

Epidemiologie

Epidemiologische Tatsachen fordern „gegenwärtigen strahlenbiologischen Kenntnisstand“ heraus

Anmerkungen zur Stellungnahme der Strahlenschutzkommission (SSK) zur KiKK-Studie

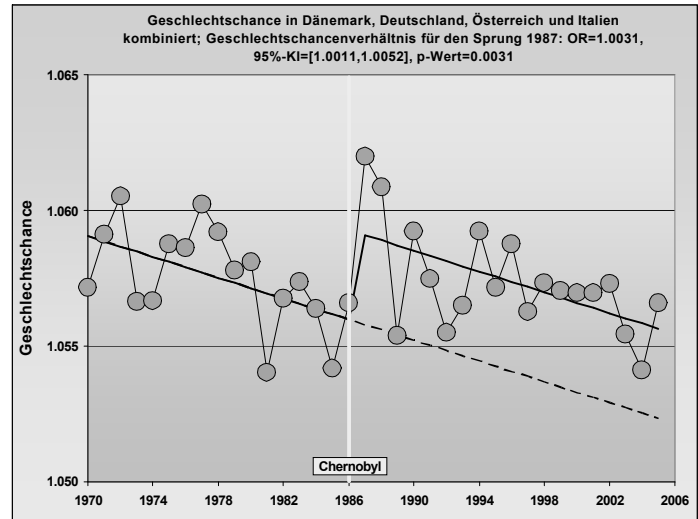
Von Hagen Scherb¹

Die Stellungnahme der Strahlenschutzkommission (SSK) zur KiKK-Studie strapaziert den Begriff „gegenwärtiger Kenntnisstand“. Es ist zu kritisieren, daß sich die SSK dabei auf ein seit Jahrzehnten veraltetes „Wissen“ beruft. Mittlerweile ist es in vielen epidemiologischen Studien gelungen, die Wirkung von Niedrigdosisstrahlung zweifelsfrei nachzuweisen. Die Verdoppelungsdosen für angeborene Fehlbildungen (genetische Effekte, zum Beispiel Down Syndrom) liegen teilweise in der Größenordnung von 1 Millisievert pro Jahr (mSv/a), und das Geschlechtschancenverhältnis (sex odds ratio) pro mSv/a liegt bei 1.015 (z.B. [1], [2], [3]). Dem muß man gegenüberstellen, daß das Wissenschaftliche Komitee der Vereinten Nationen für die Wirkung von Atomstrahlung UNSCEAR [4] von einem relativen Risiko pro Gray (Gy) für Fehlbildungen von 1,033 ausgeht. Das entspricht etwa einer Verdoppelungsdosis von 20 Gray: *The estimate of risk for congenital abnormalities is about 2,000 cases per million live births (compared to 60,000 cases per million live births).*

Damit wird das Risiko durch Niedrigdosisstrahlung in Bezug auf Fehlbildungen (con-

genital abnormalities) nach dem sogenannten „strahlenbiologischen Fachwissen“ um den Faktor 20.000 unterschätzt. Weil UNSCEAR [4] auf ältere Tierversuche mit für genetische Fragestellungen viel zu kleinen Stichproben Rekurs nimmt, verwundert das nicht. Kanzerogenese wird durch Mutation von genetischem Material hervorgerufen, insbesondere gelten strahleninduzierte DNA-Schäden in der Fetalentwicklung als Kausalfaktoren der frühkindlichen Leukämie. Offenbar leidet die Risikobewertung der Strahlenkanzerogenese unter einer grundlegenden Fehleinschätzung der strahlengenetischen Realität und Risiken. Vor diesem Hintergrund ist die feste Überzeugung der SSK, die Strahlenbelastung um Kernkraftwerke sei zur Erklärung des erhöhten Krebsrisikos um mindestens einen Faktor 1.000 zu klein, nicht nachvollziehbar und verfehlt.

Ich möchte eines von zahlreichen markanten Beispielen aufzeigen (siehe Abbildung), die es nach dem beschworenen „strahlenbiologischen Fachwissen“ nicht geben dürfte: Die vier Länder Dänemark, Deutschland, Österreich und Italien bilden quasi den Abschnitt eines groben „äquidistanten geographischen Rings“ um Tschernobyl auf dem Globus. Dort beobachten wir einen säkularen Sprung des Geschlechtsverhältnisses des Menschen bei der Geburt (sex odds) im Jahre 1987 um 0,31 Prozent. Dieser signifikante, langfristige Sprung



(change-point) entspricht circa 43.000 verlorenen Mädchen in den vier – insgesamt eher moderat kontaminierten – Ländern von 1987 bis 2005. (Bei einer möglicherweise realistischeren sex odds der verlorenen Kinder von 3:10 sind es 72.000 verlorene Kinder.) Weil dieser Effekt linear von der Dosis abhängt, muß in der Zusammenfassung von Zentral- und Osteuropa mit der Größenordnung von 500.000 „missing children“ nach Tschernobyl bis heute gerechnet werden. Effekte in dieser Dimension werden vom „strahlenbiologischen Fachwissen“ kategorisch ausgeschlossen, und dieses „Fachwissen“ widerlegt sich damit selbst. Theorie und weitere Beispiele siehe [3] und [5]. In Westeuropa, insbesondere in Frankreich und auch in den USA sind 1987 keine Sprünge in den Trends des Geschlechtschancenverhältnisses nachweisbar. Dagegen liegen in Belarus, Ukraine und Russischer Föderation ausgeprägte und hochsignifikante Sprünge vor.

Die SSK kritisiert die Methodik der KiKK-Studie. In der KiKK-Studie wurde mit gutem Grund und in Unkenntnis der Abstandsverteilungen der Daten ein reziprokes Abstandsgesetz für das Krebsrisiko postuliert: Nach UNSCEAR ist nämlich die Dosis ebenfalls einem reziproken Abstandsgesetz unterworfen. Damit ist der Abstand zum

KKW in der KiKK-Studie ein Surrogat für die Dosis und die in der KiKK-Studie nachgewiesene signifikante kontinuierliche Abstandsabhängigkeit ist ein Indiz für einen ursächlichen Zusammenhang zwischen Strahlendosis und Krebsrisiko. Natürlich ist das Dosisgesetz in realen Daten „verrauscht“. Allerdings führt die nicht-differentielle Fehlklassifikation zur Unterschätzung und nicht zur Überschätzung von Effekten.

Die SSK kapriziert sich darauf, daß der nachgewiesene Effekt nur innerhalb eines Kreises von 5 Kilometer (km) Radius um die Kernkraftwerke vorliege. Ein gravierender methodischer Mangel der SSK-Stellungnahme besteht darin, daß die statistische Power für die Fragestellungen unter Ausschluß des 5-km Innenkreises nicht berechnet wird. Diese Power (alle Krebserkrankungen) beträgt nur 36 Prozent, und deshalb kann ein weitreichender Effekt nicht mit genügend hoher Wahrscheinlichkeit ausgeschlossen werden. Die Annahme der Hypothese „Keine Fernwirkung > 5 km“ hat also eine Irrtumswahrscheinlichkeit von 64 Prozent. Trotzdem propagiert die SSK den Ausschluß einer Fernwirkung. Ab 4.800 Fällen und 14.200 Kontrollen aufwärts würde eine ausreichende Power von mindestens 80 Prozent entstehen. Erst wenn dann immer noch kein signifikanter „Ef-

¹ Dr. Hagen Scherb, Institut für Biomathematik und Biometrie, Helmholtz Zentrum München – Deutsches Forschungszentrum für Gesundheit und Umwelt, Ingolstädter Landstraße 1, D-85764 Neuherberg

fekt > 5 km“ vorläge, wäre die Annahme der SSK einigermaßen gedeckt, mit einer Irrtumswahrscheinlichkeit von dann 20 Prozent statt jetzt 64 Prozent. In einer Informationsveranstaltung zur KiKK-Studie am 17.10.2008 im Helmholtz Zentrum München wies ich auf diese Interpretationsprobleme hin. Mir wurde entgegnet: „Die Musik spielt innerhalb fünf Kilometer“.

Für die Bevorzugung der Abstandskategorie „< 5 km“ gibt es im Rahmen einer differenzierten Analyse keine vernünftige Begründung. Die Vorstellung, ein Agens oder ein Risikofaktor halte einen definierten, symmetrischen Abstand ein, ist weltfremd. Es ist offensichtlich, daß diese Kategorie nur deshalb attraktiv erscheint, weil der mutmaßliche Effekt jenseits von 5 Kilometer für sich alleine nicht mehr signifikant ist, was aber mangels Power auf keinen Fall bedeutet, dort sei ein Risiko von Null nachgewiesen (absence of evidence is not evidence of absence). Die Festlegung auf eine willkürliche Abstandskategorie ist bezüglich der statistischen Power der kontinuierlichen Abstandsauswertung unterlegen, was die SSK auch weiß und einräumt. Gleichwohl ignoriert die SSK, daß zum Beispiel das Risiko im 15 km-Umkreis nach den 5 km-Strata des KiKK-Berichts (S. 46) immer noch signifikant erhöht ist: $OR=1,200$, $95\%-KI=[1,012, 1,421]$, $p=0.0350$. Es stellt sich hier die Frage, warum die SSK eine unterlegene statistische Methode höher bewertet als die adäquate, vor Kenntnis der Daten festgelegte Methode der KiKK-Studie. Die SSK kennt nun die Daten und optimiert retrospektiv ihre Re-Analyse im Hinblick auf nichtsignifikante Resultate. Die SSK verhält sich wie ein „umgekehrter texanischer Scharfschütze“, der eine Zielscheibe neben die Einschußlöcher in der Wand malt.

Wegen der Unkenntnis der Abstandsverteilungen in der Planungsphase der KiKK-Studie ist die von der SSK betonte Unterscheidung in einen Hypothesen-generierenden und einen Hypothesen-testenden Studienteil obsolet. Die SSK verfährt ganz nach der Devise „teile (die Daten) und herrsche (über die Statistik)“. Sie scheint nicht zu bedenken, daß darunter die statistische Genauigkeit (Power) leidet.

Die SSK behauptet einerseits, man wisse nicht, woher die Krebserkrankungen um die Kernkraftwerke kommen. Andererseits wisse man aber genau, die Strahlung sei viel zu niedrig. Das ist unlogisch und unglaubwürdig. Im Einzelgespräch räumen SSK-Mitglieder durchaus ein, daß schon aus prinzipiellen Gründen Strahlung als Ursache der um die KKW's erhöhten Leukämien nicht ausgeschlossen werden kann. Man fragt sich, warum das in der KiKK-Stellungnahme der SSK nicht zum Ausdruck kommt, sondern der gegenteilige Eindruck erweckt werden soll.

Die SSK wäre gut beraten, sich auf Fakten zu berufen und nicht auf einen diffusen, von Lernfähigkeit und Einsichtsbereitschaft abhängigen „gegenwärtigen Kenntnisstand“. Allerdings ist dieser Begriff bequem: Er delegiert Verantwortung und kaschiert Unwissen. Im Sinne anerkannter Wissenschaftslogik (occam's razor) sowie angesichts der angehäuften Evidenz inklusive der KiKK-Studie und angesichts des faktisch nicht existenten „etablierten strahlenbiologischen Wissens“ auf populationsgenetischer Ebene, ist bis zu einem seriösen wissenschaftlichen Beweis des Gegenteils davon auszugehen, daß Atomkraftwerke durch Emissionen ihrer Kernreaktoren das Krebsrisiko in ihrer Umgebung erhöhen.

1. Métnekil, Czeizel (2005). Increasing total prevalence rate of

cases with Down syndrome in Hungary, *European Journal of Epidemiology*, 20: 525-535.

2. Lazjuk et al. (1998). Genetic consequences of the Chernobyl accident for Belarus Republic. *Gijutsu-to-Ningen*. 283: 26–32. (Artikel in Japanisch). Artikel in Englisch: <http://www.rri.kyoto-u.ac.jp/NSRG/reports/kr21/kr21pdf/Lazjuk.pdf> (29/10/2008).

3. Scherb, Voigt (2007). Trends in the human sex odds at birth in

Europe and the Chernobyl Nuclear Power Plant accident. *Reproductive Toxicology*, 23: 593-599.

4. UNSCEAR 2001 Report, Hereditary Effects of Radiation, Scientific Annex, p. 82.

5. Scherb, Voigt (2008). Analytical ecological epidemiology: Exposure-response relations in spatially stratified time series. *Environmetrics*, published Online: 12 Sep 2008

Epidemiologie

Feuchtgebiete – der neue Bestseller der Strahlenschutzkommission

Sind die Leukämien bei Krümmel und anderswo durch Wasserdampf entstanden?

Von Inge Schmitz-Feuerhake¹

Es war nicht zu erwarten, dass die deutsche Strahlenschutzkommission aus dem Ergebnis der Kinderkrebsstudie bei Kernkraftwerken von 2007 (KiKK) in einer nachvollziehbaren Weise Lehren ziehen würde. Als weiland 1991 die Leukämiehäufung beim AKW Krümmel unübersehbar geworden war, verkündete der damalige Umweltminister Klaus Töpfer vor Ort, Radioaktivität sei als Ursache auszuschließen, es handele sich wahrscheinlich um eine Virusinfektion. Auf erstaunte Nachfrage von Journalisten teilte er mit, das hätten ihm seine Berater in der Strahlenschutzkommission (SSK) gesagt.

Die SSK hielt es für unwissenschaftlich, der Radioaktivitätsthese überhaupt nachzugehen. Sie publizierte 1994 eine Stellungnahme, in der die Clusterbildung durch Mikroepidemien als Erklärung favorisiert wird. Auf diese wurde von ihr bis zur KiKK-Studie immer verwiesen und es hieß,

bevor neue Erkenntnisse zur Erklärung des Krümmelclusters gewonnen werden könnten, müssten die molekulargenetischen Grundlagen der Leukämieentstehung genauer erforscht werden.

Um ihrer Beurteilung der KiKK-Studie Gewicht zu verleihen, hat die SSK ausländische Expertinnen herangezogen. Prof. Sarah Darby hat in England erfolgreich die Strahlenbedingtheit der Leukämien bei der staatlichen Wiederaufarbeitungsanlage für Kernbrennstoffe Sellafield bestritten, deren Plutoniumausstöße bekanntlich die irische See und die Ostküste Irlands verseuchen. Auch bei den im Folgenden entdeckten weiteren Leukämiehäufungen bei britischen Nuklearanlagen war sie im Geschäft und stieg darüber zur Chefepidemiologin ihres Landes auf. Sie ist sozusagen die Maria Blettner Großbritanniens.

Sie ist diejenige, die einen lächerlich kleinen nicht-signifikanten Befund um Orte mit „geplanten“ aber nicht in Betrieb gegangenen Nuklearanlagen zum großen Gegenbe-

¹ Prof. Dr. Inge Schmitz-Feuerhake, ingesf@uni-bremen.de

Glossar

95%-Konfidenzintervall (KI; oder Vertrauensbereich zum Niveau 95%). Das 95%-Konfidenzintervall wird in durchschnittlich 95% aller Fälle, in denen man Daten erhebt und ein Konfidenzintervall berechnet, den wahren aber unbekanntem Parameter (z. B. das gesuchte relative Risiko) enthalten. Oder suggestiv: Das Konfidenzintervall umfaßt alle Parameter, die mit der Beobachtung verträglich sind. Es gibt eine Denkschule, die die ausschließliche Verwendung von Konfidenzintervallen propagiert, und von der Verwendung von p-Werten wegen häufigen Missbrauchs (s. u.) abrät. Neutral ist es, stets das 95%-KI und den p-Wert gemeinsam anzugeben.

Change-Point. Ein Change-Point-Modell ist ein spezielles Regressionsmodell, mit dem versucht wird, Heterogenitäten in den Daten, die sich z. B. grafisch als Sprünge (jump) und/oder Knicke (broken stick) in Verläufen darstellen, zu modellieren und diese Sprünge und Knicke zu quantifizieren bzw. deren Signifikanz zu testen bzw. entsprechende p-Werte und Konfidenzintervalle anzugeben. Technisch wird dies mit Dummy- bzw. Indikatorvariablen, die nur die Werte 0 und 1 annehmen, realisiert. Die Abbildung zeigt ein einfaches Change-Point-Modell.

Geschlechtschance der Geburten (sex odds oder weniger systematisch: sex ratio). Quotient aus einer Anzahl von männlichen Geburten (Zähler) und der entsprechenden Anzahl von weiblichen Geburten (Nenner). Beispiel: In der Kombination von Dänemark, Deutschland, Österreich und Italien (s. Abbildung) wurden 1970 1,104,062 Jungen und 1,044,367 Mädchen geboren. Die Geschlechtschance ist daher 1.0572.

Geschlechtschancenverhältnis (sex odds ratio). Quotient zweier Geschlechtschancen. Beispiel: In der Abbildung entspricht der Abstand der beiden Linien ab 1987 einem Geschlechtschancenverhältnis von 1.0031.

Gray bzw. Sievert (Sv). 1 Sv ist ein biologisch bewertetes Dosismaß, das von dem physikalischen Maß 1 Gray (1 Joule/kg) abgeleitet wird. Die Angabe einer Dosis in Sv in einer konkreten Belastungssituation beruht auf einer Vielzahl von Annahmen und ist dementsprechend subjektiv und damit insgesamt problematisch. Ein leichter objektivierbares Expositionsmaß, wie etwa das kBq/m² (Fallout) oder eine direkte Transformation davon, kann daher u. U. günstiger sein als das nicht direkt beobachtbare Sv. Die tausendsten Teile sind mGy und mSv.

Irrtumswahrscheinlichkeit. Wahrscheinlichkeit eine Hypothese anzunehmen, wenn sie falsch ist. Beispiel: Ein üblicher Test zum Niveau 5% hat höchstens eine 5%-ige Irrtumswahrscheinlichkeit, die Nullhypothese abzulehnen oder äquivalent die Alternativhypothese anzunehmen, wenn tatsächlich die Nullhypothese zutrifft.

Occam's razor. Bezeichnet eine Sparsamkeitsmaxime in den Wissenschaften. Man soll keine unnötigen Annahmen machen. Einfache wissenschaftliche Vorstellungen bzw. Modelle sind leichter zu falsifizieren als komplexe, und sind so dem wissenschaftlichen Erkenntnisfortschritt eher zuträglich.

Odds (Chance). Ist p mit $0 < p < 1$ eine Wahrscheinlichkeit, so nennt man $p/(1-p)$ Chance oder im Englischen odds. In Wahrscheinlichkeitstheoretischen und statistischen Anwendungen ist die odds eher geeignet als die Wahrscheinlichkeit p selbst, ansonsten entsprechen sich p und odds aber ein-eindeutig.

Odds ratio (Chancenverhältnis, OR). Sind p und q zwei Wahrscheinlichkeiten, so nennt man das Verhältnis ihrer odds $(p/(1-p))/(q/(1-q))$ die odds ratio oder im Deutschen das Chancenverhältnis.

Power (statistische Power, Trennschärfe). Wahrscheinlichkeit die Nullhypothese abzulehnen, wenn eine Alternativhypothese zutrifft. Gegenstück zur Irrtumswahrscheinlichkeit. Beispiel: Ein üblicher Test zum Niveau 5% hat, im Grenzübergang, falls die Alternativhypothesen stetig in die Nullhypothese übergehen, in der Nullhypothese eine Power von nur 5%.

p-Wert (p oder empirisches Signifikanzniveau/Irrtumswahrscheinlichkeit). Wahrscheinlichkeit, eine Beobachtung zu erhalten, die gleich oder extremer ist (Zweiseitigkeit) als eine aus Daten gewonnene Beobachtung. Diese Wahrscheinlichkeit wird unter der Annahme einer geeigneten Verteilung (Nullhypothese-Verteilung) berechnet. Die Fiktion in dieser Vorstellung sollte nicht übersehen werden: Unendliche Wiederholbarkeit der Beobachtungen, Vorliegen einer bestimmten Verteilung der Beobachtungen, wo tatsächlich i.d.R. nur eine einmalige Beobachtungsserie vorliegt. Beispiel: In der Abbildung ist ein zweiseitiger p-Wert von 0.0031 angegeben. D.h.: In ca. 30 von 10000 Fällen wird also ein Sprung nach oben von 1.0031 oder höher, oder ein Sprung nach unten von $1/1.0031=0.9969$ oder tiefer rein zufällig auftreten, auch wenn der Trend der Geschlechtschance nach 1987 tatsächlich nicht „echt gestört“ wurde. Weil 30 in 10000 relativ selten ist, wird man geneigt sein zu glauben, dass der Trend ab 1987 „echt gestört“ wurde. Man wird nun andere, ähnliche Datensätze untersuchen, um die Hypothese zu stützen oder zu schwächen, bzw. um eine mögliche quantitative Abhängigkeit der Sprunghöhe von der Kollektivdosis auf Länderebene zu untersuchen. Hinweis: Ein weit verbreiteter Missbrauch des p-Werts liegt darin, Evidenz für „keinen Effekt“ oder „keinen Unterschied“ zu suggerieren, wenn der p-Wert größer ist als typischerweise 0.05. Auch die SSK begeht diesen Fehler im aktuellen Fall. Diese nahe liegende und oft unbewusste Missbrauchsmöglichkeit des p-Werts ist der Grund, warum manche Wissenschaftler den völligen Verzicht auf den p-Wert propagieren. Dieser Missbrauch, so hofft man, sei mit dem alternativen Konfidenzintervall nicht so leicht möglich, weil das Konfidenzintervall stets den vollen Umfang der mit der Beobachtung verträglichen Parameter (Risiken) ausweist, und damit weniger zu einer unwissenschaftlichen Dichotomisierung Anlass gibt.

Relatives Risiko (RR). Das relative Risiko ist der Quotient aus der Wahrscheinlichkeit für Erkrankung bzw. Tod in einer exponierten Population (Zähler) dividiert durch die Wahrscheinlichkeit für Erkrankung bzw. Tod in einer entsprechenden nicht exponierten Population (Nenner). Beispiel: $RR=1.33$ pro 1mSv/a bedeutet z. B., daß die Wahrscheinlichkeit für Erkrankung in einer Population, die der Dosis 1mSv/a ausgesetzt ist, 33% höher ist als die Wahrscheinlichkeit für Erkrankung in einer vergleichbaren nicht exponierten Population. (Technischer Hinweis: Das RR wird häufig als odds ratio (OR) mittels logistischer Regression geschätzt. Bei seltenen Erkrankungen (Zutreffen der ‚rare disease assumption‘) kann nämlich die odds ratio als gute Approximation für das relative Risiko angesehen werden).

Texanischer Scharfschütze. Der Begriff des Texanischen Scharfschützen und des damit verbundenen Phänomens leitet sich aus einem amerikanischen Witz her: Ein texanischer Hinterwäldler schießt mit einer Schrotflinte auf eine Mauer und zeichnet anschließend um jedes Einschußloch eine Zielscheibe. Anschließend brüstet er sich vor seinen Freunden mit seiner Treffgenauigkeit (WIKIPEDIA). Texanischer Scharfschütze zu sein, wird Wissenschaftlern und insbesondere Statistikern gelegentlich – zu recht, manchmal zu unrecht – vorgeworfen.

Verdoppelungsdosis. Dosis, die die spontan auftretende Häufigkeit eines Merkmals verdoppelt. Beispiel: Wenn gesagt wird, 10mSv erhöht die frühkindliche Leukämie um 40% (d.h. relatives Risiko (RR)/10mSv=1.4), dann ist das RR pro mSv 1.034, und die Verdoppelungsdosis 20.6 mSv. Hinweis. Rechenweise im multiplikativen Modell. Analogie zur Zinseszinsrechnung und in der Physik zu Zerfallskonstante und Halbwertszeit.

Dr. Hagen Scherb, November 2008