisse von Petridou et al. (1996) in Griechenland dar. Ausschließlich bei der

in Utero exponierten Geburtskohorte wurde ein erhöhtes Relatives Risiko in den ersten Lebensjahren beobachtet.

Von den sieben in Tabelle 11 dargestellten Studien zu Leukämie in den höher kontaminierten Gebieten in Weißrussland, der Ukraine und Russland fand nur Noshchenko et al. (2001) einen signifikanten Anstieg der Leukämie-Inzidenzen zwischen 1986 und 1996 in der Ukraine. Für Erwachsene konnte kein Anstieg der Inzidenzen von Schilddrüsenkrebs, Leukämie oder anderen Krebslokalisationen nachgewiesen werden, der mit der Exposition in Zusammenhang steht.

Diskussion

Die Ergebnisse der ökologischen Studien nach dem Tschernobylunfall können durch verschiedene Fehlerquellen beeinflusst worden sein. Bei den meisten Studien handelt es sich um *time-trend-designs*, die besonders vulnerabel gegenüber Veränderungen in der Krebsregistrierung sind. Sowohl die Qualität der medizinischen Diagnostik als auch die Aufmerksamkeit, die den möglichen Folgeerkrankungen nach dem Tschernobylunfall gezollt wurde, können sich über die Zeit stark verändert haben und sowohl zu Über- als auch zu Unterschätzung der wahren Inzidenzrate geführt haben. In den betroffenen Gebieten ist die Veränderung in der Krebsregistrierung besonders auffällig.

Die weiter oben zu den Atomkraftwerkstudien angeführten Einschränkungen für ökologische Studien gelten auch für die ökologischen Studien nach dem Tschernobylunfall. Ökologischer Bias durch Screening und weitere mögliche Confounder wurden bereits in den vorausgegangenen Abschnitten diskutiert.

2.3.4 Zusammenfassung

Seit 1990 sind in der Epidemiologie ionisierender Strahlung weniger Studien durchgeführt worden als in den achtziger und neunziger Jahren des letzten Jahrhunderts. Zu vielen Fragestellungen wurden zunehmend gute Daten auf individueller Ebene verfügbar, so dass Fall-Kontroll und Kohortenstudien durchgeführt werden konnten. Aufgrund der intensiven theoretischen Diskussion um dieses Studiendesign ist die Sensibilität gegenüber den Einschränkungen ökologischer Studien groß. Ökologische Studien werden selten zitiert, ohne dass auf ihre eingeschränkte Validität hingewiesen wird.

Ökologische Studien wurden hauptsächlich zu den Fragestellungen: „Kinderkrebs in der Umgebung von kerntechnischen Anlagen“ und „Krebserkrankungen nach dem Tschernobylunfall“ durchgeführt. Einige Studien wurden zusätzlich zu den Fragestellungen der Auswirkung erhöhter natürlicher Hintergrundstrahlung und der Auswirkungen der Atombombenversuche bzw. anderer radioaktiver Umweltverschmutzung durchgeführt.

Ökologische Studien in der Epidemiologie ionisierender Strahlung beschränkten sich in den letzten 12 Jahren weitgehend auf qualitative Studien, in denen Inzidenzraten auf signifikante Unterschiede hin untersucht wurden. Die Strahlendosis ging mit wenigen Ausnahmen nicht in die Berechnung mit ein. Die Exposition wurde durch eine einfache räumliche oder zeitliche Zuordnung ersetzt. Die Studien von Jacob et al. (1999, 2000) stellen eine Ausnahme dar, weil ein quantitativer Zusammenhang zwischen einer Doseinheit und dem Risiko zu erkranken bestimmt wurde.

3 Simulationen zu ökologischen Studien

Nachdem im vorigen Abschnitt die methodischen Grundlagen und der Wissensstand zu ökologischen Studien in der Literatur dokumentiert wurden, stellt dieser Abschnitt eigene Forschungsergebnisse vor, die mit Computersimulationen gewonnen wurden. Sie beziehen sich im Wesentlichen auf den „pure specification bias“, der ein grundsätzliches Problem in jeder ökologischen Studie darstellt, selbst in Abwesenheit von Confounder-Variablen. Als wesentliche Auswertemodelle für zusammengefasste Daten werden die logistische Regression und später auch in Abschnitt 5 die Poissonregression eingesetzt. Individuelle Daten können mit einer entsprechenden Likelihood-Funktion oder mit Cox-Modellen analysiert werden. Während die Likelihood mit individuellen Daten mit vertretbarem Rechenaufwand gute Ergebnisse liefert (s. Abschnitt 5), ist der Mehraufwand bei der Umsetzung von Cox-Modellen erheblich. Auf ihren Einsatz wurde deshalb im Rahmen dieses Projektes verzichtet.

3.1 *Pure specification bias*

Einige der Probleme in ökologischen Studien werden dadurch verursacht, dass häufig die Expositionen, denen die Individuen ausgesetzt sind, innerhalb einzelner Regionen stark variieren (*within-area variability*). In diesem Fall lässt sich ein auf individueller Ebene verwendetes Risikomodell nur dann ohne Verzerrung auf die Gruppenebene übertragen und zur Analyse aggregierter Daten verwenden, wenn es linear ist (Salway & Wakefield, 2003). Für andere Risikomodelle (siehe Abschnitt 2.1) gilt dies nur, wenn es keine (oder nur eine sehr geringe) Variabilität der Exposition innerhalb der Regionen gibt.

Allerdings sind zur Untersuchung seltener Erkrankungen, wie z.B. Schilddrüsenkrebs, multiplikative Modelle – die als Link-Funktion den natürlichen Logarithmus oder den Logit besitzen – besser geeignet (Wakefield & Salway, 2001) und in der Epidemiologie üblich. Verwendet man aber diese nicht-linearen Risikomodelle auch auf Gruppenebene kann – bei variabler Exposition innerhalb einzelner Gruppen – der sogenannte *Pure Specification Bias* auftreten. Diese Verzerrung liegt auch dann vor, wenn weder Confounder noch Missklassifikation eine Rolle spielen. Der Ausdruck *pure specification bias* wurde 1992 von Greenland geprägt. Die Verwendung eines Risikomodells zur Analyse aggregierter Daten ist also nur dann ratsam, wenn ein linearer Zusammenhang zwischen individueller Exposition und individueller Erkrankungswahrscheinlichkeit unterstellt werden kann oder die Exposition innerhalb der Gruppen konstant ist. In diesem Fall repräsentiert die mittlere Exposition einer Gruppe auch die Exposition jedes Individuums. Im Folgenden wird die Herleitung der entsprechenden Formeln dargestellt.

Betrachtet werden m Gruppen mit n_j Individuen in Gruppe j ($j=1, \dots, J$). Der Krankheitsstatus von Individuum i ($i=1, \dots, n_j$) aus Gruppe j wird repräsentiert durch Y_{ij} mit den Ausprägungen 1 für den Zustand *krank* und 0 für den Zustand *gesund*, die individuelle Erkrankungswahrscheinlichkeit wird mit p_{ij} bezeichnet. x_{ij} steht für die individuelle Exposition von Individuum i aus Gruppe j und α für die Baseline-Erkrankungswahrscheinlichkeit. Für den Zusammenhang zwischen individueller Exposition und individueller Erkrankungswahrscheinlichkeit werden in den drei betrachteten Risikomodellen folgende Zusammenhänge angenommen: (1) Lineares Risikomodell auf individueller Ebene; (2) Exponentielles Risikomodell auf individueller Ebene, (3) Logistisches Risikomodell auf individueller Ebene. Stehen die Daten nur auf Gruppenebene zur Verfügung, sind zum einen

nur die Erkrankungsraten λ_j und zum anderen nur die durchschnittlichen Expositionen $\bar{x}_{.j}$ pro Gruppe bekannt. Die drei Risikomodelle, die zur Analyse der aggregierten Daten verwendet werden können, haben damit folgende Gestalt:

Lineares Risikomodell auf Gruppenebene:

$$\lambda_j = \alpha_{eco} + \beta_{eco} \bar{x}_{.j} \quad (3.1)$$

Exponentielles Risikomodell auf Gruppenebene:

$$\lambda_j = \exp(\alpha_{eco} + \beta_{eco} \bar{x}_{.j}) \quad (3.2)$$

Logistisches Risikomodell auf Gruppenebene:

$$\lambda_j = \exp(\alpha_{eco} + \beta_{eco} \bar{x}_{.j}) / [1 + \exp(\alpha_{eco} + \beta_{eco} \bar{x}_{.j})] \quad (3.3)$$

Im linearen Risikomodell liefert die Analyse auf Gruppenebene einen unverzerrten Schätzer des individuellen Risikoparameters, da folgender Zusammenhang gilt:

$$\begin{aligned} \frac{1}{n_j} \sum_{i=1}^{n_j} (\alpha_{ind} + \beta_{ind} x_{ij}) &= \alpha_{eco} + \beta_{eco} \left(\frac{1}{n_j} \sum_{i=1}^{n_j} x_{ij} \right) \\ \Leftrightarrow \frac{1}{n_j} \sum_{i=1}^{n_j} \alpha_{ind} + \frac{1}{n_j} \sum_{i=1}^{n_j} \beta_{ind} x_{ij} &= \alpha_{eco} + \frac{1}{n_j} \sum_{i=1}^{n_j} \beta_{eco} x_{ij} \end{aligned} \quad (3.4)$$

Für das exponentielle Risikomodell dagegen gilt die Gleichheit in der Regel nicht:

$$\frac{1}{n_j} \sum_{i=1}^{n_j} [\exp(\alpha_{ind} + \beta_{ind} x_{ij})] \neq \exp \left[\alpha_{eco} + \beta_{eco} \left(\frac{1}{n_j} \sum_{i=1}^{n_j} x_{ij} \right) \right] \quad (3.5)$$

Die Gleichheit trifft in der Regel auch für das logistische Risikomodell nicht zu:

$$\frac{1}{n_j} \sum_{i=1}^{n_j} \left[\frac{\exp(\alpha_{ind} + \beta_{ind} x_{ij})}{1 + \exp(\alpha_{ind} + \beta_{ind} x_{ij})} \right] \neq \frac{\exp \left[\alpha_{eco} + \beta_{eco} \left(\frac{1}{n_j} \sum_{i=1}^{n_j} x_{ij} \right) \right]}{1 + \exp \left[\alpha_{eco} + \beta_{eco} \left(\frac{1}{n_j} \sum_{i=1}^{n_j} x_{ij} \right) \right]} \quad (3.6)$$

Die Hauptursache für den *pure Specification Bias* ist die Variabilität der Exposition *innerhalb* der Gruppen (Salway & Wakefield (2003)) und das Verhältnis dieser zur Variabilität der Exposition *zwischen* den Gruppen.

3.2 Zweck der Untersuchung

Mit Hilfe der Simulationsstudie wird untersucht, wie sich *der pure specification bias* bei der Auswertung aggregierter Daten auswirkt, deren zugrunde liegende Individualdaten einem logistischen Risikomodell folgen. Besonderes Augenmerk gilt dabei dem Einfluss der Varianz zwischen den Gruppen und der Varianz innerhalb der Gruppen, da nach einer Aussage von Salway & Wakefield (2003) der Pure Specification Bias um so kleiner ist, je größer der Quotient aus diesen beiden Varianzen ist.

3.3 Simulationskonstellationen

Mittels SAS wurden zunächst log-normalverteilte, individuelle Expositionen x_{ij} für je 400 Individuen ($i=1, \dots, 400$) in 51 Gruppen ($j=1, \dots, 51$) generiert. Die Berechnung der individuellen Erkrankungswahrscheinlichkeiten erfolgte auf Grundlage des logistischen Risikomodells mit der Formel:

$$p_{ij} = \frac{\exp(\alpha_{ind} + \beta_{ind} x_{ij})}{1 + \exp(\alpha_{ind} + \beta_{ind} x_{ij})} \quad (3.8)$$

Als Baseline-Erkrankungswahrscheinlichkeiten p_0 wurden die Werte 0,01, 0,05 und 0,10 betrachtet, so dass gilt: $\exp(\alpha_{ind}) = p_0$. Als Werte für β_{ind} wurden 1,00, 1,25 und 1,50 gewählt. Auf Grundlage der so berechneten individuellen Erkrankungswahrscheinlichkeiten der insgesamt 20.400 Individuen wurden die Kranken bzw. Gesunden mittels der SAS-Zufallszahlenfunktion RANBIN als Bernoulli-verteilte Zufallsvariable erzeugt.

Um den Einfluss der Expositionsvarianz innerhalb der Gruppen und zwischen den Gruppen auf den *pure specification bias* zu untersuchen, wurden bei der Generierung der individuellen Expositionen außerdem die Werte für diese Varianzen vorgegeben. Dabei wurden für die Varianzen innerhalb der Gruppen die Werte 0,01, 0,25, 1, 5 sowie 10 und für die Varianzen zwischen den Gruppen die Werte 0,02, 0,14, 0,54, 1,22 und 2,17 verwendet. Insgesamt konnten damit 25 verschiedene Konstellationen aus Varianz innerhalb der Gruppen und Varianz zwischen den Gruppen betrachtet werden. Bei der Vorgabe der Varianzen wurde darauf geachtet, dass für alle 25 Konstellationen der Expositions Mittelwert über alle 51 Gruppen gleich groß – nämlich 2,51 – blieb. Die betrachteten Varianzen zwischen den Gruppen sowie der entsprechende kleinste (in Gruppe 1), mittlere (in Gruppe 26) und größte (in Gruppe 51) Expositions Mittelwert sind in Tabelle 13 aufgelistet.

Tabelle 13: Varianz innerhalb und zwischen den Gruppen für die Simulationsstudie

Varianz <i>zwischen</i> den Gruppen	Mittlerer Expositions-mittelwert (in Ort 1)	Mittlerer Expositions-mittelwert (in Ort 26)	Größter mittlerer Expositions-mittelwert (in Ort 51)
0,02	2,260	2,51	2,760
0,14	1,885	2,51	3,135
0,54	1,260	2,51	3,760
1,22	0,635	2,51	4,385
2,17	0,010	2,51	5,010

Aus dem Datensatz mit den 20.400 individuell generierten Beobachtungen wurde anschließend durch Berechnung der Mittelwerte über die individuellen Expositionen und Berechnung der Erkrankungsrate (Anzahl Kranke pro Gruppe dividiert durch Anzahl Individuen in der entsprechenden Gruppe) je Gruppe ein aggregierter Datensatz mit je 51 Beobachtungen erstellt. Außerdem wurden für jede Gruppe der natürliche Logarithmus sowie der Logit der Erkrankungsrate berechnet.

3.4 Datenauswertung

Auf Grundlage der so generierten Datensätze mit individuellen und aggregierten Daten wurden drei Analysen durchgeführt. Die individuellen Daten wurden mittels der SAS-Prozedur PROC LOGISTIC ausgewertet, wobei die binäre Variable, die den Krankheitszustand repräsentiert, als abhängige Variable und die individuellen Expositionen als unabhängige Variable verwendet wurden. Die aggregierten Daten wurden zweifach ausgewertet: Zum einen in der Annahme, dass der Zusammenhang zwischen individuellen Expositionen und Erkrankungswahrscheinlichkeiten einem logistischen Risikomodell folgt, und zum anderen in der Annahme, dass dem Zusammenhang ein exponentielles Risikomodell zugrunde liegt. Die erste Annahme ist plausibel, da die Individualdaten in der Tat auf Grundlage des logistischen Risikomodells generiert worden sind. Allerdings ist nach den hier gemachten Ausführungen offensichtlich, dass die Parameter nicht exakt übereinstimmen. Die zweite Annahme wird gemacht, da häufig in ökologischen Studien ein exponentielles Risikomodell untersucht wird, während bei der Auswertung von individuellen Daten meistens von einem logistischen Modell ausgegangen wird.

Dabei ergibt sich allerdings das Problem, dass die Risikoparameter der verschiedenen Risikomodelle nicht ohne weiteres miteinander vergleichbar sind. So stellt im linearen Modell der Risikoparameter die Erhöhung der abhängigen Variablen bei Erhöhung der unabhängigen Variablen um eine Einheit dar, während im exponentiellen bzw. logistischen Risikomodell der Exponent des Risikoparameters als Relatives Risiko bzw. Odds Ratio zu interpretieren ist (siehe Abschnitt 2.1). Da das lineare Risikomodell damit der Klasse der additiven Modelle, das exponentielle und logistische Risikomodell aber der Klasse der multiplikativen Modelle zuzuordnen sind, lässt sich das lineare Modell mit den anderen beiden Modellen nicht vergleichen. Daher wurden die aggregierten Daten, deren zugrunde liegende Individualdaten einem logistischen Risikomodell folgen, *nicht* mit einem ökologisch-linearen Risikomodell ausgewertet. Die Vergleichbarkeit der beiden multiplikativen Modelle ist nur unter dem Aspekt gewährleistet, dass bei kleinen Erkrankungswahrscheinlichkeiten das Relative Risiko des exponentiellen Modells dem Odds Ratio des logistischen Modells entspricht.

Da im logistischen bzw. exponentiellen Risikomodell für den Zusammenhang zwischen individueller Exposition und individueller Erkrankungswahrscheinlichkeit Folgendes gilt

$$p_{ij} = \exp(\alpha_{ind} + \beta_{ind} x_{ij}) \Leftrightarrow \log(p_{ij}) = \alpha_{ind} + \beta_{ind} x_{ij} \quad (3.9)$$

lassen sich – in Analogie zu den entsprechenden individuellen Modellen – aggregierte Daten nur anhand der folgenden Gleichungen analysieren:

$$\text{logit}(\lambda_j) = \alpha_{eco} + \beta_{eco} \bar{x}_{.j} \quad (3.10)$$

bzw.

$$\log(\lambda_j) = \alpha_{eco} + \beta_{eco} \bar{x}_{.j} \quad (3.11)$$

Dabei bezeichnet λ_j die Erkrankungsrate (die sich als Mittelwerte über die mit 1 bzw. 0 codierten Kranken bzw. Gesunden ergeben) und $\bar{x}_{.j}$ die durchschnittliche Exposition in Gruppe j . Mit diesen Vorgaben wurden die aggregierten Daten im Rahmen einer linearen Regression mittels der SAS-Prozedur PROC REG ausgewertet, wobei der Logit bzw. der natürliche Logarithmus der Erkrankungsrate die abhängige Variable und die durchschnittliche Exposition je Gruppe die unabhängige Variable darstellten.

Ziel der Analysen war zu überprüfen, inwieweit der bei der Datengenerierung zugrunde gelegte („wahre“) Wert für den Regressionsparameter $\beta_{true} = \beta_{ind}$ in den zwei Analysen reproduzierbar ist. (Für die individuelle Analyse ist die Erwartungstreue des Schätzers β_{true} bekannt und wird nicht weiter untersucht.

Simulationen auf Grundlage des linearen und exponentiellen Risikomodells wurden nicht durchgeführt. Für das lineare Risikomodell kann analytisch gezeigt werden, dass es keinen *pure specification bias* gibt.

Für alle Kombinationen aus Baseline-Erkrankungswahrscheinlichkeit und „wahren“ Risikoparameter sowie für jede der 25 Konstellationen wurden 1000 Simulationen durchgeführt. Berechnet wurden anschließend der Mittelwert über die 1000 geschätzten Regressionsparameter $\frac{1}{1000} \sum \hat{\beta}_{eco}$, die mittlere quadratische Abweichung vom „wahren“ Wert des Regressionsparameters $\frac{1}{1000} \sum (\beta_{true} - \hat{\beta}_{eco})^2$ sowie der prozentuale Anteil der Simulationen, bei denen der „wahre“ Risikoparameter β_{true} im 95%-Konfidenzintervall des geschätzten Parameters $\hat{\beta}_{eco}$ liegt. Der Pure Specification Bias ist dann als gering anzusehen, wenn der Mittelwert über die 1000 Simulationen durchläufe nahe am wahren Wert des Risikoparameters liegt, wenn die mittlere quadratische Abweichung nahe Null ist und wenn der Prozentsatz der Durchläufe, bei denen sich der „wahre“ Regressionsparameter im 95%-Konfidenzintervall der geschätzten Parameter befindet, beinahe 100 beträgt.

3.5 Ergebnisse

Im Anhang 7 werden in den Tabellen 1 bis 18 die Ergebnisse der Simulationen im Detail dargestellt. Die Ergebnisse der Auswertungen der aggregierten Daten mit dem logistischen Risikomodell (siehe Tabellen 1 bis 9) zeigen, dass bei fester Varianz zwischen den Gruppen der Mittelwert der 1000 ökologischen Parameterschätzer umso weiter vom wahren Wert des Risikoparameters abweicht – und damit der *pure specification bias* um so größer wird – je größer die Varianz innerhalb der Gruppen wird. Dies gilt – mit wenigen Ausnahmen entsprechend für die mittlere quadratische Abweichung der Parameterschätzer. Je größer die Varianz innerhalb der Gruppen ist, desto größer wird die mittlere quadratische Abweichung.

Betrachtet man jedoch für feste Varianzen innerhalb der Gruppen den Mittelwert der Parameterschätzer für größer werdende Varianzen zwischen den Gruppen, ergibt sich *kein* einheitliches Bild. Die Ergebnisse hängen auch noch von der Baseline-Erkrankungswahrscheinlichkeit und dem angenommenen (wahren) Wert der Risikoparameter ab. In den Abbildungen 1 bis 36 werden die individuellen Expositionen gegenüber den individuellen Erkrankungswahrscheinlichkeiten als auch die durchschnittlichen Expositionen gegen die Erkrankungsraten pro Gruppe dargestellt. Die Abbildungen korrespondieren also mit den Werten in den vier Ecken der Tabellen 1 bis 18. Bei „kleinen“ Varianzen entspricht der ökologische Parameterschätzer ziemlich genau dem „wahren“, individuellen Risikoparameter. Je größer die Werte der wahren Risikoparameter und Baseline-Erkrankungswahrscheinlichkeiten werden, desto größer werden auch die individuellen Erkrankungswahrscheinlichkeiten bzw. Erkrankungsraten je Gruppe.

Bei großer Varianz innerhalb der Gruppen und kleiner Varianz zwischen den Gruppen liegen nur die Gruppenmittelwerte nahe zusammen (siehe Abbildungen 4 bis 6, 16 bis 18 und 28 bis 30). Für diese Konstellation aus Varianz innerhalb der Gruppen und Varianz zwischen den Gruppen sind hinsichtlich der Parameterschätzungen die größten Abweichungen vom „wahren“ Wert des Risikoparameters zu verzeichnen, die um so größer werden, je größer die Baseline-Erkrankungswahrscheinlichkeiten sind.

Im Fall einer kleinen Varianz innerhalb der Gruppen und einer großen Varianz zwischen den Gruppen (siehe Abbildungen 7 bis 9, 19 bis 21 und 31 bis 33) ist zu beobachten, dass bedingt durch die große Varianz zwischen den Gruppen – genauso wie die aggregierten Daten – auch die Individualdaten relativ weit auseinander liegen. Dabei folgen die aggregierten Werte der durch die Individualdaten vorgegebenen Kurve. Entsprechend stimmen die Schätzungen für die ökologischen Risikoparameter ziemlich genau mit den individuellen Risikoparametern überein.

Bei großer Varianz innerhalb der Gruppen und großer Varianz zwischen den Gruppen (siehe Abbildungen 10 bis 12, 22 bis 24 und 34 bis 36) liegen aufgrund der großen Varianz zwischen den Gruppen die aggregierten Daten weit auseinander. Allerdings folgen sie nicht – wie im Fall einer kleinen Varianz innerhalb der Gruppen und einer großen Varianz zwischen den Gruppen – der von den individuellen Daten vorgegebenen Kurve. Stattdessen wird der Verlauf der aggregierten Daten von der Kurve der individuellen Daten geschnitten, wobei die Werte für (durchschnittliche) Exposition und Erkrankungswahrscheinlichkeit bzw. -rate am Schnittpunkt umso höhere Werte annehmen, je größer die Baseline-Erkrankungswahrscheinlichkeit ist.

Allen Abbildungen ist jedoch gemeinsam, dass sich für größer werdende Varianzen innerhalb der Gruppen die Punkte, die die aggregierten Daten darstellen, vertikal in Richtung x-Achse verschieben, während die horizontale Position in etwa gleich bleibt. Die aggregierten Daten werden also hinsichtlich der Erkrankungsraten immer kleiner, während die Werte für die durchschnittlichen Expositionen in etwa unverändert sind.

3.6 Fazit

Insgesamt lässt sich sagen, dass der *pure specification bias* umso größer wird, je größer die Baseline-Erkrankungswahrscheinlichkeit und der wahre Wert des Risikoparameters sind. Des

Weiteren wird bei gleich bleibender Varianz zwischen den Gruppen der Pure Specification Bias umso größer, je größer die Varianz innerhalb der Gruppen wird. Die Aussage von Salway & Wakefield 2003, nach der der *pure specification bias* um so kleiner wird, je größer der Quotient aus Varianz zwischen den Gruppen und Varianz innerhalb der Gruppen ist, lässt sich auf Grundlage der Simulationsergebnisse allerdings nur bedingt bestätigen.

Für die “Tschernobyl-Daten” lassen sich daraus folgende Schlüsse ziehen: Wenn, wie in der Strahlenepidemiologie häufig von einem linearen Modell ausgegangen werden kann, spielt der *pure specification bias* keine Rolle. Wenn von einem logistischen Modell ausgegangen wird, ist die Verzerrung gering, wenn die Baseline-Erkrankungswahrscheinlichkeit gering ist, wovon ausgegangen werden kann, da Schilddrüsenkrebs bei Kindern selten ist. Allerdings muss davon ausgegangen werden, dass die Varianz der Exposition zwischen den Regionen und die Varianz innerhalb der Regionen vergleichbar groß ist. Der *pure specification bias* (ohne Betrachtung von Confoundern, Messfehlern und Migration) kann also hier zu einer nicht vernachlässigbaren Verzerrung führen.

4 Schilddrüsenkrebs nach dem Tschernobylunfall

Nach dem Kernkraftwerksunfall in Tschernobyl im April 1986 breiteten sich Radionuklide durch die Luft aus und führten zu flächenhaften Kontaminationen großer Gebiete in Weißrussland, Russland und der Ukraine. Neben ^{137}Cs war vor allem ^{131}I für die Strahlenexposition der Bevölkerung verantwortlich. Dieses Radionuklid gelangte hauptsächlich mit der Nahrung in den menschlichen Körper und führte nach Anlagerung in der Schilddrüse dort zu erheblichen Strahlendosen. Als hochbelastetes Nahrungsmittel erwies sich Milch, die in höheren Mengen von Kindern und Jugendlichen getrunken wurde. Als Folge der Ingestion von ^{131}I stieg in der Geburtskohorte 1968-86 die Inzidenz von Schilddrüsenkarzinomen nach 1990 beständig an. In den meisten Ländern ist die Hintergrundinzidenz für diese Krebsart recht gering und schwankt zwischen 1–20 Fälle je Million Kinder und Jahr. Dieser Wert wurde in den Jahren nach 1990 in den betroffenen Gebieten beständig signifikant überschritten, wobei das zusätzliche Risiko im Wesentlichen auf drei Ursachen zurückzuführen ist. Dabei ist zunächst die Strahlenexposition zu nennen. Außerdem muß eine Anzahl von zusätzlichen Krebsfällen jedoch auch der verbesserten Nachweis- und Berichtsrate zugeschrieben werden. Verbesserungen dieser Rate wurden unter anderem durch den systematischen Aufbau von Kinderkrebsregistern und durch den Einsatz von leistungsfähigen Ultraschallgeräten in den Kliniken erzielt. Nicht zuletzt treten zusätzliche Fälle auch durch die Alterung der Geburtskohorte auf. Die Beiträge dieser drei Ursachen zur Inzidenzrate quantitativ zu erfassen, erweist sich als äußerst schwierig und soll in den beiden folgenden Abschnitten (und den dazugehörigen Anhängen) unternommen werden. Dazu wird im Abschnitt 4 die vorhandene Literatur analysiert und die Datengrundlage vorgestellt. In Abschnitt 5 werden dann die Ergebnisse von Simulationen diskutiert für Szenarios, die mit den vorhandenen Daten kompatibel sind und auf plausiblen Annahmen beruhen.

4.1 Schilddrüsenkrebsrisiko durch ^{131}I Exposition während der Kindheit

Allgemein ist der Kenntnisstand über das Schilddrüsenkrebsrisiko nach Strahlenexposition im Kindesalter gering.

Eine zusammenfassende Studie (Ron et al. 1995) von fünf Kohorten- und zwei Fall-Kontrollstudien zum Schilddrüsenkrebsrisiko nach externen Expositionen im Kindesalter ergab ein zusätzliches relatives Risiko je Schilddrüsendosis von 7,7 (95% Konfidenzintervall (KI): 2,1; 28,7) Gy⁻¹ und ein zusätzliches absolutes Risiko je Schilddrüsendosis von 4,4 (95% KI: 1,9; 10,1). Die Risikoschätzungen der einzelnen Studien unterschieden sich um mehr als eine Größenordnung. Da die einzelnen Studien unterschiedliche Altersverteilungen bei Exposition und unterschiedliche Beobachtungszeiten nach Exposition hatten, konnten in der zusammenfassenden Studie nur bedingt Aussagen über die Abhängigkeit vom Alter bei Exposition oder von der Zeit nach Exposition gemacht werden.

Studien zum Schilddrüsenkrebsrisiko nach ^{131}I Exposition im Kindesalter liegen für diagnostische Applikationen, Umweltkontaminationen durch Atomwaffentests und durch den Tschernobylunfall vor.

Die Ergebnisse von Studien nach diagnostischen Applikationen von ^{131}I im Kindesalter wurden in UNSCEAR Berichten zusammengefasst (UNSCEAR 1994, 2000). Die Studien ergaben keine signifikanten Aussagen zum Schilddrüsenkrebsrisiko. Eine neuere schwedische Studie (Dickman et al. 2003) schloss 2367 Personen mit Alter bei Exposition kleiner 20 ein.

Die drei beobachteten Schilddrüsenkrebsfälle ergaben keine erhöhte Inzidenz im Vergleich zur schwedischen Bevölkerung: Es wurde ein ‚Standardized incidence ratio‘ (SIR) von 1,01 (95% KI: 0,21; 2,96) gefunden.

Unter 2473 Personen, die im Schulalter zwischen 1951 und 1958 durch Atomwaffentests auf dem Nevada Testgelände exponiert wurden, wurden in mehreren Untersuchungen im Zeitraum 1965 bis 1986 insgesamt acht Schilddrüsenkrebsfälle gefunden (Kerber et al. 1993). Das damit verbundene relative Risiko war nicht-signifikant erhöht.

Insgesamt haben die bisher vorliegenden Studien zum Schilddrüsenkrebsrisiko nach Expositionen durch Atomwaffentests oder nach diagnostischen Applikationen im Kindesalter eine geringe statistische Aussagekraft. Die starke Erhöhung der Schilddrüsenkrebsinzidenz in den durch den Tschernobylunfall kontaminierten Gebieten erlaubt eine Durchführung von Studien mit einer höheren statistischen Aussagekraft.

Zwei Fall-Kontrollstudien von Personen, die im Kindesalter oder während der Jugend durch den Tschernobylunfall exponiert wurden, haben eine eindeutige Assoziation zwischen der ^{131}I - Exposition und einer erhöhten Rate von Schilddrüsenkrebs festgestellt (Astakova et al. 1998, Davis et al. 2004). Allerdings erlaubten die Studien nur beschränkte Aussagen über die Höhe der Risikowerte.

Gegenwärtig wird eine Kohortenstudie mit 25 161 weißrussischen und ukrainischen Personen durchgeführt, die zum Zeitpunkt des Tschernobylunfalls im Kindesalter waren und deren ^{131}I – Gehalt in der Schilddrüse im Mai/Juni 1986 gemessen wurde (Stezhko et al. 2004). Obwohl von der Studie wesentliche Ergebnisse zum Schilddrüsenkrebsrisiko nach ^{131}I – Exposition im Kindesalter zu erwarten sind, wird sie wichtige Fragen zum Risiko in der Bevölkerung nicht beantworten können. Durch die intensiven Schilddrüsenuntersuchungen der Kohortenmitglieder ist die Krebsinzidenz im Vergleich zur Bevölkerung stark erhöht, d.h. Aussagen zum absoluten Risiko werden sich auf die Kohorte beziehen, aber nicht auf die Bevölkerung. Auch ist es nicht klar, inwieweit die intensiven Untersuchungen die Schätzungen des relativen Risikos beeinflussen. In einer Studie des Schilddrüsenkrebsrisikos nach Bestrahlung von gutartigen Erkrankungen in Kopf- und Halsbereich erniedrigte sich der Punktschätzer des relativen Risikos durch eine Intensivierung der Schilddrüsenuntersuchungen um einen Faktor 5 (Schneider et al. 1993). Die Erniedrigung war allerdings statistisch nicht signifikant.

Quantitative Abschätzungen des Schilddrüsenkrebsrisikos nach dem Tschernobylunfall wurden mit ökologischen Studien für Siedlungen durchgeführt, für die eine relativ gute Dosimetrie vorlag (Jacob et al. 1998; 1999; 2000). Für die Geburtskohorte 1971-85 ergab sich für den Beobachtungszeitraum 1991-95 ein zusätzliches relatives Risiko je Schilddrüsendosis von 23 (95% KI: 8,6; 82) Gy⁻¹ und ein zusätzliches absolutes Risiko je Schilddrüsendosis von 2,1 (95% KI: 1,0; 4,5). Die Ergebnisse stimmen innerhalb der Konfidenzgrenzen (zweiseitiger Gausstest) mit den Risiken überein, die in der zusammenfassenden Studie für externe Expositionen (siehe oben) beobachtet wurden. Allerdings erhebt sich die Frage, inwieweit die Ergebnisse der ökologischen Studien durch einen ökologischen Bias verfälscht sind.

4.1.1 Reihenuntersuchungen

Die Erhöhung der Schilddrüsenkrebsinzidenz in der Ukraine ist teilweise direkt der Strahlung durch den Tschernobylunfall und teilweise einer erhöhten Nachweis- und Berichtsrate von Schilddrüsenkrebsfällen zuzuschreiben. Es erhebt sich die Frage, welchen Anteil Reihenuntersuchungen (mass screening) an der Erhöhung der Nachweis- und Berichtsrate haben.

Tabelle 14 fasst die veröffentlichten Daten von Reihenuntersuchungen zusammen, die in der Ukraine durchgeführt wurden. Demnach wurden in den Oblasts Chernihiv, Zhytomyr, und Kyiv in den Zeiträumen 1992-94 und 1996-2000 durch die Reihenuntersuchungen 5, bzw. 79 Schilddrüsenkrebsfälle gefunden. Die Prävalenzen in den Reihenuntersuchungen lagen im ersten Zeitraum bei 400 Fällen je 10^6 Personen, im zweiten Zeitraum bei 2400 Fällen je 10^6 Personen. Aus diesen hohen Prävalenzen kann geschlossen werden, dass die in den Reihenuntersuchungen gefundenen Fälle zum überwiegenden Teil ohne die Reihenuntersuchungen gar nicht oder erst einige Jahre später bekannt und gemeldet worden wären. In Anbetracht der Fallzahlen tragen somit die Reihenuntersuchungen vor 1998 nur zu einem unwesentlichen, seit 1998 jedoch zu mehr als 20% zu den im Krebsregister für die Oblasts Chernihiv, Kyiv (ohne die Stadt Kyiv) und Zhytomyr gemeldeten Fälle bei.

Tabelle 14: Anzahl und Prävalenz der Schilddrüsenkrebsfälle in Reihenuntersuchungen (mass screening), die in der Ukraine bis zum Jahr 2000 durchgeführt wurden. AR steht für Alter bei der Reihenuntersuchung, AU für Alter zur Zeit des Unfalls

Reihenuntersuchung	Gebiet	Altersgruppe	Zeitintervall	Fälle in der Reihenuntersuchung	Prävalenz (Fälle je 10^6 Personen)	Fälle im Krebsregister
IPHECA (WHO, 1996)	Kyiv and Zhytomyr	AR \leq 15	1992-1994	5	400	110
Sasakawa (Nikoforova et al., 2002) und Danilyuk et al., 2002)	Kyiv und Zhytomyr	AU \leq 18 AU \leq 14	1996-2000 1996-2000	25 11	2300 1300	310 67
Ukraine-USA cohort study (Tronko et al., 2003)	Kyiv ^a , Chernihiv and Zhytomyr	AU \leq 18	1998-2000	43	3200	197

^aEinwohner der Stadt Kyiv zum Zeitpunkt des Unfalls wurden nicht untersucht.

4.2 Literaturübersicht: Epidemiologie und Morphologie von Karzinomen der Schilddrüse

4.2.1 Epidemiologie und Risikofaktoren für Schilddrüsenkrebs

Nach dem Tschernobylunfall wurde nur für Schilddrüsenkrebs eine signifikante Erhöhung der Inzidenzraten beobachtet, während für andere Krebsformen keine Erhöhung durch Strahlenexposition festgestellt wurde. Zur Schätzung des Strahlenrisikos müssen das Grundrisiko in der nicht exponierten Bevölkerung und noch weitere Risikofaktoren in das Risikomodell eingehen. Im vierten Kapitel wird der Literaturübersicht der Wissensstand zu diesem Thema dokumentiert (Anhang 1). Zunächst wird eine Definition für Schilddrüsenkarzinome und deren Erscheinungsformen gegeben. Dann wird eine deskriptive

Epidemiologie der klinisch manifesten Tumore vorgestellt mit europaweiten Angaben zur Alters- und Geschlechtsverteilung. Grafische Abbildungen dokumentieren die Inzidenz- und Mortalitätsraten in Europa sowie deren zeitliche Trends. Interessant ist die relative günstige Prognose nach der einer klinischen Behandlung, die mit der Überlebensrate nach fünf Jahren quantifiziert wird.

Der größte Risikofaktor für Schilddrüsenkrebs dürfte die Strahlenexposition sein, die sowohl intern als auch extern erfolgen kann. Eine Reihe weiterer Faktoren werden in der Literatur diskutiert, die in der Literaturübersicht Erwähnung finden. Dazu gehören Ernährung, hormonale und reproduktive Faktoren bei Frauen, assoziierte Neoplasmen, genetische Disposition und sozioökonomischer Status.

4.2.2 Latente papilläre Karzinome der Schilddrüse

Die große Mehrheit von 98 % aller nach dem Tschernobylunfall registrierten Schilddrüsenkarzinome gehört zur papillären Variante. Schon früh wurde bemerkt, dass es sich bei einigen gefundenen Karzinomen um sehr kleine (< 1 cm), latente Karzinome handeln könnte, die unter Umständen klinisch nicht relevant geworden wären. Diese Tumoren werden auch als latente Mikrokarzinome bezeichnet. Zur Prävalenz dieser Mikrokarzinome wurde eine Literaturrecherche durchgeführt, deren Ergebnisse im zweiten Kapitel des Anhangs 1 dargestellt werden.

Es ist klar, dass durch diese Mikrokarzinome die Risikoschätzung stark beeinflusst werden kann, da ihr Auftreten von der Intensität und Gründlichkeit des Screening abhängt. Dieses Problem wird auch in den Simulationen zu diesem Projekt modellhaft untersucht. Allerdings stellte sich bei der Literaturrecherche heraus, dass quantitative Aussagen zur Prävalenz latenter Mikrokarzinome mit großen Unsicherheiten behaftet sind.

So lag die Prävalenz für papilläre Mikrokarzinome in den ausgewerteten Autopsiestudien zwischen 1,5% (Griechenland) und 35,6% (Finnland). Für die Altersgruppe 40 Jahre und jünger lag die Prävalenz zwischen 3,5% (Island) und 14% (Finnland).

Die Prävalenzschätzungen für papilläre Mikrokarzinome aus Autopsiestudien sind etwa 100mal größer als die Inzidenzraten für klinisch relevanten Schilddrüsenkrebs in den entsprechenden Ländern. Obwohl diese beiden Größen nicht direkt vergleichbar sind, deutet der große Unterschied darauf hin, dass die meisten der papillären Mikrokarzinome zu Lebzeiten einer Person keine gesundheitlichen Einschränkungen verursacht und deswegen nicht diagnostiziert wird.

4.3 Datengrundlagen

In den folgenden Abschnitten werden das Studiengebiet, die Zahl der registrierten Krebsfälle und die Dosisabschätzungen für die untersuchten Ortschaften beschrieben. Die Simulationen mit realistischen Ausgangsdaten basieren auf zwei verschiedenen Dosisabschätzungen: der *vorläufigen* Schätzung und der *endgültigen* Schätzung (Likhtarov et al. 2005). Der Grund für die Nutzung zweier voneinander abweichender Schätzungen besteht darin, dass die Daten der endgültigen Schätzung erst gegen Ende der Projektdauer verfügbar waren. Deswegen mussten die Simulationen zu den methodologischen Untersuchungen der Verzerrung von Risikoschätzern in ökologischen Studien zunächst mit der vorläufigen Schätzung

durchgeführt werden. Die Simulationen zu ausgewählten Screening-Szenarien konnten schon mit der endgültigen Schätzung erfolgen. Bei der Darstellung der Ergebnisse wird jedoch angegeben, mit welcher Dosisschätzung gearbeitet wurde.

4.3.1 Studiengebiet



Abbildung 4: Studiengebiet mit den ukrainischen Oblasts Zhytomyr, Kyiv und Chernihiv (schraffiert) sowie der Stadt Kyiv

Das Studiengebiet liegt in der Ukraine und umfasst die drei stark kontaminierten Oblasts Zhytomyr, Kyiv und Chernihiv sowie die Stadt Kyiv (Abbildung 4). Im Norden hat es eine gemeinsame Grenze mit Weißrussland. Das ehemalige Kernkraftwerk Chernobyl befindet sich ganz im Norden der mittleren Oblast Kyiv. Zur Risikogruppe gehören alle Mitglieder der Geburtskohorte 1968-85, die aus Siedlungen stammen, in denen mehr als zehn individuelle Aktivitätsmessungen der Schilddrüse vorgenommen worden sind. Im Jahre 1986 bestand die Geburtskohorte 1968-85 aus ca. einer Million exponierten Kindern und jungen Erwachsenen unter achtzehn Jahren. Circa 65% dieser Personen lebten in der Stadt Kyiv.

4.3.2 Personen und Krebsfälle in der vorläufigen Dosisschätzung

Die Datengrundlage der *vorläufigen* Schätzung umfasst 743 Ortschaften in den hoch kontaminierten Oblasts Zhytomyr, Kyiv und Chernihiv. In diesen Ortschaften wurden in den Monaten Mai und Juni des Jahres 1986 mehr als zehn individuelle Messungen der Schilddrüsensdosis an Kindern und jungen Erwachsenen der Geburtskohorte 1968-95 vorgenommen. Tabelle 15 zeigt die Ausgangsdaten der vorläufigen Dosisschätzung zusammengefasst für jede Oblast. Im Beobachtungszeitraum 1990-99 wurden in den drei Oblasts insgesamt 160 Fälle von Schilddrüsenkrebs registriert. Aus den individuellen Messungen wurde für jeden Geburtsjahrgang der geschlechtsspezifische geometrische Dosismittelwert in einer Ortschaft geschätzt, wobei sich die Verteilung der Dosis lognormal verhielt. Da die Dosismittelwerte nur schwach vom Geschlecht abhingen, wurden sie für jeden Geburtsjahrgang zusammengefasst. Als geometrische Standardabweichung der Dosisverteilung in einem Geburtsjahrgang wurde der Wert 2,2 angenommen. Die Mittelwerte der Dosis sinken mit ansteigendem Alter, weil damit die Masse der Schilddrüse wächst. Die arithmetischen Ortsdosismittelwerte für die 743 Ortschaften der vorläufigen Dosisschätzung werden in Abbildung 5 dargestellt. Die zwei größten Städte sind Zhytomyr und Chernihiv mit jeweils mehr als 80000 mit Risiko behafteten Personen, die Mehrzahl der betroffenen Ortschaften verfügt über einige Hundert Personen in der Geburtskohorte 1968-86. Die mittlere Bevölkerungsdosis für die 366397 Personen aus der Risikogruppe betrug 0,173 Gy.

Tabelle 15: Anzahl der Siedlungen, geschätzte Anzahl der exponierten Kinder im Jahr 1986, beobachtete Anzahl von Schilddrüsenkrebsfällen von 1990-99 und geschätzte arithmetische Dosismittelwerte in der vorläufigen Dosisschätzung

Region	Anzahl Siedlungen	Anzahl Kinder in 1986	Anzahl Fälle von 1990-99	Arithmetisches Dosismittel [Gy]
Chernihiv O.	255	118655	56	0,163
Kyiv O.	220	82900	54	0,279
Zhytomyr O.	268	164842	50	0,127
Gesamt	743	366397	160	0,173

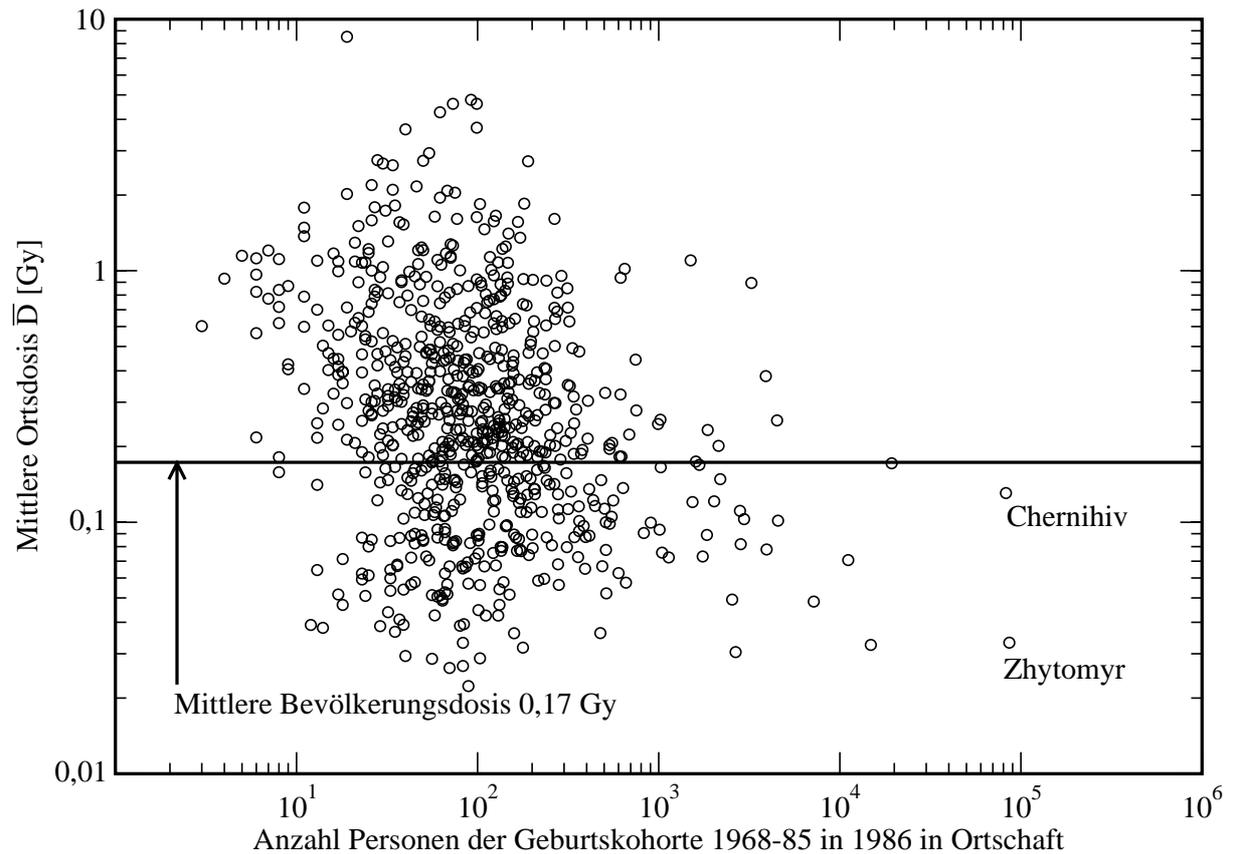


Abbildung 5: Arithmetische Dosismittelwerte für 743 Ortschaften als Funktion der Anzahl der Personen der Geburtskohorte 1968-85 im Jahre 1986 in der vorläufigen Dosisschätzung

4.3.3 Personen und Krebsfälle in der endgültigen Dosisschätzung

Für die endgültige Dosisschätzung wurden 669 Ortschaften ausgewählt, in denen mehr als 10 individuelle Dosismessungen im Mai und Juni 1986 durchgeführt wurden, und zwar jeweils für männliche und weibliche Personen getrennt. Deswegen gehören weniger Ortschaften zur endgültigen Dosisschätzung als zur vorläufigen Dosisschätzung. Hinzugenommen wurde die Hauptstadt Kyiv, die in den Simulationen als Oblast behandelt wird, die nur eine Ortschaft enthält. Tabelle 16 enthält die Ausgangsdaten der endgültigen Dosisschätzung, zusammengefasst auf der Ebene der Oblasts. Insgesamt werden 1002706 Personen erfasst mit einer mittleren Dosis von 0,08 Gy. Die arithmetischen Dosismittelwerte für die 670 Ortschaften zeigt Abbildung 6.

Tabelle 16: Anzahl der Siedlungen, geschätzte Anzahl der exponierten Kinder im Jahr 1986, beobachtete Anzahl von Schilddrüsenkrebsfällen von 1990-99 und geschätzte arithmetische Dosismittelwerte in der endgültigen Dosisschätzung

Region	Anzahl Siedlungen	Anzahl Kinder in 1986	Anzahl Fälle von 1990-99	Arithmetisches Dosismittel [Gy]
Chernihiv O.	222	119712	56	0,080
Kyiv Stadt	1	650078	198	0,044
Kyiv O.	203	74296	46	0,113
Zhytomyr O.	244	158620	50	0,122
Gesamt	670	1002706	350	0,080

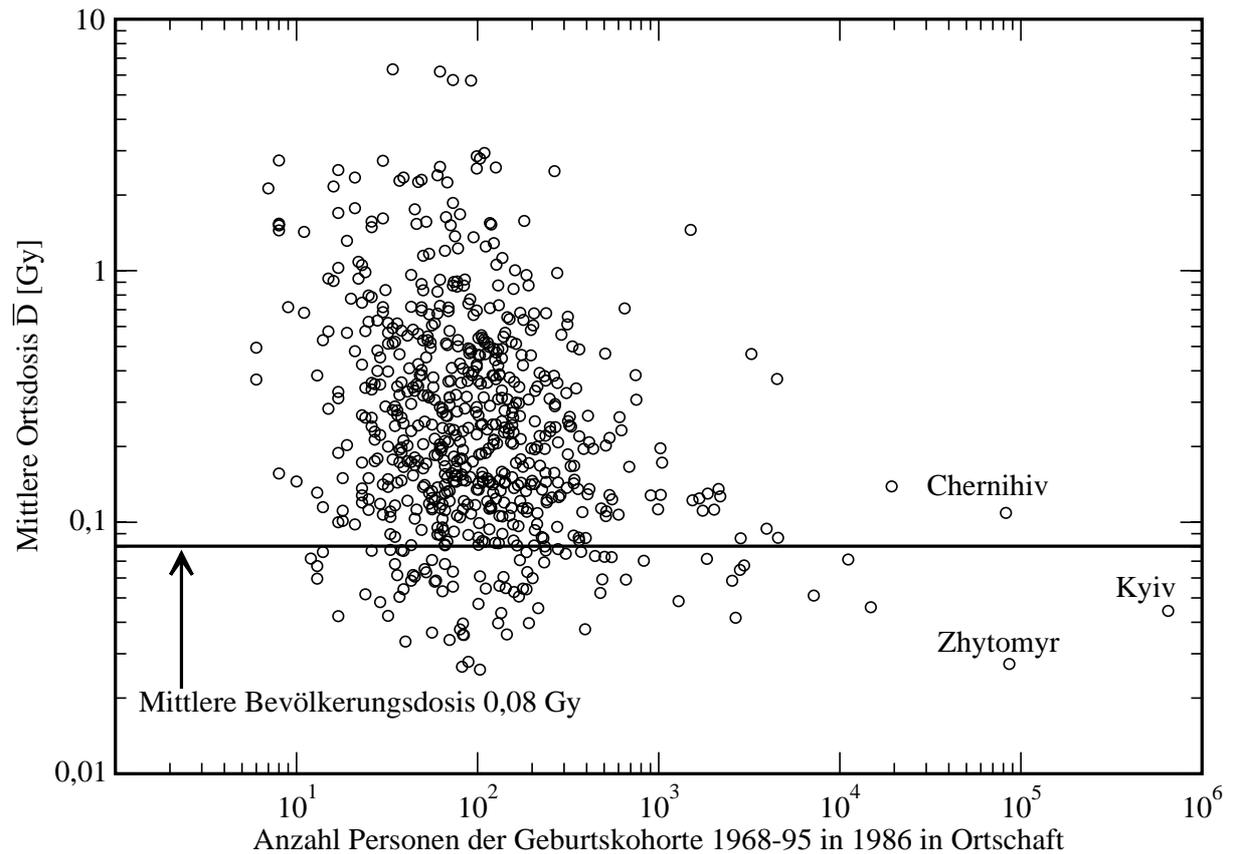


Abbildung 6: Arithmetische Dosismittelwerte für 670 Ortschaften als Funktion der Anzahl der Personen der Geburtskohorte 1968-85 im Jahre 1986 in der endgültigen Dosisschätzung

4.3.4 Abschätzung des Hintergrundrisikos

Mit der endgültigen Dosisschätzung aus Abschnitt 4.3.3 wurden realitätsnahe Screening-Szenarien simuliert. Dazu ist es notwendig, auch das Hintergrundrisiko möglichst genau zu schätzen. Man benötigt das oblast-spezifische Hintergrundrisiko

$$\bar{h}_{0k} = \frac{I_k^{90-99}}{(1 + \bar{\eta}_k^{90-99})(1 + \gamma \bar{D}_k)} \quad (2.1)$$

der Geburtskohorte 1968-85 für den Beobachtungszeitraum 1990-99. Dabei wird angenommen, dass es innerhalb einer Oblast k konstant ist. Es wird bestimmt durch Division der Inzidenz I_k^{90-99} im Zeitraum 1990-99 in einer Oblast k durch zwei Faktoren, die den Effekt von Strahlenexposition und von erhöhter medizinischer Überwachung beschreiben.

Der erste Faktor

$$1 + \bar{\eta}_k^{90-99} \quad (2.2)$$

entsteht durch die verbesserte Diagnose und Meldung von Fällen gegenüber dem Zeitraum vor dem Kernkraftwerksunfall. Dabei misst der zusätzliche Anstiegsfaktor

$$\bar{\eta}_k^{90-99} = \frac{4I_{0k}^{90-93} + 4I_{0k}^{94-97} + 2I_{0k}^{98-01}}{10I_{0k}^{86-89}} \quad (2.3)$$

den Anstieg der Hintergrundinzidenz im Zeitraum 1990-99 gegenüber dem Zeitraum 1986-89, der als Referenzzeitraum für eine Situation ohne Auswirkungen von Strahlung und Screening benutzt wird. Die Hintergrundinzidenzen für einen Zeitraum P von jeweils vier Jahren in einer Oblast k werden mit Gewichtungsfaktoren versehen, die der Jahresdauer entsprechen, mit der sie im Beobachtungszeitraum 1990-99 liegen. Sie wurden mit der von Jacob et al. (2004) empirisch gefundenen Beziehung

$$I_{0ks}^P = c_{ks}^P + d_{ks}^P a^{e_s} \quad (2.4)$$

berechnet. Hierbei sind c_{ks}^P , d_{ks}^P und e_s vom Geschlecht s abhängige Fitparameter, deren Zahlenwerte Jacob et al. (2004) zu entnehmen sind. Die Hintergrundinzidenz steigt mit dem Alter a deutlich an. Zur Berechnung der Hintergrundinzidenz wurde nicht die genaue Altersverteilung der Geburtskohorte verwendet sondern die Altersabhängigkeit wurde nur näherungsweise berücksichtigt. Die empirische Beziehung wurde mit dem mittleren Alter von 18 Jahren für die Geburtskohorte 1968-85 im Zeitraum 1990-99 ausgewertet. Zur weiteren Vereinfachung wurden die geschlechtsspezifischen Unterschiede vernachlässigt. Bei der Berechnung der geschlechtsunabhängigen Hintergrundinzidenz I_{0k}^P wurden die männlichen und weiblichen Beiträge gleich gewichtet.

Der zweite Faktor

$$1 + \gamma \bar{D}_k \quad (2.5)$$

beschreibt das relative Strahlenrisiko in einer Oblast k mit einer mittleren Dosis \bar{D}_k . Dabei bezeichnet der Faktor γ das Zusatzrisiko pro Einheitsdosis (engl. Excess Relative Risk Per unit Dose ERRPD). In den Simulationen wurden für das ERRPD zwei Werte angenommen, die zwei verschiedenen Simulationsvarianten entsprechen. Die zwei Werte von 10 Gy^{-1} (Variante 1) und 20 Gy^{-1} (Variante 2) sind motiviert durch Ergebnisse der Arbeiten von Ron et al. (1995) und Jacob et al. (1999). Tabelle 17 fasst die oblast-spezifischen Eingangsdaten für die Simulationen mit der endgültigen Dosisverteilung zusammen.

Tabelle 17: Schätzer für das Hintergrundrisiko \bar{h}_{0k} aus der Inzidenz I_k^{90-99} mit den zusätzlichen Anstiegsfaktoren $\bar{\eta}_k^{90-99}$, den Dosismittelwerten \bar{D}_k und einem ERRPD γ von 10 Gy⁻¹ (Variante 1) und 20 Gy⁻¹ (Variante 2)

	Inzidenz I_k^{90-99}	zusätzlicher Anstiegsfaktor $\bar{\eta}_k^{90-99}$	Dosismittel- wert \bar{D}_k	Hintergrundinzidenz \bar{h}_{0k}	
				Variante 1	Variante 2
Oblast	[10 ⁻⁶ PY ⁻¹]		[Gy]	[10 ⁻⁶ PY ⁻¹]	
Chernihiv	46,8	1,1	0,080	12,4	8,6
Kyiv Stadt	30,5	1,1	0,044	10,1	7,7
Kyiv Obl.	61,9	1,1	0,113	13,8	9,0
Zhytomyr	31,5	0,5	0,122	9,5	6,1
Mittelwert	34,9	1,0	0,080	10,5	7,7

4.3.5 Statistische Eigenschaften der zeitintegrierten Schilddrüsenaktivitäten aus den Dosisschätzungen der ersten und zweiten Kategorie

Das System zur Dosisrekonstruktion für die Gesamtbevölkerung der Ukraine, die zum Zeitpunkt des Unfalles jünger als 18 Jahre war, unterscheidet drei Kategorien der Dosisschätzung (Anhang 2).

- 1) Zur ersten Kategorie gehören Personen, deren Schilddrüsendosis auf der Grundlage von individuellen Messungen im Mai und Juni 1986 berechnet wurde, und für die zusätzlich Informationen über ihre Wohnorte und Essgewohnheiten vorlagen. Diese Informationen stammen idealerweise aus einer persönlichen Befragung. Konnte diese nicht durchgeführt werden, wurden Erfahrungswerte für Personen aus dem gleichen Wohnort benutzt.
- 2) In der zweiten Kategorie wurden alters- und geschlechtsspezifische Dosismittelwerte für Siedlungen berechnet, in denen nur direkte individuelle Messungen der Schilddrüsenaktivität vorlagen, aber sonst keine weiteren Informationen
- 3) Zur dritten Kategorie gehören Siedlungen, für die alters- und geschlechtsspezifische Dosismittelwerte berechnet wurden, ohne dass dort individuelle Schilddrüsenmessungen gemacht wurden.

Nur die Daten für Siedlungen der ersten und zweiten Kategorie sind genau genug und können für eine Risikoanalyse benutzt werden.

Die Mittelwerte der zeitintegrierten Schilddrüsenaktivitäten, die bei den Dosisschätzungen der ersten und zweiten Kategorie verwendet wurden, zeigen für jeden Geburtsjahrgang und beide Geschlechter eine gute Übereinstimmung. Eine Analyse der statistischen Verteilung der gemessenen tagesgenauen Schilddrüsenaktivitäten und der zeitintegrierten Aktivitäten mit dem Lilliefors-Test zeigt, dass die überwiegende Mehrzahl lognormal verteilt ist.

In der epidemiologischen Regressionsanalyse werden meist arithmetische Mittelwerte benutzt, während für die alters- und geschlechtsspezifischen Dosisschätzungen in einer Ortschaft zunächst geometrische Mittelwerte angegeben wurden. Zur Umrechnung der Mittelwerte wird ein zusätzlicher Formparameter für eine angenommene Lognormalverteilung benötigt. Dieser lässt sich aus dem Verhältnis von arithmetischem und geometrischem Mittelwert bestimmen. Für die tagesgenauen Schilddrüsenaktivitäten liegt das Verhältnis in dem engen Bereich von 1,35 bis 1,53. Werden diese Aktivitäten über die Zeit integriert, verbreitert sich der Bereich von 1,28 bis 1,61.

Schließlich wurde ein weiterer Vergleich der zeitintegrierten Schilddrüsenaktivitäten der ersten und zweiten Kategorie durchgeführt. Für jeden Geburtsjahrgang und für beide Geschlechter wurden zunächst für jede einzelne Siedlung der arithmetische und der geometrische Mittelwert für die zeitintegrierten Schilddrüsenaktivitäten der ersten Kategorie gebildet. Dann wurde für jeden dieser Mittelwerte das Verhältnis mit dem siedlungsbezogenen geometrischen Aktivitätsmittelwert der zweiten Kategorie gebildet. Für die siedlungsbezogenen Verhältnisse wurde wiederum der Mittelwert über alle Siedlungen gebildet. Wenn man das Verhältnis der geometrischen siedlungsbezogenen Mittelwerte aus der ersten und zweiten Kategorie betrachtet, lag für jeden Geburtsjahrgang und für beide Geschlechter der Mittelwert über alle Siedlungen immer nahe bei eins. Dies ist ein weiterer Hinweis auf die gute Übereinstimmung der Dosis-schätzungen mit Daten der ersten und zweiten Kategorie.

4.3.6 Morphologische Analyse der Schilddrüsenkrebsfälle in den hoch kontaminierten Gebieten der Ukraine

Dieser Abschnitt fasst die Ergebnisse des Anhangs 3 zusammen.

Die Siedlungen mit mehr als zehn Messungen der ^{131}I Aktivität in der menschlichen Schilddrüse, für die die vorliegende Simulationsstudie durchgeführt wurde, liegen in den drei Oblasts (Regionen) Chernihiv, Kyiv und Zhytomyr. Zusätzlich gehört die Stadt Kyiv zu den untersuchten Siedlungen. Für die Geburtskohorte 1968-85 und für den Operationszeitraum 1990-99 wurden dem klinisch-morphologischen Register am *Institut für Endokrinologie und Metabolismus der Akademie der medizinischen Wissenschaften der Ukraine* (Tronko et al. 1999, 2002a, 2002b) 191 Schilddrüsenkrebsfälle mit Wohnsitz zum Zeitpunkt des Unfalls in den drei Oblasts, und 202 Fälle mit Wohnsitz in der Stadt Kyiv gemeldet.

Es liegen Ergebnisse morphologischer Analysen für 147 Fälle (77%) in den drei Oblasts, und für 96 Fälle (48%) in der Stadt Kyiv vor. Insgesamt waren 52% dieser Fälle Kinder (Alter bei Operation unter 15) und 29% der Fälle junge Erwachsene (Alter bei Operation 19 bis 31). Diese Altersverteilung ist nicht repräsentativ für die Altersverteilung aller Fälle, für die deutlich mehr Fälle unter jungen Erwachsenen auftraten.

Der Anteil papillärer Karzinome ist in den drei Oblasts mit 98% außerordentlich hoch (Tabelle 18). Dies entspricht der Tatsache, dass früh nach Strahlenexpositionen zusätzlich auftretende Tumoren im Wesentlichen papilläre Karzinome sind. In der Stadt Kyiv war die Strahlenexposition geringer als in den drei Oblasts. Dementsprechend ist der Anteil papillärer Karzinome geringer. Die Anzahl follikulärer Karzinome ist gering, zeigt jedoch eine zunehmende Tendenz mit zunehmendem Alter bei Operation. Dies kann mit einem höheren relativen Beitrag der spontanen Tumoren bei höherem Alter erklärt werden.

Tabelle 18: Haupttypen von im Zeitraum 1990-99 operierten Schilddrüsenkarzinomen in der Geburtskohorte 1968-85

Histologischer Typ	Chernihiv, Kyiv und Zhytomyr Oblasts	Stadt Kyiv
Papillär	144 (98%)	86 (90%)
Follikulär	2 (1,4%)	7 (7%)
Medullär	1 ^a (0,7%)	1 (1%)
Anaplastisch	-	2 (2%)
Insgesamt	147 (100%)	96 (100%)

^ain diesem Fall wurde gleichzeitig ein papilläres Karzinom gefunden

Die relativen Häufigkeiten der Haupttypen von Schilddrüsenkarzinomen im gesamten Untersuchungsgebiet (drei Oblasts und die Stadt Kyiv) entspricht den relativen Häufigkeiten in der gesamten Ukraine, und auch in Weißrussland (Tronko et al. 1999, 2002a, 2002b; Demidchik & Demidchik 2002).

Ergebnisse spezifischer histologische Untersuchungen lagen für 130 Fälle in den drei Oblasts und für 78 Fälle in Kyiv vor. Kleine Tumoren (≤ 1 cm) wurden in 8% der analysierten Fälle gefunden (Tabelle 19). Ihre Häufigkeit nahm mit der Zeit nach dem Tschernobylunfall zu, von 3% im Zeitraum 1990-93 auf 12% im Zeitraum 1997-99. Die relative Häufigkeit kleiner Tumoren im gesamten Studiengebiet entspricht der relativen Häufigkeit in der gesamten Ukraine, ist allerdings deutlich geringer als in Weißrussland, wo eine relative Häufigkeit kleiner Tumoren von 38% angegeben wurde. Die Ursache dieses Unterschiedes zwischen der Größe operierter Tumoren in den beiden Ländern ist noch nicht abschließend geklärt.

Tabelle 19: Größenverteilung von im Zeitraum 1990-99 operierten Schilddrüsenkarzinomen in der Geburtskohorte 1968-85

Tumorgröße	Chernihiv, Kyiv und Zhytomyr Oblasts	Stadt Kyiv
≤ 1 cm	10 (8%)	7 (9%)
1,1 – 4 cm	108 (83%)	58 (74%)
> 4 cm	12 (9%)	13 (17%)
Insgesamt	130 (100%)	78 (100%)

Eine Ausbreitung der Tumoren über die Schilddrüse hinaus (T4-Kategorie der TNM Klassifikation) wurde in 61% der Fälle unter Kindern und in 27% der Fälle unter jungen Erwachsenen beobachtet. Regionale Metastasen traten in 58% der Fälle unter Kindern, und in 41% der Fälle unter jungen Erwachsenen auf. Ähnliche Verteilungen wurden für die gesamte Ukraine gefunden.

Zusammenfassend ergibt sich in den drei Oblasts und in der Stadt Kyiv ein ähnliches Bild der morphologischen Typisierung der Schilddrüsentumoren wie in der gesamten Ukraine. Auffallend ist allerdings der extrem hohe Anteil von papillären Karzinomen in den drei Oblasts. Im Vergleich zu Weißrussland fällt der relative geringe Anteil von kleinen Tumoren (≤ 1 cm) auf, der ein Hinweis auf eine relativ geringe Nachweiserate von Schilddrüsentumoren sein kann.

5 Simulationen zu ökologischen Studien zum Schilddrüsenkrebsrisiko nach dem Tschernobylunfall

Mit Simulationsstudien lassen sich Einsichten in die Wirkungszusammenhänge gewinnen, unter denen eine Risikoschätzung zustande gekommen ist, die auf realen epidemiologischen Daten basiert. Diese Tatsache ist eine wichtige Motivation für die Simulationen in diesem Projekt, mit denen das Entstehen eines ökologischen Bias verstanden und seine Größe quantifiziert werden soll. Hierzu werden nicht bekannte Daten, von denen jedoch die Ergebnisse einer Risikostudie abhängen, auf der Grundlage von plausiblen Annahmen simuliert. Speziell für Risikostudien nach dem Tschernobylunfall wurden drei wichtige Ursachen identifiziert: 1) der Screening-Effekt, 2) Messfehler für die Strahlendosis sowie 3) Abweichungen von einer linearen Dosiswirkungsbeziehung identifiziert. Im Folgenden werden die Ergebnisse der Simulationsrechnungen zusammenfassend dargestellt. Die detaillierten Darstellungen sind jeweils in den Anhängen 4-6 bzw. 8 und 9 zu finden. Die Auswirkungen der Überdiagnose (Abschnitt 5.1.1) von Schilddrüsenkrebskarzinomen auf die Risikoschätzer wurden mit synthetisch generierten Daten untersucht, die nicht auf realen Inzidenzraten und Dosis-schätzungen beruhen. Aussagen für die Risikoschätzer können deshalb nur von qualitativer Natur sein. Die übrigen Simulationsrechnungen zum Screening-Effekt (Abschnitt 5.1.3), zu Messfehlern (Abschnitte 5.2.1, 5.2.2) und zu einem linear-quadratischem Risikomodell (Abschnitt 5.3) basieren auf realen Daten und können auch ansatzweise quantitative Aussagen zur Verzerrung von Risikoschätzer liefern.

5.1 Screening

Eine verbesserte Nachweis- und Berichtsrate (Screening-Effekt) ist neben der Strahlenexposition eine Hauptursache für den Anstieg der Inzidenz von Schilddrüsenkrebs in der Geburtskohorte 1968-85. Die Vorhersagevariable Exposition kann nun durch den Screening-Effekt als Confounder beeinflusst werden. Die Erhebung von verlässlichen Daten zur Stärke des Screening-Effektes in der gesamten exponierten Bevölkerung hat sich als sehr schwierig herausgestellt. Noch vergleichsweise einfach zu ermitteln sind Daten zur Verbreitung von Ultraschall Diagnosegeräten. Aber auch die gewachsene Empfindsamkeit in der Bevölkerung für Fragen der gesundheitlichen Folgen des Tschernobyl-Unfalles führt zu vermehrten Arztbesuchen und damit zu einer gesteigerten Nachweisrate. Hier scheint eine direkte Quantifizierung der Auswirkung auf die Nachweisrate nicht möglich. Deswegen wurde die gesteigerte Nachweis- und Berichtsrate entweder mit plausiblen Annahmen simuliert oder indirekt abgeschätzt durch den Vergleich von geeigneten Inzidenzraten, die für die Situation vor und nach dem Unfall repräsentativ sind.

5.1.1 Überdiagnose

In diesem Abschnitt werden die Simulationsrechnungen aus Anhang 4 zusammenfassend dargestellt.

Einleitung: Symptomatischer papillärer Schilddrüsenkrebs ist im jungen Erwachsenenalter eine seltene Krebserkrankung, in Europa liegen die alterstandardisierten Inzidenzraten pro 100.000 Personen für die Altersgruppe 15 bis 44 Jahre ungefähr bei 2 für Frauen und ungefähr bei 0,7 für Männer im Jahr. Die Prävalenzen von Schilddrüsenkarzinomen in Autopsiestudien liegen dagegen zwischen 2% und 30%. Im Gegensatz zu den klinisch relevant gewordenen Schilddrüsenkarzinomen handelt es sich bei den in Autopsiestudien gefundenen Karzinomen fast ausschließlich um so genannte Mikrokarzinome mit einem

Durchmesser $<1\text{cm}$. Es ist deswegen zu erwarten, dass in einer Bevölkerung, in der durch eine Exposition eine erhöhte Aufmerksamkeit gegenüber Schilddrüsenkrebs herrscht und in der mit zunehmend verbesserten diagnostischen Methoden nach dieser Tumorform gesucht wird, ein Teil dieser Mikrokarzinome diagnostiziert wird. Die Inzidenzrate würde dadurch gegenüber einer nicht exponierten Bevölkerung erhöht, unabhängig von dem kausalen Einfluss der Exposition auf die Krebsraten. Eine solche Situation liegt nach dem Tschernobylunfall in den betroffenen Gebieten der Ukraine vor.

Methoden: In dieser Simulationsstudie wurde die Auswirkung von zusätzliche diagnostizierten, strahlenunabhängigen Mikrokarzinome auf den Risikoschätzer der Wirkung von radioaktiver Strahlung simuliert. Es wurde eine stark vereinfachte Standardpopulation simuliert, in der für jedes Individuum neben einer teilweise expositionsabhängigen Wahrscheinlichkeit, an einem klinisch auffälligen Tumor zu erkranken, eine expositionsunabhängige Wahrscheinlichkeit für ein Mikrokarzinom vorlag. Die Häufigkeit dieser Mikrokarzinome wurde, wie in den Autopsiestudien beschrieben, zwischen 2% und 30% variiert. Die vorhandenen Mikrokarzinome wurden mit einer Wahrscheinlichkeit von 1% bzw. 20% aufgefunden. Das Auffinden der Mikrokarzinome wurde im ersten Schritt expositionsunabhängig simuliert, während die Wahrscheinlichkeit im zweiten Schritt von der Höhe der Exposition abhängig war.

Ergebnisse: Zusätzliche, expositionsunabhängige Mikrokarzinome führten dazu, dass die β -Schätzer der individuellen und der ökologischen Regression einen Bias in Richtung 0 aufweisen. Die ökologische Analyse ist in diesem Fall gegenüber der individuellen kaum benachteiligt. Besteht ein Zusammenhang zwischen Exposition und Diagnosewahrscheinlichkeit der Mikrokarzinome, entstand bei der individuellen Diagnose ein Bias in Richtung 0, bei der ökologischen Analyse stiegen die β -Schätzer dagegen gegenüber dem simulierten „wahren“ Wert an.

Fazit: In der Ukraine bestand nach dem Tschernobyl-Unfall in den betroffenen Gebieten eine erhöhte Aufmerksamkeit gegenüber Schilddrüsenkrebs. Es wurden, z.T. mit ausländischer Hilfe, Ultraschallgeräte angeschafft, die das Auffinden von kleinen Tumoren erleichterten. Eine Abhängigkeit der Aufmerksamkeit und der räumlichen Verteilung der verbesserten Diagnostik von der Exposition ist deswegen denkbar. Die Simulationen geben einen ersten Eindruck, wie sich diese veränderten Bedingungen auf den Risikoschätzer für strahlenabhängige Tumore auswirken können. Dabei zeigen die Ergebnisse der letzten Simulation, dass unterschiedliche Auswirkungen auf den Schätzer auf Gruppenebene und auf individueller Ebene möglich sind. Diese Frage sollte, unter Einbezug der realistischen Größenverteilungen der Schilddrüsentumoren im Untersuchungsgebiet, weiter untersucht werden.

5.1.2 Methodenentwicklung mit realitätsnahen Daten

In diesem Abschnitt werden die Ergebnisse der methodischen Untersuchungen zusammengefasst, die im Manuskript des Anhangs 5 ausführlich dargestellt sind. Die Rechnungen wurden mit der vorläufigen Dosissschätzung aus Abschnitt 4.3.2 durchgeführt. Eine mathematische Formulierung der Wirkung eines Confounders auf einen Risikoschätzer in ökologischen Studien wurde für den Fall des Einflusses von Rauchen auf das Lungenkrebsrisiko von Radon exponierten Personen von Lubin (1998, 2002) vorgestellt. Für das Schilddrüsenrisiko nach Tschernobyl wurde dieser Ansatz übernommen und weiter entwickelt. Allerdings muss ein konzeptioneller Unterschied beachtet werden. Bei

Lungenkrebs erhöhen sowohl Rauchen als auch die Exposition mit Radon das Risiko. Der Confounder Screening ist jedoch nicht biologisch wirksam sondern steigert lediglich die Nachweis- und Berichtsrate für Schilddrüsenkrebs.

Die Gesamtzahl n_c der im Zeitraum 1990-99 im Studiengebiet registrierten Krebsfälle setzt sich aus vier Beiträgen zusammen gemäß

$$n_c = n_{0n} + n_{0s} + n_m + n_{rs}. \quad (5.1)$$

Die n_{0n} spontanen Fälle würden registriert werden ohne den Kraftwerksunfall bei unveränderter Nachweis- und Berichtsrate nach dem Jahre 1990. Die n_m Strahleninduzierten Fälle wären klinisch relevant geworden nach dem Unfall bei immer noch unveränderter Nachweis- und Berichtsrate. Die $n_{0s}+n_{rs}$ zusätzlichen Fälle sind auf den Screening-Effekt zurückzuführen.

Motiviert durch die Einteilung der registrierten Krebsfälle nehmen wir an für das individuelle Risiko einer Person j aus einer Ortschaft i

$$h_{ij} = (1 + \eta_{ij})h_0 + (1 + \kappa_{ij})\beta D_{ij}. \quad (5.2)$$

Dabei ist

η_{ij} der zusätzliche Anstiegsfaktor für das Hintergrundrisiko, hervorgerufen durch die gesteigerte Nachweis- und Berichtsrate,

h_0 das Hintergrundrisiko ohne erhöhte medizinische Überwachung,

κ_{ij} der zusätzliche Anstiegsfaktor für das Strahlenrisiko, hervorgerufen durch die gesteigerte Nachweis- und Berichtsrate,

β das absolute Zusatzrisiko pro Einheitsdosis für den Fall ohne gesteigerte medizinische Überwachung und

D_{ij} die individuelle Schilddrüsendosis.

Es ist prinzipiell unmöglich, die individuellen Daten zu Exposition und Screening für *alle* Mitglieder der Geburtskohorte 1968-85 aus dem Studiengebiet zu sammeln. Für den idealen Fall, dass diese Information verfügbar wäre, lautet das mittlere bevölkerungsbezogene Risiko in der Risikogruppe

$$\langle h \rangle = \langle h_0 \rangle_{pop} + \beta_{pop} \langle D \rangle \quad (5.3)$$

mit dem absoluten Zusatzrisiko pro Einheitsdosis (EARPD) in der Bevölkerung

$$\beta_{pop} = (1 + \langle \kappa \rangle) \beta \left(1 + \frac{\langle \text{cov}_I(\kappa, D) \rangle + \text{cov}_S(\bar{\eta}, \bar{D})}{(1 + \langle \kappa \rangle) \langle D \rangle} \right), \quad (5.4)$$

wobei die Größen in eckigen Klammern immer die Mittelwerte über alle Mitglieder der Risikogruppe bezeichnen. Für das Hintergrundrisiko gilt ein entsprechender Ausdruck. In der Formel für das EARPD β_{pop} stehen zwei Kovarianzterme für die Korrelation zwischen Exposition und Screening innerhalb von Ortschaften oder zwischen den Ortschaften. Sie entstehen zum Beispiel, wenn in stark kontaminierten Ortschaften mehr Personen zu medizinischen Untersuchungen gehen.

Die ökologische Analyse basiert auf dem mittleren Risiko

$$\bar{h}_i = \langle h_0 \rangle_{eco} + \beta_{eco} \bar{D}_i \quad (5.5)$$

für eine Siedlung i . Jegliche Information zur medizinischen Überwachung wird bei diesem Ansatz ignoriert. Der arithmetische Mittelwert \bar{D}_i wird benutzt, um das ökologische Hintergrundrisiko $\langle h_0 \rangle_{eco}$ und das ökologische EARPD β_{eco} zu schätzen. Mit der Poissonregression wird das mittlere Risiko $\langle h \rangle$ korrekt reproduziert. Jedoch wird die Aufteilung in ein Hintergrundrisiko und ein Strahleninduziertes Risiko in den meisten Fällen von der wahren Aufteilung abweichen, die durch die Terme $\langle h_0 \rangle_{pop}$ und $\beta_{eco} \langle D \rangle$ gegeben ist. Der ökologische Bias wird nun definiert als das Verhältnis $\beta_{eco} / \beta_{pop}$, wobei der Wert eins für einen unverzerrten Risikoschätzer steht.

Mit drei stark vereinfachenden Modellannahmen für die Korrelation zwischen Exposition und Screening wurde untersucht, wie sich diese Korrelationen auf die Stärke des Bias auswirken. Die Untersuchungen wurden zunächst mit Simulationsrechnungen durchgeführt. Parallel dazu wurde ein analytisches Lösungsverfahren entwickelt, mit dem die Risikoschätzer aus einem Gleichungssystem bestimmt werden können.

Ein wesentliches Ergebnis dieser Rechnungen war, dass der Bias durch eine komplizierte Korrelationsstruktur zwischen den konstituierenden Größen des individuellen Risikos h_{ij} bestimmt wird. Allein das Vorliegen von Korrelationen ist allerdings nicht hinreichend für das Auftreten eines Bias. Treten nur solche Korrelationen auf, die schon in der Definition des bevölkerungsbezogenen Risikos β_{pop} enthalten sind, würde sich kein Bias ergeben. Als Beispiel kann ein spezielles Korrelationsmodell gelten, das analog zu dem von Lubin (1998, 2002) benutzten Modell aufgebaut wurde. Hier konnten die ökologischen Schätzer das bevölkerungsbezogene Risiko unverzerrt wiedergeben trotz der Anwesenheit von Korrelationen innerhalb von Ortschaften.

5.1.3 Szenarienrechnungen mit realitätsnahen Daten

Dieser Abschnitt fasst die Ergebnisse des Manuskriptes aus Anhang 6 zusammen, die Ergebnisse wurden auf der Grundlage der endgültigen Dosissschätzung (Abschnitt 4.3.3) erzielt. Eine Aufgabe des Projektes war die Entwicklung von Szenarien zum Screening-Effekt, die eng an die im Studiengebiet vorliegende Situation angelehnt sind. Deshalb wurden die Szenarien so aufgebaut, dass sie gesicherte Fakten möglichst genau wieder geben. Dazu gehört die Gesamtzahl der im Studiengebiet im Beobachtungszeitraum registrierten 350 Fälle. Aufwendiger ist die Abschätzung der zusätzlichen Anstiegsfaktoren für die erhöhte Nachweis- und Berichtsrate. Die dazu gemachten Annahmen sind im Abschnitt 4.3.4

dargestellt, in Tabelle 17 befinden sich Zahlenwerte für die oblastbezogenen zusätzlichen Anstiegsfaktoren und die „ungestörten“ Hintergrundrisiken. Wenn das gesamte bevölkerungsbezogene Risiko und das Hintergrundrisiko bekannt sind, ergibt sich das strahlenabhängige Risiko durch Subtraktion. Betrachtet wurden zwei mögliche Varianten für die Größe des Hintergrundrisikos. In der ersten Variante beträgt das mittlere Hintergrundrisiko 21,2 Fälle pro 10^6 Personenjahre, in der zweiten Variante ist es um einen Faktor 0,73 kleiner. Die Werte für das bevölkerungsbezogene EARPD β_{pop} lauten entsprechend 1,7 Fälle pro 10^4 Personenjahre und Gy bzw. 2,4 Fälle pro 10^4 Personenjahre und Gy.

Zwei Szenarien zu den Auswirkungen einer erhöhten Nachweis- und Berichtsrate wurden entwickelt, die beide die getroffene Aufteilung von Hintergrundrisiko und strahlenabhängigem Risiko beachten. Für beide Szenarien wurde angenommen, dass das Hintergrundrisiko innerhalb einer Oblast konstant ist und zwischen den Oblasts variiert. Die daraus resultierende Korrelation zwischen den zusätzlichen Anstiegsfaktoren und dem Hintergrundrisiko ist jedoch zu vernachlässigen. Bemerkbar macht sich nur die Korrelation zwischen Exposition und Screening, weil in der Zhytomyr Oblast die höchste mittlere Dosis auf den kleinsten Anstiegsfaktor trifft. Im ersten Szenario treten also nur Korrelationen *zwischen* den Oblasts auf. Im zweiten Szenario wird noch zusätzlich vorausgesetzt, dass auch Korrelationen zwischen Exposition und Screening *innerhalb* einer Ortschaft vorhanden sind. Hier wird zusätzlich für größere Städte mit mehr als 1000 Kindern aus der Risikogruppe angenommen, dass ca. 1% kurz nach dem Unfall in hoch kontaminierte ländliche Gebiete ausgewichen ist. Für diese zahlenmäßig kleine Gruppe wird eine mittlere Dosis angesetzt, die fünfmal höher als der Dosismittelwert der ganzen Ortschaft ist. Die Hälfte der Mitglieder dieser hoch exponierten Gruppe lässt sich in diesem Szenario ärztlich auf ein Schilddrüsenkarzinom untersuchen. Dieser Anteil ist deutlich höher als in der Gesamtbevölkerung, was zu einer positiven Korrelation innerhalb größerer Städte führt.

In den Szenarien wurden jeweils zwei verschiedene Werte für das Verhältnis des individuellen zusätzlichen Anstiegsfaktors für das strahlenabhängige Risiko und für das spontane Risiko κ_{ind}/η_{ind} berücksichtigt. Der Wert $\kappa_{ind}/\eta_{ind} = 1$ bedeutet, dass der Anstieg für spontane Tumoren und für strahlenabhängige Tumoren gleich stark ist. Für $\kappa_{ind}/\eta_{ind} = 0,2$ wird angenommen, dass durch das Screening fünfmal mehr spontane Tumoren gefunden werden. Motiviert wird dieser Wert durch die hohe Anzahl von okkulten Tumoren, die in verschiedenen Autopsiestudien (Abschnitt 4.2) obduziert wurden.

Die Poissonregression der Rohdaten zu Exposition (Likhtarov et al. 2005) und Inzidenz auf der Ebene der Ortschaften ergab für das EARPD β_{raw} den Wert 1,7 (1,2; 2,3) Fälle pro 10^4 Personenjahre und Gy und für das ERRPD γ_{raw} den Wert 8,4 (5,9; 12) pro Gy. Die Werte in Klammern bedeuten hier die obere und untere Grenze des 95% Konfidenzintervalles (KI). Das entspricht genau den Punktschätzern, die in den Simulationen des ersten Szenarios für die erste Variante mit dem Verhältnis $\kappa_{ind}/\eta_{ind} = 0,2$ das ökologische EARPD β_{eco} und das ERRPD γ_{eco} gefunden wurden. Die Punktschätzer für die Risikofaktoren der zweiten Variante liegen oberhalb der KI der Punktschätzer aus den Rohdaten. Deshalb werden die Ergebnisse der zweiten Variante im Folgenden nicht mehr diskutiert.

Der ökologische Bias war im zweiten Szenario größer als im ersten Szenario. Er betrug maximal -9% für das EARPD und -15% für das ERRPD. Die bevölkerungsbezogenen Risikofaktoren werden durch die ökologischen Risikofaktoren unterschätzt, wobei der Bias

besonders für das EARPD als moderat anzusehen ist. Auch ein weiterer statistischer Test, bei dem gezählt wird, wie oft bei einer gegebenen Anzahl von Simulationsläufen der „wahre“ Wert eines Risikofaktors im KI des entsprechenden ökologischen Risikofaktors liegt, deutet auf nur geringe Verzerrungen hin.

Die Punktschätzer für das EARPD zeigen nur eine schwache Abhängigkeit vom Verhältnis κ_{ind}/η_{ind} , die Unterschiede zwischen den Werten 1 und 0,2 blieben unter 10%. Für das ERRPD wurden größere Unterschiede bis zu 20% beobachtet. Allerdings waren sie in keinem Fall statistisch signifikant.

Zum Vergleich wurde auch ein Szenario für eine Kohortenstudie entwickelt, das sich an eine laufende Kohortenstudie (Tronko et al. 2003) anlehnt. In der Kohorte wurde in den Jahren 2000-02 eine um den Faktor 20 höhere Inzidenz als in der ökologischen Studie der Jahre 1990-99 beobachtet. Hierfür gibt es drei Ursachen: ein höhere mittlere Exposition in der Kohorte von ca. 0,65 Gy, ein hohe Nachweis- und Berichtsrate für *alle* Mitglieder der Kohorte und ein höheres mittleres Alter. Im Zeitraum 1990-99 hat die Geburtskohorte 1968-85 ein mittleres Alter von 18 Jahren, im Zeitraum 2000-02 für die Kohortenstudie betrug es 24 Jahre. Wiederum ist es schwierig, die Wirkung aller drei Ursachen zu quantifizieren. Für das EARPD finden wir in der Kohorte eine Erhöhung um einen Faktor 4-5, die zum größten Teil auf intensives Screening zurück zu führen ist. Aber sicher trägt auch das höhere Alter zu dem Unterschied bei. Für das ERRPD finden wir keinen Unterschied, wenn das Verhältnis $\kappa_{ind}/\eta_{ind} = 1$ ist, weil hier der Anstieg gleich auf das spontane und das strahlenabhängige Risiko auswirkt. Für das Verhältnis $\kappa_{ind}/\eta_{ind} = 0,2$ sinkt das ERRPD deutlich gegenüber dem ökologischen ERRPD. Ein ähnliches Verhalten konnten auch Schneider et al. (1993) feststellen.

5.2 Messfehler

In vielen epidemiologischen Studien kann die interessierende Exposition für jedes einzelne Individuum nur abgeschätzt und nicht genau gemessen werden. Dadurch ergibt sich ein Messfehler, der den Zusammenhang, der in Regressionsmodellen untersucht werden soll, stören kann. Die Untersuchung der Wirkung von Messfehlern oder Fehlklassifikation (beim Vorliegen von qualitativen Expositionsvariablen) auf die Risikoschätzung hat eine lange Tradition in der Epidemiologie. Dabei werden zwei unterschiedliche Modelle zur Berücksichtigung von Messfehlern betrachtet: das klassische Messfehlermodell und das Modell nach Berkson. Im Folgenden wird eine kurze Darstellung dieser Modelle im Rahmen der einfachen linearen Regression gegeben. Eine ausführliche Darstellung findet man bei Fuller (1987).

Sei die interessierende wahre Regressionsgleichung gegeben durch

$$Y = \alpha + \beta Z + e \quad (5.6)$$

Ein Messfehlerproblem entsteht dadurch, dass die wahren Werte der erklärenden Variable Z nicht beobachtbar oder messbar sind. Anstelle von Z wird daher eine Surrogatvariable X verwendet. Beim klassischen additiven Messfehlermodell geht man davon aus, dass X sich aus dem wahren Wertes Z und dem Messfehler ε zusammensetzt, d.h.

$$X = Z + \varepsilon, \quad Z \text{ und } \varepsilon \text{ unabhängig, } \varepsilon \sim N(0, \sigma^2) \quad (5.7)$$

Es wird also angenommen, dass der Fehler normalverteilt ist und unabhängig von dem beobachteten Wert Z . Die Variation von X ist damit größer als die von Z . Aufgrund von (5.7) wird aus der Regressionsgleichung (3.6)

$$Y = \alpha + \beta Z + e = \alpha + \beta(X - \varepsilon) + e = \alpha + \beta X + u \quad \text{mit } u = e - \beta\varepsilon \quad (5.8)$$

d.h. bei Verwendung von X erhält man eine Regressionsgleichung, in der die erklärende Variable und die Residuen u korreliert sind. Es gilt hierbei

$$\text{Cov}(X, u) = -\beta\sigma^2 \quad (5.9)$$

Durch diese Korrelation erhält man in der einfachen linearen Regression einen nach unten verzerrten Schätzer für den Regressionskoeffizienten β (Attenuation Bias). Zur Korrektur dieser Verzerrung gibt es im linearen Modell verschiedene Methoden (Fuller 1987). Für komplexere Modelle (logistische Regression, Cox-Regression) sind die Eigenschaften der Schätzer bisher weniger gut untersucht (Bender und Blettner 2000).

Beim Berkson-Modell geht man davon aus, dass man den Effekt von gewissen Dosiswerten X untersuchen möchte, die man aber in der Praxis nicht exakt einhalten kann. Die wahre individuelle Dosis Z setzt sich also zusammen aus der vorgegebenen Dosis X und einem individuellen von X unabhängigen normalverteilten Messfehler ε

$$Z = X + \varepsilon, \quad X \text{ und } \varepsilon \text{ unabhängig, } \varepsilon \sim N(0, \sigma^2) \quad (5.10)$$

Die Variation von Z ist damit größer als die von X . Aufgrund von (5.10) wird aus der Regressionsgleichung (5.6)

$$Y = \alpha + \beta Z + e = \alpha + \beta(X + \varepsilon) + e = \alpha + \beta X + u \quad \text{mit } u = e + \beta\varepsilon \quad (5.11)$$

also eine Regressionsgleichung zwischen Y und X mit Residuen, die mit X nicht korreliert sind. Aus diesem Grund ist β im einfachen linearen Regressionsmodell auch bei Vorliegen von Berkson-Fehlern unverzerrt schätzbar. Diese Eigenschaft gilt aber nicht unbedingt auch in nicht-linearen Modellen (Carroll et al. 1995). Sowohl das klassische als auch das Berkson-Modell lassen sich auch mit multiplikativen Fehlern formulieren. Die Verwendung eines additiven Berkson-Modells hat den Nachteil, dass dieses Modell auch negative Expositionen zulässt. Da dies theoretisch nicht möglich ist, empfiehlt sich vermutlich die Anwendung eines multiplikativen Berkson-Modells. Messfehler in ökologischen Studien wurden von Brenner et al. (1992) untersucht.

5.2.1 Klassischer und Berkson-Fehler

Im Anhang 8 wurden Simulationen zur Auswirkung von zufälligen individuellen Messfehlern auf die Risikoschätzer durchgeführt. Im klassischen Fehlermodell schwankt der notierte Dosiswert zufällig um den wahren Wert. Die Schwankung wurde durch die Multiplikation einer angenommenen Wahrscheinlichkeitsdichte mit dem wahren Messwert realisiert. Dadurch wird die Verteilung der notierten Dosen für alle Personen breiter als die Verteilung der wahren Dosen. Ein solches Verhalten ist zu erwarten, wenn der wahre Dosiswert mit einem Messgerät ungenau gemessen wird. Im Gegensatz dazu schwankt im Berkson-Fehlermodell der wahre Wert der Dosis um den gemessenen Dosiswert. Das kann passieren,

wenn die Exposition theoretisch mit einem unsicheren Modell berechnet wird oder einer bestimmten Personengruppe ein fester Messwert zugeordnet wird (Carroll et al. 1995).

Grundlage der Simulation der Dosisverteilung waren die Daten der vorläufigen Dosissschätzung nach Abschnitt 4.3.2. Die Wahrscheinlichkeitsdichte der Verteilung des individuellen Messfehlers wurde als lognormal angenommen mit einem geometrischen Mittelwert $\mu_g = 1$. Die Breite der Wahrscheinlichkeitsdichte wurde variiert mit drei verschiedenen geometrischen Standardabweichungen $\sigma_g = 1,5, 2,25$ und 3 . Die Krebsfälle wurden mit einem linearen Risikomodell

$$h_{ij} = h_0 + \beta D_{true,ij} \quad (5.12)$$

ermittelt unter Verwendung der wahren Dosis $D_{true,ij}$ für eine Person i aus einer Ortschaft j . Für das wahre Hintergrundrisiko h_0 wurde ein realistischer Werte von 10 Fällen pro 10^6 Personenjahren angenommen, für das Zusatzrisiko pro Einheitsdosis (EARPD) β ein Wert von 2 pro 10^4 Personenjahren und Gy. Die Risikoschätzer $h_{0,eco}$ und β_{eco} für die wahren Werte wurden mit der Poissonregression aus zusammengefassten Daten für Ortschaften bestimmt. Zum Vergleich wurden die Schätzer $h_{0,ind}$ und β_{ind} aus der Regression der individuellen Likelihood auf der Grundlage der Individualdaten ermittelt.

Mit der Poissonregression ergibt sich für den Bias des EARPD

$$\frac{\beta_{eco}}{\beta} = \frac{\mu_g}{\mu_a} \text{ im klassischen Fehlermodell und } \frac{\beta_{eco}}{\beta} = \frac{\mu_a}{\mu_g} \text{ im Berkson - Fehlermodell.} \quad (5.13)$$

Hierbei sind μ_g und μ_a der geometrische Mittelwert und der arithmetische Mittelwert der angenommenen Wahrscheinlichkeitsdichte für die Verteilung des individuellen Fehlers. Bei konstantem μ_g wächst μ_a monoton mit steigender geometrischer Standardabweichung σ_g . Diese einfachen Beziehungen wurden empirisch mit Simulationen gefunden und danach auch analytisch exakt bewiesen. Außerdem stellte sich heraus, dass das Hintergrundrisiko h_0 unverzerrt geschätzt werden kann.

Die Schätzer aus der Regression der individuellen Likelihood verhalten sich im Berkson-Fehlermodell genau wie die Schätzer aus der Poissonregression. Im klassischen Fehlermodell werden sowohl das Hintergrundrisiko als auch das EARPD verzerrt wieder gegeben. Die Verzerrung des EARPD ist deutlich stärker, wenn man sie mit dem Schätzer aus der Poissonregression vergleicht. Diese Ergebnisse wurden nur mit Simulationen erreicht, ein analytischer Beweis steht noch aus.

5.2.2 Ökologischer Messfehler

Hier wurden die Simulationen (s. Anhang 8.2) auf Grundlage der endgültigen Dosissschätzung aus Abschnitt 4.3.3 (Likhtarov et al. 2005) durchgeführt. Allerdings wurden von den ursprünglich 670 Ortschaften nur 581 Ortschaften verwendet, weil nur hierfür eine detaillierte Liste mit der genauen Anzahl der gemessenen Personen vorlag. Die Anzahl der Personen in

der Risikogruppe sank dadurch nur unwesentlich auf knapp unter eine Million. Aus der Liste geht hervor, dass in der größten Stadt Kyiv nur 0,66% der Personen gemessen wurden. In der Stadt Chernihiv betrug der Anteil 7% und in der Stadt Zhytomyr 0,32%. In den vielen kleineren Ortschaften lag der Anteil der gemessenen Personen immer über 10%. Insgesamt wurde ein Anteil von 7% an der gesamten Risikogruppe gemessen. Die Krebsfälle wurden wieder mit einem linearen Risikomodell simuliert, wobei angenommen wurde, dass die individuelle Dosis für alle Personen bekannt ist. Ein individueller Messfehler wurde nicht noch zusätzlich betrachtet.

Weil nicht alle Personen in einer Ortschaft gemessen wurden, kann der Mittelwert der Ortsdosis nur ungenau bestimmt werden. Diese Unsicherheit wird eine Verzerrung der Risikoschätzer aus der Poissonregression hervorrufen. Weil die Mittelwerte aus den unvollständigen Messungen um die wahren Mittelwerte der Ortsdosis schwanken, liegt die Vermutung nahe, dass sich der ökologische Messfehler wie ein klassischer Fehler auswirkt und einen Bias kleiner als eins erzeugt (Brenner et al. 1992). Die Vermutung wurde mit Simulationen getestet.

Tatsächlich sinkt der Bias β_{eco}/β des EARPD auf Werte kleiner als eins, wenn der relative Anteil der gemessenen Personen in jeder Ortschaft gleich verringert wird. Wenn in jeder Ortschaft (und in der Stadt Kyiv) nur 7% der Personen gemessen werden, liegt der Wert des Bias jedoch bei 0,96 und ist damit vernachlässigbar. In der Liste der gemessenen Personen schwankt der Anteil der gemessenen Personen von Ort zu Ort und liegt im Mittel bei 7%. Diese zusätzliche Variabilität hat auf den Bias des EARPD einen geringen Einfluss, er liegt jetzt bei 0,97. Jedoch treibt diese Variabilität den Bias des Hintergrundrisikos von 1,05 auf 1,23. Damit wächst zugleich der Bias des relativen Zusatzrisikos. Wenn man den Extremfall annimmt, dass in jeder Ortschaft nur eine Person gemessen wurde, sinkt das Verhältnis β_{eco}/β auf den Wert 0,61. Die Vermutung, dass sich der ökologische Messfehler wie ein klassischer Fehler auswirkt, konnte mit den Simulationen bestätigt werden. Verglichen mit den individuellen Messfehlern im klassischen Fehlermodell und im Berkson-Fehlermodell ist die Wirkung des ökologischen Messfehlers jedoch vergleichsweise gering.

5.3 Linear-quadratisches Risikomodell in ökologischen Studien

In der Strahlenepidemiologie benutzt man oftmals ein linear-quadratisches Risikomodell, um die Abweichung einer Dosis-Wirkungsbeziehung von einem linearen Zusammenhang zu testen. Hier werden die Ergebnisse des Anhangs 9 zusammengefasst, in dem Simulationen mit der realitätsnahen endgültigen Dosis-schätzung (Abschnitt 4.3.3) durchgeführt wurden. Wie wir gesehen haben, kann allein die Gruppierung den *pure specification bias* verursachen selbst wenn keine Confounder wie z.B. Screening vorliegen.

Für die Simulation der Krebsfälle wurde das Risikomodell

$$h_{ij} = h_0 + bD_{ij} + cD_{ij}^2 \quad (5.14)$$

für eine Person j aus einer Ortschaft i und einer individuellen Schilddrüsendosis D_{ij} benutzt. Das Gesamtrisiko setzt sich zusammen aus dem Hintergrundrisiko h_0 und einem Strahlen bedingtem Risiko, dessen linearer und quadratischer Beitrag mit den Koeffizienten b und c gewichtet werden. Die Werte für die Koeffizienten h_0 , b und c wurden so gewählt, dass die Zahl von 350 Neuerkrankungen im Studiengebiet im Zeitraum 1990-99 im Mittel über 100

Simulationsläufe reproduziert wird. Es zeigte sich, dass für positive Werte des quadratischen Koeffizienten c die Erkrankungswahrscheinlichkeit nicht mehr hinreichend genau durch den Ausdruck ΔTh_{ij} approximiert werden kann. Die Voraussetzung, dass dieser Ausdruck klein sein muss, ist nicht mehr gegeben, weil das individuelle Risiko h_{ij} quadratisch mit der Dosis wächst. Falls c negativ ist, kann auch das individuelle Risiko negativ werden. Um das zu verhindern, wurde der Minimalwert auf das Hintergrundrisiko h_0 begrenzt. Negative Werte für c werden meist strahlenbiologisch motiviert, um den Effekt des Zellsterbens bei großen Dosen zu beschreiben.

Die Simulationen haben ergeben, dass schon bei geringen Abweichungen vom linearen Verhalten starke Abweichungen zwischen den Risikoschätzern aus der Poissonregression mit zusammengefassten Daten von den Koeffizienten des individuellen Risikomodells auftreten. Die größten Verzerrungen bis zu einem Faktor 3 wurden bei dem quadratischen Koeffizienten c festgestellt. Daraus muss gefolgert werden, dass für den Fall, dass ein linear-quadratisches Risikomodell das individuelle Schilddrüsenkrebsrisiko nach Tschernobyl beschreibt, die Koeffizienten dieses Modells mit zusammengefassten Daten nicht verlässlich bestimmt werden können.

6 Zusammenfassung

6.1 Allgemeine Ergebnisse für ökologische Studien

In ökologischen Studien erfolgt die Datensammlung und Analyse auf Gruppenebene. Die Eigenschaften „Exposition“, „Erkrankung“ und mögliche Kovariable gelten für Gruppen von Individuen und nicht für die risikobehafteten Individuen selber. Fehlen Daten auf individueller Ebene, können ersatzweise ökologische Risikostudien durchgeführt werden, deren Studiendesign oft als unvollständig bezeichnet wird.

Weil bei der Zusammenfassung von Daten die eindeutige (mathematisch: bijektive) Beziehung zwischen Exposition und Erkrankung verloren geht, können ökologische Studien in der Regel keine Aussagen zum individuellen Risiko treffen. Inwieweit sie doch einen Beitrag zur Risikoforschung leisten können, ist Gegenstand intensiver Diskussionen in der epidemiologischen Literatur. Ausgelöst wurden sie durch die Beobachtung, dass ökologische und individuelle Studien unterschiedliche, manchmal sogar widersprüchliche Ergebnisse geliefert haben. Als neueres prominentes Beispiel kann Cohens Analyse zu Lungenkrebs und Radon gelten, wobei hier das Rauchverhalten als so genannter Confounder einen ökologischen Bias für den Risikoschätzer verursacht hat.

Eine allgemeine Theorie zur Beschreibung des ökologischen Bias existiert nicht, wohl auch weil die Problemstellung noch verhältnismäßig neu ist. Deshalb wurde in diesem Projekt eine Literaturrecherche durchgeführt, um den Wissensstand zur Anwendbarkeit von ökologischen Studien zu erfassen. Es wurden die wichtigsten Faktoren ermittelt, die die Risikoschätzer aus ökologischen Studien beeinflussen und möglicherweise verzerren.

Dabei wurde klar, dass ein einheitliches Konzept für eine ökologische Studie nicht existiert. Vielmehr muss bei einer Bewertung unterschieden werden, auf welcher Ebene (engl. level) Daten erhoben (level of measurement), die Analyse durchgeführt (level of analysis) und Schlussfolgerungen gezogen (level of inference) werden. Liegen nur zusammengefasste Daten vor, muss eine vollständige ökologische Analyse durchgeführt werden. Bei einer teilweise ökologischen Analyse werden z.B. für die Variablen Alter, Geschlecht und Krankheitsstatus individuelle Werte verwendet, jedoch ist die Exposition nur auf Gruppenebene bekannt. Bei der Analyse auf individueller Ebene liegen alle relevanten Einzeldaten vor, mit denen ein individuelles Risiko abgeschätzt werden kann. Die Risikoschätzer einer ökologischen Analyse müssen auf der Gruppenebene interpretiert werden. Im Allgemeinen lässt sich aus ihnen kein individuelles Risiko ableiten. Werden die Schätzer jedoch als individuelle Risiken interpretiert, spricht man von cross-level-inference. Die Übertragung der Ergebnisse von der Gruppenebene zur individuellen Ebene kann zu einer Verzerrung des wahren individuellen Risikos, d.h. zum ökologischen Bias, führen. Sie ist Gegenstand vieler Kritiken an ökologischen Studien.

Die Gründe für das Auftreten eines ökologischen Bias sind vielfältiger Natur, die im Abschnitt 2.2 ausführlich diskutiert werden. Die wichtigsten Gründe seien hier kurz angeführt:

Bias innerhalb einer Gruppe, z.B. durch Missklassifikation oder Confounding. Hier würde jedoch auch bei der individuellen Analyse ein Bias entstehen.

Confounding auf Gruppenebene, z.B. durch die Korrelation von Hintergrundinzidenz und Exposition. Auf individueller Ebene muss dazu nicht unbedingt ein Confounding vorliegen.

Effekt-Modifikation auf Gruppenebene, z.B. wenn Kovariable mit der Exposition zwischen Gruppen korreliert sind.

Als weitere Gründe können Probleme bei der Confounder-Kontrolle, das Fehlen von adäquaten Daten oder die Migration zwischen den Gruppen genannt werden.

Wenn die Varianz der Exposition innerhalb der Gruppen die Varianz zwischen den Gruppen übersteigt, ist eine ökologische Analyse mit nicht-linearen Risikomodellen wenig effizient, weil dann der pure specification bias auftritt. Dieser Zusammenhang wurde für das logistische Modell mit synthetisch generierten Daten in einer Simulationsstudie untersucht. Dabei zeigt sich, dass der Bias mit der Baseline-Erkrankungswahrscheinlichkeit anwächst. Insgesamt wurde eine komplexe Abhängigkeit des Bias von den absoluten Werten der Varianzen deutlich. Eine Abhängigkeit nur allein vom Quotienten konnte nicht bestätigt werden.

6.2 Spezielle Ergebnisse für ökologische Studien zum Schilddrüsenkrebsrisiko nach dem Tschernobylunfall

Bei der Untersuchung des Schilddrüsenkrebsrisikos durch ^{131}I Exposition wurde in verschiedenen ökologischen Studien in der Ukraine und in Weißrussland eine klare Dosiswirkungsbeziehung gefunden.

Andererseits ist es in der epidemiologischen und biometrischen Forschung bekannt, dass ökologische Studien, also Korrelationsstudien, die aggregierte Daten heranziehen, selten von sich aus schlüssig sind. Das Hauptproblem ist der so genannte ökologische Trugschluss. Dieser resultiert daher, dass für die Individuen, die erkranken, selbst nicht bekannt ist, in welchem Maße sie dem Risiko ausgesetzt sind, d.h. die mittlere (oder aggregierte) Exposition sagt nur in Ausnahmefällen (genau: bei der Einpunktverteilung) etwas über die Exposition der Individuen aus. Außerdem ist die Exposition häufig nicht das einzige Merkmal, durch das sich die exponierten Personen von den nicht exponierten Personen unterscheiden. Letzteres bedeutet, dass ein Confounder vorliegt, der nicht berücksichtigt werden kann.

Für die spezifische Situation der Exposition im Kindesalter gegenüber ^{131}I nach dem Tschernobyl-Unfall und dem resultierendem Auftreten von Schilddrüsenkrebs liegen allerdings Bedingungen vor, die eine quantitative Abschätzung des Strahlenrisikos mit Hilfe von aggregierten Daten erlauben könnten. Zu diesen Bedingungen gehört, dass die ^{131}I Exposition mit Abstand der dominierende Faktor für die Schilddrüsenkrebsinzidenz in den höher kontaminierten Gebieten ist, und dass eine große Anzahl (mehr als 200 000) von Messungen der ^{131}I Aktivität in der menschlichen Schilddrüse durchgeführt wurde, die es erlaubt, Bevölkerungsgruppen zu definieren, deren mittlere Dosen eine größere Variabilität haben als die Individualdosen innerhalb der Bevölkerungsgruppen. Ziel des Forschungsprojektes war es deshalb zu untersuchen, wie stabil und zuverlässig die Risikoschätzungen des Schilddrüsenkrebsrisikos mit ökologischen Studien sind.

In der vorliegenden Studie wurden folgende mögliche Quellen für einen ökologischen Trugschluss betrachtet:

1. Die Verteilung der Strahlendosis innerhalb einer Region kann ungleichmäßig sein. Die Varianz der Dosisverteilung in einer Siedlung kann in der gleichen Größenordnung liegen wie die Varianz zwischen den Siedlungen.
2. Die Dosis-Wirkungsfunktion muss nicht notwendigerweise linear sein.

3. Messfehler können zu Fehler bei der Bestimmung der mittleren Exposition in einer Region führen.
4. Durch verbesserte Nachweismöglichkeiten und durch verändertes Berichtsverhalten kann die Zahl der gemeldeten Fälle ansteigen.
5. Die Nachweis- und Berichtsrate kann von der Strahlenexposition abhängen (in Regionen mit erhöhter Strahlenexposition war die Aufmerksamkeit erhöht)
6. Gezielte Screening-Programme können die Zahl der beobachteten Fälle erhöhen

Die Probleme, die aus der Bedingung in Punkt 1 entstehen, werden auch als „pure specification bias“ bezeichnet und wurden bisher in der epidemiologischen Literatur nur wenig untersucht. Diese Verzerrung spielt keine Rolle, wenn das Risiko tatsächlich linear von der Dosis abhängig ist. Geht man aber von einem nicht-linearen Modell aus (Punkt 2), so konnte gezeigt werden, dass schon geringe Abweichungen vom linearen Modell (z.B. eine quadratische Komponente) zu großen Verzerrungen der ökologischen Regressionsparameter nach sich ziehen können. Allerdings dürfte die Auswirkungen, die sich aus den Punkten 1 und 2 ergeben, für das hier untersuchte Problem weniger gravierend sein, da in dem betrachteten Dosisbereich die lineare Beziehung als adäquat angesehen wird.

Aus der Literatur war bekannt, dass sich Messfehler (Punkt 3) in ökologischen Studien nicht so auswirken wie in Studien mit individuellen Daten. Es konnte gezeigt werden, dass sich der ökologische Messfehler wie ein klassischer Messfehler verhält. Die Wirkung ist jedoch – unter den gemachten, realistischen Annahmen – eher gering.

Die Auswertung von Autopsiestudien belegte eine Prävalenz von Mikrokarzinomen, die unter normalen Umständen unbekannt bleiben, die aber durch intensivere Untersuchungsprogramme teilweise gefunden werden können. Dementsprechend wurde neben der Strahlenexposition die verbesserte Nachweis- und Berichtsrate als Hauptursache für den Anstieg der Schilddrüsentumorraten genannt (Punkt 4). Leider lagen uns keine detaillierten Daten über den Anteil der Fälle vor, der aufgrund einer verbesserten Nachweis- und Berichtsrate registriert worden ist. Deshalb mussten plausible Annahmen über die Größe dieses Effektes gemacht werden. Als Parameter dienen dafür zusätzliche Anstiegsfaktoren, die in dem Projekt StSch4240 Schilddrüsenexpositionen von weißrussischen und ukrainischen Kindern und resultierendes Schilddrüsenkrebsrisiko abgeleitet wurden. Diese Faktoren sind mit den Strahlenexposition schwach negativ korreliert (Punkt 5), und verursachen deshalb nur eine kleine Verzerrung.

Im Studiengebiet wurden zusätzlich im Beobachtungszeitraum 1990-99 gezielte Screening-Programme durchgeführt, die jedoch erst ab 1998 zur deutlichen Erhöhung der berichteten Fallzahlen führten. Insgesamt wurde die Inzidenzrate durch die Screening-Programme um ungefähr zehn Prozent erhöht. In den Simulationen wurde dieser Effekt deswegen noch nicht berücksichtigt (Punkt 6).

Auf Grundlage dieser Informationen wurden Simulationsszenarien entwickelt, die mit den vorliegenden Daten zur Strahlenexposition und Krebsinzidenz kompatibel sind. Mit den Ergebnissen lassen sich folgende Aussagen treffen.

Die Verzerrung der Risikoschätzer einer ökologischen Studie durch eine verbesserte Nachweis- und Berichtsrate und mögliche Korrelationen mit der Strahlenexposition ist moderat. In den betrachteten Szenarien lag sie für das absolute Zusatzrisiko pro Einheitsdosis bei -9 % und für das relative Zusatzrisiko pro Einheitsdosis bei -15%. Betragsmäßig ähnliche große Verzerrungen werden auch durch Messfehler verursacht.

Das absolute Zusatzrisiko einer ökologischen Studie ist deutlich kleiner als das Risiko einer Kohortenstudie, bei der die Inzidenz durch intensives Screening gegenüber der Inzidenz in der Bevölkerung stark erhöht ist.

Das relative Zusatzrisiko hängt davon ab, ob durch ein verbessertes Screening zusätzlich gefundene Mikrokarzinome eher spontane als strahleninduzierte Karzinome sind. Wenn dieser Effekt eine Rolle spielt, wird der Schätzer für das relative Risiko einer ökologischen Studie deutlich größer sein als in einer Kohortenstudie.

Zu den letzten beiden Punkten ist zu beachten, dass sich die Risikoschätzer in einer Kohorte mit intensivem Screening auf das Vorliegen eines solchen Screenings beziehen, während die ökologischen Studien Aussagen zum Risiko in der Bevölkerung zum Ziel haben.

6.3 Fazit

Die Probleme von ökologischen Studien wurden vielfach diskutiert und müssen bei der Interpretation von ökologischen Studien beachtet werden. Durch die Verzerrungen (ökologischer Trugschluss) sind in vielen Situationen Aussagen aus ökologischen Studien nicht zuverlässig. Allerdings sollte in Einzelfällen geprüft werden, wie groß die potenzielle Verzerrung ist. Zur Bewertung von quantitativen Risikoschätzern aus ökologischen Studien sind Simulationsstudien erforderlich und hilfreich, um die möglichen Quellen der Verzerrung einzeln und gemeinsam zu untersuchen. Allgemein gültige Aussagen über die Qualität lassen sich nicht machen, sondern es muss in jedem Fall eine Bewertung der Annahmen und ihrer Auswirkungen erfolgen. Für die Tschernobyl-Daten zeigten die Simulationsstudien jedoch, dass unter realistischen Annahmen, die Verzerrungen als moderat angesehen werden können.

7 Literatur

Aldrich TE, Andrews KW, Liboff AR. (2001) Brain cancer risk and electromagnetic fields (EMFs): assessing the geomagnetic component. *Arch Environ Health* 56(4):314-9.

Appetecchia M., Mecule A., Sciarretta F. (2002) A long-standing cystic lymph-node metastasis from occult thyroid carcinoma. *Journal of Experimental and Clinical Cancer Research* 2002; 21: 137-138.

Astakova LN, Anspaugh LR, Beebe GW, Bouville A, Drozdovitch VV, Garber V, Gavrillin YI, Khrouch VT, Kuvshinnikov AV, Kuzmenkov YN, Minenko VP, Moschik KV, Nalivko AS, Robbins J, Shemiakina EV, Shinkarev S, Tochitskaya SI, Waclawiw MA (1998), Chernobyl-related thyroid cancer in children of Belarus: A case-control study. *Radiat Res* 150: 349-356

Auvinen A., Hakama M., Arvela H., Hakulinen T., Rahola T., Suomela M., Soderman B., Rytomaa T. (1994): Fallout from Chernobyl and incidence of childhood leukaemia in Finland, 1976-92. *British Medical Journal* 309(6948):151-4.

Barton C. (2001): The incidence of childhood leukaemia in West Berkshire. *Med Confl Surviv.* 17 (1): 48-55.

BEIR VI (1999) Health Effects of Exposure to Radon. National Academic Press, Washington, D.C

Bender R and Blettner M (2000), Simulationen zum Cox-Modell mit Berkson-Fehler: Probleme und erste Ergebnisse. Gemeinsamer Workshop "Überlebenszeitanalyse" der AGs "Statistische Methoden in der Medizin" und "Statistische Modellierung" der Deutschen Region der IBS, "Statistische Methoden in der Epidemiologie" der DAE und "Statistische Methodik der klinischen Forschung" der GMDS. Heidelberg, 23.11.2000.

Benigni R, Giaimo R, Matranga D, Giuliani A. (2001) The potential impacts of climate variability and change on health impacts of extreme weather events in the United States. *Environ Health Perspect* 109(7):705-9.

Bithell JF, Dutton SJ, Draper GJ, Neary NM (1994): Distribution of childhood leukaemias and Non-Hodgkin's lymphomas near nuclear installations in England and Wales. *British Medical Journal* 309 (6953): 501-5.

Björk J, Strömberg U (2002) Effects of systematic exposure assessment errors in partially exologic case-control studies. *International Journal of Epidemiology* 31: 154-160.

Black RJ, Sharp L, Harkness EF, Mc Kinney PA (1994): Leukaemia and Non-Hodgkin's lymphoma: incidence in children and young adults resident in the Dounreay area of Caithness, Scotland in 1968-91. *J Epidemiol Community Health* 48 (3):232-6.

Bleuer JP, Averkin YI., Okeanov AE., Abelin T. (1997): The epidemiological situation of thyroid cancer in Belarus. *Stem Cells* 15(2): 251-254.

Boice J.D., Land C.E., Preston D.L. (1997): Ionizing radiation In: Schottenfeld D. Fraumeni J.F. (eds.) *Cancer Epidemiology and Prevention* (second edition).

Bouges S., Daures JP., Hebrard M. (1999): Incidence of acute leukaemias, lymphomas and thyroid cancers in children under 15 years, living around the Marcoule nuclear site from 1985 to 1995. *Rev Epidemiol Sante Publique* 47 (3): 205-17.

Brenner H, Savitz DA, Jöckel KH, Greenland S (1992), Effects of Nondifferential Exposure Misclassification in Ecologic Studies, *American Journal of Epidemiology* 135, 85-95

Breslow NE, Day NE (1980), *Statistical Methods in Cancer Research, Volume 1 – The Analysis of Case-Control Studies*. IARC

Breslow NE, Day NE (1987), *Statistical Methods in Cancer Research, Volume II – The Design and Analysis of Cohort Studies*. IARC

Buglova E.E., Kenigsberg, J.E., Sergeeva N.V. (1996) : Cancer risk estimation in Belarussian children due to thyroid irradiation as a consequence of the Chernobyl nuclear accident. *Health Physics* 71 (1): 45-49.

Carroll RJ, Ruppert D, Stefanski LA (1995), *Measurement Error in Nonlinear Models*, Chapman & Hall, London

Clarke EA., McLaughlin J., Anderson TW. (1991): *Childhood leukaemia around Canadian nuclear facilities - Phase I*. Ottawa, Canada: AECB.

Cohen B.L. (1990) Ecologic versus case-control studies for testing a linear no-threshold dose-response relationship . *International Journal of Epidemiology* 19, 680-684.

Cohen B.L. (1994) Invited Commentary: In Defense of Ecologic Studies for Testing a Linear-No Threshold Theory. *American Journal of Epidemiology* 139: 765-768. Cohen B.L. (1995) Test of the linear-no threshold theory of radiation carcinogenesis for inhaled radon decay products. *Health Physics* 68, 157-174.

Cohen B.L. (1995) Test of the linear-no threshold theory of radiation carcinogenesis for inhaled radon decay products. *Health Physics* 68: 157-174.

Cook-Mozaffari PJ., Darby SC., Doll R., Forman D., Hermon C., Pike MC., Vincent T. (1989): Geographical variation in mortality from leukaemia and other cancers in England and Wales in relation to proximity to nuclear installations, 1969-78. *Br J Cancer* 60 (2): 270.

Cotterill S.J., Pearce M.S. Parker I. (2001) Thyroid cancer in children and young adults in the North of England. Is increase incidence related to the Chernobyl accident? *European Journal of Cancer* 37, 1020-1026.

Danilyuk VV, Saiko S, Mikhailov RB. (2002) Recent achievements of Korosten inter-area medical diagnostic center after completion of the Chernobyl Sasakawa health and medical cooperation project: 1996-2000. In: Yamashita S, Shibata Y, Hoshi M, et al., eds. *Chernobyl: Message for the 21st century*. Amsterdam: Elsevier, International Congress Series 2002: 39-47.

Darby SC., Olsen JH., Doll R., Thakrar B., Brown PD., Storm HH., Barlow L., Langmark F., Teppo L., Tulinius H. (1992): Trends in childhood leukaemia in the Nordic countries in relation to fallout from atmospheric nuclear weapons testing. *British Medical Journal* 304 (6833): 1005-9.

Darby S, Deo H, Doll R (2001) A parallel analysis of individual and ecological data on residential radon and lung cancer in south-west England. *Journal of the Royal Statistical Society A* 164: 193-203.

Demidchik YuE, Demidchik EP. (2002) Thyroid carcinomas in Belarus 16 years after the Chernobyl disaster. In: *Proceedings of Symposium on Chernobyl - related Health Effects*. Tokyo: Radiation Effects Association, 2002: 66-77.

Dickman PW, Holm LE, Lundell MD, Boice JD, Hall P (2003), Thyroid cancer risk after thyroid examination with ¹³¹I: a population-based cohort study in Sweden. *Int J Cancer* 106: 580-587.

Dieckmann, H. (1992): Incidence of leukemia in the Elbmarsch area. *Das Gesundheitswesen* 54 (10): 592-6.

Diez Roux AV (2001) Investigating neighborhood and area effects on health. *Am J Public Health* 91(11):1783-9 Related Articles, Links.

Draper GJ., Stiller CA., Cartwright RA., Craft AW., Vincent TJ. (1993): Cancer in Cumbria and the vicinity of the Sellafield nuclear installation, 1963-90. *British Medical Journal* 306 (6870):89-94.

Durkheim E. (1951) *Suicide: a study in Sociology*. New York, Free Press.

Flegal KM; Brownie C; Haas J.D. (1986) The effect of exposure misclassification on estimation of relative risk. *American Journal of Epidemiologie* 123; 736-751.

Fuller WA (1987), *Measurement Error Models*. Wiley, New York

Garau J. (2001) Have guidelines had an impact on antimicrobial resistance? *Semin Respir Infect Sep*;16(3):210-4.

Gardner M.J., Snee M.P., Hall, A.J., Powell C.A., Downes S., Terrell J.D. (1990) Result of case-control study of leukaemia and lymphoma among young people near Sellafield nuclear plant in West Cumbria. *British Journal of Medicine* 300, 423-429.

Gilbert ES., Tarone R., Bouville A., Ron E. (1998): Thyroid Cancer rates and 131I doses from Nevada Atmosphere nuclear bomb tests. *J NatL Cancer Inst.* 90(21): 1654-60.

Gilman EA. and Knox EG. (1998): Geographical distribution of birth places of children with cancer in the UK. *Br J Cancer* 77: 842-849.

Goldsmith JR. (1992): Nuclear installations and childhood cancer in the UK: mortality and incidence for 0-9-year-old children, 1971-1980. *Sci Total Environ* 127 (1-2): 13-35.

Gordis L. (2000) *Epidemiology*. W.B. Saunders Company, Philadelphia, USA.

Grant WB. (2002) An estimate of premature cancer mortality in the U.S. due to inadequate doses of solar ultraviolet-B radiation. *Cancer* 5;94(6):1867-75.

Grant WB.(2002) An ecologic study of dietary and solar ultraviolet-B links to breast carcinoma mortality rates. *Cancer* 1;94(1):272-81.

Greenland S., Morgenstern H. (1989) Ecologic bias, confounding and effect modification. *International Journal of Epidemiology* 18, 269-274.

Greenland S. (1992) Divergent biases in ecologic and individual level studies. *Statistics in Medicine* 11, 1209-1223.

Greenland S, Robin J (1994) Acceting the Limits of Ecologic Studies: Drs.Greenland and Robins Reply to Drs. Piantadosi and Cohen. *American Journal of Epidemiology* 139: 769-771.

Greenland S, Samet JM (1994) Invited Commentary: Ecologic Studies - Biases, Misconceptions, and Counterexamples. *American Journal of Epidemiology* 139: 747-760.

Greenland, S. (Ed.) (1987). *Evolution of Epidemiologic Ideas*. Epidemiology Resources.

Greenland S. (2001) Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *European journal of cancer* 30: 1343-1350.

Greenland S (2002) A review of multilevel theory for ecologic analyses. *Statistics in Medicine* 21: 389-395.

Grosche B., Lackland D., Mohr L., Dunbar J., Nicholas J., Burkart W., Hoel D. (1999): Leukaemia in the vicinity of two tritium releasing facilities: a comparison of the Kruemmel Site, Germany and the Savannah River Site, South Carolina, USA. *J Radiol Prot* 19 (3): 201-2.

Guizard AV., Boutou O., Pottier D., Troussard X., Pheby D., Launoy G., Slama R., Spira A., ARKM (2001): The incidence of childhood leukaemia around the La Hague nuclear waste processing plant (France): a survey for the years 1978-1998. *J Epidemiol Community Health* 55 (7): 469-74.

Gunay U., Meral A., Sevinir B. (1996): Pediatric malignancies in Bursa, Turkey. *J Environ Pathol Toxicol Oncol* 15: 263-65.

Hattchouel JM., Laplanche A. (1995): Leukaemia mortality around French nuclear sites. *Br J Cancer*, 71: 651-3.

Heidenreich WF.; Kenigsberg J., Jacob P., Buglova E., Goulko G., Paretzke HG., Demidchik EP., Golovneva A. (1999): Time trends of thyroid cancer incidence in Belarus after the Chernobyl accident. *Radiation Research* 151: 617-625.

Hill C., Laplanche A. (1992): Mortalité par cancer autour d'installations nucléaires francaises entre 0-24 ans, France, Paris. INSERM/Douin.

Hjalmar U., Kulldorff M., Gustafsson G. (1994): Risk of acute childhood leukaemia in Sweden after the Chernobyl reactor accident. Swedish Child leukaemia group. British Medical Journal. 309 (6948): 154-7.

Hoffmann W., Dieckmann H., Schmitz-Feuerhake I. (1997): A cluster of childhood leukaemia near a nuclear reactor in northern Germany. *Archiv Environ Health* 52 (4):275-80.

Hoffmann W. (2002): Has fallout from the Chernobyl accident caused childhood leukaemia in Europe? A commentary on the epidemiologic evidence. *Eur J Public Health* 12 (1): 72-6.

Hopenhayn C, Stump ML, Browning SR. (2002) Regional assessment of atrazine exposure and incidence of breast and ovarian cancers in Kentucky. *Arch Environ Contam Toxicol* 42(1):127-36.

Hull BP, McIntyre PB, Sayer GP.(2001) Factors associated with low uptake of measles and pertussis vaccines--an ecologic study based on the Australian Childhood Immunisation Register. *Aust N Z J Public Health* 25(5):405-10.

IARC (2000): evaluation of carcinogenic risks to humans. Ionizing Radiation, Part I: X- and Gamma (Y)- Radiation, and Neutrons. Volume 75. Lyon, France.

Ivanov VK., Tolochko G., Lazarev VS., Shuvaeva L. (1993): Child Leukaemia after Chernobyl. *Nature* 365: 702.

Ivanov VK., Tolochko GV., Shuvaeva LP., et al. (1996): Childhood leukaemia in Belarus before and after the Chernobyl accident. *Radiat Environ Biophys* 35: 75-80.

Ivanov VK., Tsyb AF., Nilova EV., et al. (1997): Cancer risks in the Kaluga oblast of the Russian federation 10 years after the Chernobyl accident. *Radiat Environ Biophys* 34 : 3-6.

Ivanov VK., Tolochko GV., Shuvaeva LP., et al. (1998): Infant leukaemia after the Chernobyl accident. *Radiat Environ Biophys* 35: 75-80.

Ivanov VK., Gorsky AI., Tsyb AF., Maksyutov MA., Rastopchin EM. (1999): Dynamics of thyroid cancer incidence in Russia following the Chernobyl accident. *J Radiol Prot.* 19(4): 291-2.

Ivanov VK., Gorski AI., Maksyutov MA., Vlasov OK., Godko AM., Tirmarche M., Valenty M., Verger P. (2003): Thyroid cancer incidence among adolescents and adults in the Bryansk region of Russia following the Chernobyl accident. *Health Phys* 84 (1): 46-60.

Iwasaki T., Nishizawa K., Murata M. (1995): Leukaemia and lymphoma mortality in the vicinity of nuclear power stations in Japan, 1973-1987. *J Radiol Protect*, 15: 271-88.

Jablon S., Hrubec Z., Boie JD. Jr. (1991): Cancer in the population living near nuclear facilities. A survey mortality nationwide and incidence in two states. *JAMA* 7, 265(11): 1403-8.

Jacob P, Goulko G, Heidenreich W, et al. (1998), Radioactive iodine risk to children estimated. *Nature* 392: 31-32.

Jacob, P. Goulko, G., Heidenreich, WF., Likhtarev, I., Kairo, I., Tronko, N.D., Bogdanova, T.I., Kenigsberg, J., Buglova, E., Drotzdovich, V., Golovneva, A., Demitchik, E.R., Balonov, M., Zonova, I., Beral, V. (1998). Thyroid cancer risk to children calculated. *Nature* 392, 31-32.

Jacob, P. Kenigsberg, J. Zonova, I, Goulko, G., Buglova, E., Heidenreich, WF., Golovneva, A, Bratilova, A., Drotzdovich, V., Kruk, J., Pochtnaja, G.T., Balonov, M., Demitchik, E.R., Paretzke, H.G. (1999). Childhood exposure due to the Chernobyl accident and thyroid cancer risk in contaminated areas of Belarus and Russia. *British journal of cancer* 80(9), 1461 – 1469.

Jacob P, Kenigsberg Y, Goulko G, Buglova E, Gering F, Golovneva A, Kruk J, Demidchik EP (2000), Thyroid cancer risk in Belarus after the Chernobyl accident: Comparison with external exposures. *Radiat Environ Biophys* 39, 25-31.

Kaatsch P., Kaletsch U., Meinert R., Michaelis J. (1998): Krebserkrankungen bei Kindern in der Umgebung von deutschen kerntechnischen Anlagen. *Cancer Causes Control* 9: 529-533.

Kerber RA, Till JE, Simon SL, Lyon JL, Thomas DC, Preston-Martin S, Rallison ML, Lloyd RD, Stevens W (1993), *JAMA* 270: 2076-2082

Kinlen LJ. (1993): Can paternal preconceptional radiation account for the increase of leukaemia and non-Hodgkin's lymphoma in Seascale? *British Medical Journal* 306 (6894): 1718-21.

Kinlen LJ. (1993): Childhood leukaemia and non-Hodgkins lymphoma in young people living close to nuclear reprocessing sites. *Biomed Pharmacother* 47 (10): 429-34.

Kinlen LJ., Dickson M., Stiller CA. (1995): Childhood leukaemia and non-Hodgkin's lymphoma near large construction sites, with a comparison with Sellafield nuclear site. *British Medical Journal* 310 (6982): 763-8.

Kleinbaum D.G., Kupper L.L., Morgenstern H. (1982) *Epidemiologic research: principles and quantitative methods* Van Nostrand Reinhold, New York.

Koifman S, Koifman J, Meyer A. (2002) Human reproductive system disturbances and pesticide exposure in Brazil. *Cad Saude Publica* 18(2):435-45.

Körblein A. and Hoffmann W. (1999): Childhood Cancer in the vicinity of German nuclear power plants. Online im Internet: WWW: <http://mars.healthnet.org/MGS/MGSV6N1Korblein.html>. (05.10.99).

Körblein A. (2001): Stellungnahme zu "Krebsrate bei Kindern im Umkreis bayerischer Kernkraftwerke" von Dr. A. Körblein, 4.7.2001, Umweltinstitut München. Online im Internet WWW: http://www.kinderkrebsregister.de/stellungnahme_0801.html (2003-06-16).

Koushik A, King WD, McLaughlin JR. (2001) An ecologic study of childhood leukemia and population mixing in Ontario, Canada. *Cancer Causes Control* 12(6):483-90.

Kritidis P. & Florou H. (2001): Radiological impact in Greece of the Chernobyl accident- a retrospective synopsis. *Health Phys.* 80(5): 440-6.

Künzli, N., Tager I.B. (1997) The semi-individual study in air pollution epidemiology: a valid design as compared to ecologic studies. *Environmental Health Perspectives*, 105; 1078-1083.

Lagarde F., Pershagen G. (2001) Parallel analysis of individual and ecologic data on residential radon, cofactors, and lung cancer in Sweden. *American Journal of Epidemiology* 149, 268-274.

Laurier D., Rommens C., Drombry-Ringard C., Merle-Szeremeta A., Degrange J. (2000): Assessment of the risk of radiation-induced leukaemia in the vicinity of nuclear installations: the Nord- Cotentin radio-ecological study. *Rec Epidemiol Sante Publique* 48 (2): 224-36.

Laurier D., Grosche B., Hall P. (2002): Risk of Childhood leukaemia in the vicinity of nuclear installations – findings and recent controversies. *Acta Oncol.* 41 (1): 14-24.

Likhtarov I, Kovgan L, Vavilov S., Chepurny M., Bouville A., Luckyanov N., Jacob P., Voillequé P and Voigt G. (2005), Post-Chornobyl Thyroid Cancers in Ukraine. Report 1: Estimation of Thyroid Doses, *Radiation Research* 163, 125-136.

Little MP. (2002): The proportion of thyroid cancers in the Japanese Journal of radiological protection. *Journal of Radiological Protection* 22(3): 279-292.

Lopez-Abente G., Aragonés N., Pollán M., Ruiz M., Gandarillas A. (1999): Leukemia, lymphomas and myeloma mortality in the vicinity of nuclear power plants and nuclear fuel facilities in Spain. *Cancer Epidemiolo Biomarkers Prev*, 8 (10): 925-34.

Lopez-Abente G., Aragonés N., Pollán M. (2001): Solid-tumor mortality in the vicinity of uranium cycle facilities and nuclear power plants in Spain. *Environ Health Perspect* 109(7): 721-9.

Lubin J.H. (1998) On the discrepancy between epidemiologic studies in individuals of lung cancer and residential radon and Cohen's ecologic regression. *Health Physics* 75: 4-10.

Lubin J.H. (2002) The potential for bias in Cohen's ecological analysis of lung cancer and residential radon. *Journal of Radiological Protection* 22, 141-148

Lund E., Galanti MR. (1999): Incidence of Thyroid cancer in Scandinavia following fallout from atomic bomb testing: an analysis of birth cohorts. *Cancer Causes Control* 10 (3): 181-7.

Mangano JJ., Gould JM., Sternglass EJ., Sherman JD., Brown J., McDonnell W. (2002): Infant death and childhood cancer reductions after nuclear plant closings in the United States. *Archiv Environ Health.* 57 (1): 23-31.

Matanoski G.M., Sartwell P. Elliott E. et al. (1984) Cancer risks in radiologists and radiation workers. In: Boice J.D., Fraumeni J.E. (eds) *Radiation Carcinogenesis. Epidemiology and Biological Significance.* New York, Raven Press, 83-96.

McLaughlin JR., Clarke EA., Nishri ED., Anderson TW. (1993): Childhood leukaemia in the vicinity of Canadian nuclear facilities. *Cancer Causes Control.* 4 (1): 51-8.

Mettler FA., Williamson MR., Royal HD. et al. (1992): Thyroid nodules in the population living around Chernobyl. *Jama* 268:616-19.

Michaelis J., Kaletsch U., Burkart W., Grosche B. (1997): Infant Leukaemia after the Chernobyl accident. *Nature* 387: 246.

Michaelis J. (1998): Recent epidemiological studies on ionizing radiation and childhood cancer in Germany. *Int. J. Radiat. Biol* 73, 4: 377-81.

Morgenstern H. (1998): Ecologic Studies. In: Rothman K.J., Greenland S. (Hrsg.): *Modern Epidemiology*. Lippincott-Raven Publishers, Philadelphia.

Nair MK., Nabi KS., Amma NS., Gangadharan P., Jayalekshmi P., Jayadevan S., Cherian V., Reghuram KN. (1999): Population study in the high natural background radiation area in Kerala, India. *Radiat Res.* 152(6): 145-8.

Nikiforova NV, Elagin VV, Sivachenko TP, et al. (2002) Thyroid diseases among children and adolescents in Kiev region 15 years after the Chernobyl accident. In: Yamashita S, Shibata Y, Hoshi M, et al., eds. *Chernobyl: Message for the 21st century*. Amsterdam: Elsevier, International Congress Series 2002: 57-65

Noshchenko A., Moysich K., Bondar A. et al. (2001): Patterns of acute leukaemia occurrence among children in the Chernobyl region. *Int J Epidemiol* 30: 125-129.

Parkin D.M., Muir C.S., Whelan S.L. et al. (1992) *Cancer incidence in five continents, Vol 6*, IARC Scientific publication No 120, Lyon, International Agency for research on cancer.

Parkin DM., Cardis E., Masyer E., Friedl HP., Hansluwka H., Bobev D., Ivanov E., Sinnaeve J., Augustin J., Plesko I., et al. (1992): Childhood leukaemia following the Chernobyl accident: the European Childhood Leukaemia-Lymphoma Incidence Study (ECLIS). *European Journal Cancer*, 29a(1):87-95.

Parkin DM., Cardis E., Masuyer E., et al. (1993): Childhood leukaemia following the Chernobyl accident: the European Childhood Leukaemia –Lymphoma Incidence Study (ECLIS). *Eur J Cancer* 29A: 87-95.

Parkin DM., Clayton D., Black RJ., et al. (1996): Childhood leukaemia in Europe after Chernobyl: 5 year follow-up. *Br J Cancer* 73: 1006-12.

Parkin, D.M., Pisani, P. und Ferlay, J. (1999) Estimates of the worldwide incidence of twenty five major cancers in 1990. *International journal of cancer*, 80(6), 827-841.

Petridou E., Proukasis C., Tong D., Kassimos D., Athanassiadou-Piperopoulou F., Haidas S., Kalmanti M., Kolioukas D., Kosmidis H., Louizi A., et al. (1994): Trends in geographical distribution of childhood leukaemia in Greece in relation to the Tschernobyl accident. *Scand J Soc Med* 22(2): 127-31.

Petridou E., Trichopoulos D., Dessypris N., Flytzani V., Haidas S., Kalmantii MM., Kolioukas D., Kosmidis H., Piperopoulou F., Tzortzatou F. (1996): Infant leukaemia after in utero exposure to radiation from Chernobyl. *Nature* 25, 382 (6589): 352-3.

Piantadosi S (1994) Invited Comentary: Ecologic Biases. *American Journal of Epidemiology* 1994; 139: 761-764.

Prentice, R.L., Sheppard L. (1990) Dietary fat and cancer, consistency of the epidemiologic data and disease prevention that may follow from a practical reduction in fat consumption. *Cancer Causes and Controll*1, 81-89.

Prentice R.L. , Sheppard L. (1995) Aggregate data studies in disease risk factors. *Biometrica* 82, 113-115.

Prisyazhiuk A., Pjatak O.A., Buzanov Y.A., Reeves G.K., Beral V. (1991) Cancer in the Ukraine, post-Chernobyl. *Lancet* 338: 1334-1335.

Prisyazhiuk A., Gristchenko V., Zarkodonets V. et al. (1995): The time trends of cancer incidence in the most contaminated regions of the Ukraine before and after the Chernobyl accident. *Radiat Environ Biophys* 34: 3-6.

Reynolds P, von Behren J, Gunier RB, Goldberg DE, Hertz A, Harnly ME. (2002) Childhood cancer and agricultural pesticide use: an ecologic study in California. *Environ Health Perspect* 110(3):319-24.

Roetzheim RG, Gonzalez EC, Ramirez A, Campbell R, van Durme DJ. (2001) Primary care physician supply and colorectal cancer. *J Fam Pract* 50(12):1027-31.

Roman E., Beral V., Carpenter L., Watson A., Barton C., Ryder H., Aston DL. (1987): Childhood leukaemia in the West Berkshire and Basingstoke and North Hampshire District Health Authorities in relation to nuclear establishments in the vicinity. *Br Med J* 294(6572): 597-602.

Rommens C., Laurier D., Sugier A. (2000) Methodology and results of the Nord-Cotentin radioecological study. *J Radiol Prot* 20(4): 361-80.

Ron E., Lubin J., Schneider A.B. (1992) Thyroid cancer incidence. *Nature* 360: 113-113.

Ron, E., Lubin, J., Shore, R.E., Mabuchi, K., Modan, B., Pottern, L.M., Schneider, A.B., Tucker, M.A., Boice, J.D. (1995). Thyroid cancer after exposure to external radiation: a pooled analysis of seven studies. *Radiation Research* 141, 259-277.

Schneider AB, Ron E, Lubin J, Stovall M, Gierlowski TC (1993), Dose-response relationships for radiation-induced thyroid cancer and thyroid nodules: Evidence for the prolonged effects of radiation on the thyroid. *J Clin Endocrinol Metab* 77: 362-369.

Schwartz S (1994) The fallacy of the ecological fallacy: The potential misuse of a concept and the consequences. *American Journal of Public Health* 84: 819-824.

Sharp L., Black R.J., Harkness E.F., McKinney P.A. (1996): Incidence of childhood leukaemia and non-Hodgkin's lymphoma in the vicinity of nuclear sites in Scotland, 1968-93. *Occup Environ Med* 53 (12): 823-31.

Sheppard L, Prentice RL (1995) On the reliability and Precision of Within- and Between-Population Estimates of Relative Rate Parameters. *Biometrics* 51: 853-863.

Sheppard L, Prentice RL, Rossing MA (1996) Design considerations for estimation of exposure effects on disease risk, using aggregate data studies. *Statistics in Medicine* 15: 1849-1858.

Shibata Y., Yamashity S., Masyakim V.B., Panasyuk G.N., Nagataki S. (2001) 15 years after Chernobyl: new evidence of thyroid cancer. *Lancet* 2001; 358: 1965-1966.

Smerhovsky Z, Landa K, Rossner P, Brabec M, Zudova Z, Hola N, Pokorna Z, Mareckova J, Hurychova D (2001) Risk of cancer in an occupationally exposed cohort with increased level of chromosomal aberrations. *Environ Health Perspect* 109(1):41-5.

Sobolev, B., Heidenreich W.F., Kairo, I., Jacob P. (1997): Thyroid cancer incidence in the Ukraine after the Chernobyl accident: comparison with spontaneous incidences. *Radiat Environ Biophys* 36: 195-199.

Steiner GG. (2002) Cancer incidence rates and environmental factors: an ecological study. *J Environ Pathol Toxicol Oncol* 21(3):205-12.

Steiner M., Burkart W., Grosche B., Kaletsch U., Michaelis J. (1998): Trends in infant leukaemia in West Germany in relation to in utero exposure due to Chernobyl accident. *Radiation Environ Biophys.* 37(2):87-93.

Stidley CA, Samet JM (1994) Assessment of Ecologic Regression in the Study of Lung Cancer and Indoor Radon. *American Journal of Epidemiology* 139: 312-322.

Stezhko VA, Buglova EE, Danilova LI, et al. (2004), A cohort study of thyroid cancer and other thyroid diseases after the Chornobyl accident: Objectives, design and methods. *Radiat. Res.* 161: 481-492.

Sun Q., Akiba S., Tao Z., Yuan Y., Zou J., Morishima H., Kato H., Zha Y., Sugahara T., Wei L. (2000): Excess relative risk of solid cancer mortality after prolonged exposure to naturally occurring high background radiation in Yangjiang, China. *J Radiat Res* 41: 43-52.

Susser M. (1994) The logic in ecological. *American Journal of Public Health* 84, 5; 825-834.

Szklo M., Nieto F.J. (2000): *Epidemiology beyond the basics*. Aspen Publishers, Gaithersburg, Maryland.

Tao Z., Cha Y., Sun Q. (1999): Cancer Mortality in high background radiation area of Yangjiang, China, 1979-1995. *Zhonghua Yi Xue Za Zhi* 79(7):487-92.

Tao Z., Zha Y., Akiba S., Sun Q., Zou J., Li J., Liu Y., Kato H., Sugahara T., Wei L. (2000): Cancer mortality in the high background radiation areas of Yangjiang, China during the period between 1979 and 1995. *J Radiat Res* 41: 31-41.

Tondel M., Carlsson G., Hardell L., Eriksson M., Jakobsson S., Flodin U., Skoldestig A., Axelson O. (1996): Incidence of neoplasms in ages 0-19 y in parts of Sweden with high ¹³⁷Cs fallout after the Chernobyl accident. *Health Phys.* 71(6): 947-50.

Tong S (2000) Migration bias in ecologic studies. *European Journal of Epidemiology* 365-369.

Tronko MD, Bogdanova T.I., Komissarenko I.V.(1999) Thyroid carcinoma in Children and Adolescents in Ukraine after the Chernobyl nuclear accident. *Cancer* 86: 149-156.

Tronko MD., Bogdanova TI., Komissarenko IV., Epstein OV., Oliynyk V., Kovalenko A., Likhtarev IA., Kairo I., Peters SB., LiVolsi VA. (1999): Thyroid carcinoma in children and adolescents in Ukraine after the Chernobyl nuclear accident. Statistical data and clinicomorphologic characteristics. *Cancer* 86: (1): 149-156.

Tronko ND, Bogdanova TI, Epstein OV, et al. (2002a) Thyroid cancer in children and adolescents of Ukraine having been exposed as a result of the Chernobyl accident (15-year expertise of investigations). *Int J Radiat Med*; 4:222-232.

Tronko ND, Bogdanova TI, Likhtarev IA, et al. (2002b) Summary of the 15-year observation of thyroid cancer among Ukrainian children after the Chernobyl accident. In: Yamashita S, Shibata Y, Hoshi M, et al., eds. *Chernobyl: Message for the 21st century*. Amsterdam: Elsevier, International Congress Series 2002: 77-83.

Tronko ND, Bobilyova OO, Bogdanova TI, et al. (2003) Thyroid gland and radiation (Ukrainian-American Thyroid Project). In: Shibata Y, Yamashita S, Watanabe M, et al., eds. *Radiation and Humankind*. Amsterdam, Elsevier, International Congress Series;1258:91-104.

Tukiendorf A. (2001): An ecological analysis of leukaemia incidence around the highest Cs concentration in Poland. *Cancer causes and control* 12: 653-659.

UNSCEAR (1994), Sources and effects of ionizing radiation. United Nations Scientific Committee on the Effects of Atomic Radiation. UNSCEAR 1994 Report to the General Assembly. New York: United Nations.

UNSCEAR (2000) Exposures and effects of the Chernobyl accident, Annex J.

Viel JF., Pobel D., Carre A., (1995): Incidence of leukaemia in young people around the La Hague nuclear processing plant: a sensitivity analysis. *Stat Med*. 14: 2459- 72.

Waller LA., Turnbull BW., Gustaffsson G., Hjalmarsson U., Andersson B. (1995): Detection and assessment of clusters of disease: an application to nuclear power plant facilities and childhood leukaemia in Sweden. *Stat Med* 14 (1): 3-16.

Walter SD (1991) The Ecologic Method in the Study of Environmental Health. II. Methodologic Issues and Feasibility. *Environmental Health Perspectives* 94.

Webster T. (2000) Can semi-individual studies have ecologic bias? *Epidemiology* 11; 95.

Webster T. (2002) Does the spectre of ecologic bias haunt epidemiology? *International Journal of Epidemiology* 31; 161-162.

Wei L. & Sugahara T. (2002): Risk assessment based on an epidemiological study in a high Background radiation area – A china-Japan cooperative research. *International Congress Series* 1225: 267-275.

Wiegersma PA, Hofman A, Zielhuis GA (2001) Evaluation of community-wide interventions: the ecologic case-referent study design. *Eur J Epidemiol* 17(6):551-7.

Woolner L.B., Lwmmmon M.L., Beahrs O.H., Black B.M., Keating F.R. (1960) Occult papillary carcinomas of the thyroid gland: a study of 140 cases observed in a 30 year period. *Journal of Clinical Endocrinology* 20, 89-105.

World Health Organization (1996). Medical consequences of the Chernobyl accident. Scientific Report by International Project «IPHECA». Geneva, World Health Organization.

Young FW, Lyson TA (2001) Structural pluralism and all-cause mortality. *Am J Public Health* Jan; 91(1):136-8.

Zaridze DG., Li N., Men T., Duffy SW. (1994): Childhood Cancer incidence in relation to distance from the former nuclear testing site in Semipalatinsk, Kazakhstan. *International Journal of Cancer*. 59(4): 471-5.