

Bewertung genetischer Forschungsergebnisse¹

Methoden und vermeidbare Fehler

von Anja Victor*

ABSTRACT

Viele menschliche Erkrankungen sind ganz oder teilweise auf genetische Ursachen zurückzuführen. Diese Erkenntnis hat in Verbindung mit technischen Fortschritten in der medizinischen Genetik einen regelrechten „Genetik-Boom“ ausgelöst. Neue Methoden erlauben es, Gene zu charakterisieren und genetische Varianten zu identifizieren, die für Erkrankungen verantwortlich sind. Zahlreiche Studien, die mithilfe von Markern nach genetischen Veranlagungen von Erkrankungen suchen, liegen bereits vor. Die Ergebnisse dienen nicht nur der weiteren epidemiologischen Grundlagenforschung zu Ursachen, Folgen und Verbreitung von Krankheiten. Verstärkt finden sie auch Eingang in der medizinischen Praxis bei der Auswahl der für einen Patienten optimalen Therapie. Umso mehr ist deshalb eine kritische Revision und Bewertung entsprechender Studien geboten. Dies gilt nicht nur für die biologisch-technischen Aspekte der Untersuchungen. Vielmehr sind es auch die statistischen Auswertungen und die dafür verwendeten Methoden, die hinterfragt werden müssen.

Schlüsselworte: Genetische Epidemiologie, genetische Assoziationsstudien, statistische Verfahren und Bewertungen

¹ Der vorliegende Artikel beruht auf einem vom AOK-Bundesverband in Auftrag gegebenen Gutachten (IMBEI, Universität Mainz: Kriterien zur Bewertung genetischer Testverfahren, Februar 2007).

Many human diseases are totally or partly due to genetic causes. This fact, coupled with technical progress, has triggered a virtual „genetics boom“ in the area of genetics in medicine. New methods allow to characterise genes and identify genetic variants which are responsible for diseases. Numerous studies looking for genetic predisposition of illnesses by means of genetic markers already exist. The results not only advance further epidemiological research on causes, consequences and spreading of diseases. More and more, they are also used in medical practice when selecting the optimal therapy for a patient. Therefore, a critical revision and evaluation of genetic studies is all the more required. This does not only apply to the biological-technical aspects of research – the statistical evaluations and the methodology must be analysed as well.

Keywords: Genetic epidemiology, genetic association studies, statistical procedures and evaluations

¹ The article presents the findings of an expert opinion commissioned by the federal Association of the AOK (IMBEI, Universität Mainz: Kriterien zur Bewertung genetischer Testverfahren, Februar 2007).

1 Genetische Epidemiologie

Es gibt heute genetische Tests, die Erkrankungen oder Erkrankungswahrscheinlichkeiten von Menschen vorhersagen können, lange bevor die ersten Symptome auftreten. Man geht davon aus, dass genetische Tests Therapieentscheidungen unterstützen können. Beispielsweise in der Pharmakogenetik, deren Ausgangsthese es ist, dass viele Medikamentenwirkungen durch genetische Merkmale beeinflusst werden. Die Fortschritte in der Genomanalyse lassen den Einsatz genetischer Testverfahren in

Diagnostik und Prognostik immer stärker in den Vordergrund treten. Das Gebiet der genetischen Epidemiologie versucht dabei, den gemeinsamen Einfluss von Genen und Umwelteinflüssen auf die Entstehung von Krankheiten zu quantifizieren. Die genetische Epidemiologie ist zwar noch relativ jung, erlebt jedoch derzeit einen starken Entwicklungsschub.

Genetische Studien in der Epidemiologie haben viele Ähnlichkeiten mit klassischen epidemiologischen Untersuchungen. So wird auf klassische epidemiologische Verfahren wie zum Beispiel Fall-Kontroll-Studien aufgebaut. Auf der anderen Seite gibt es

* Dr. rer. physiol. Anja Victor

Institut für medizinische Biometrie, Epidemiologie und Informatik
Universität Mainz, Obere Zahlbacher Straße 69, 55101 Mainz

Telefon: 06131 17-6853 · Fax: 06131 17-476853 oder 17-2968

E-Mail: victor@imbei.uni-mainz.de

Aspekte, die eine Abgrenzung notwendig machen und es rechtfertigen, die genetische Epidemiologie als ein Spezialgebiet der Epidemiologie zu bezeichnen. Allem voran unterscheiden sich die Art sowie auch die Anzahl der zugrunde liegenden Daten bei genetischen Untersuchungen.

Im Mittelpunkt vieler genetischer Untersuchungen stehen vor allem sogenannte genetische Marker (siehe auch das Glossar auf Seite 19). Sie können als Teil eines Gens innerhalb einer bekannten DNA-Sequenz für den Phänotyp eines Organismus mitverantwortlich sein. Von Interesse sind etwa die Single Nucleotide Polymorphisms (SNPs), die häufig in sogenannten Assoziationsstudien betrachtet werden. SNPs sind Variationen in den DNA-Sequenzen, kleinste Veränderungen in der Abfolge der vier Komplementärbasen Adenin, Guanin, Cytosin und Thymin, die den genetischen Code charakterisieren. Tritt eine Abweichung in der Abfolge der Basen in einem DNA-Strang bei mehr als einem Prozent im Genpool einer Population auf, so spricht man von einem SNP. Es lassen sich zum Beispiel durch SNP-Studien Gene oder Prädispositionsloki der genetisch bedingten Anlage oder Empfänglichkeit für Krankheiten aufdecken und auch die Krankheitsentstehung genauer als bisher charakterisieren.

In jüngster Zeit entdeckte man jedoch, dass sich viele der publizierten Ergebnisse genetischer Studien nicht reproduzieren ließen oder dass die Assoziation zwischen Gen und Krankheit überschätzt wurde (Ioannidis et al. 2001; Hirschhorn et al. 2002; Ioannidis

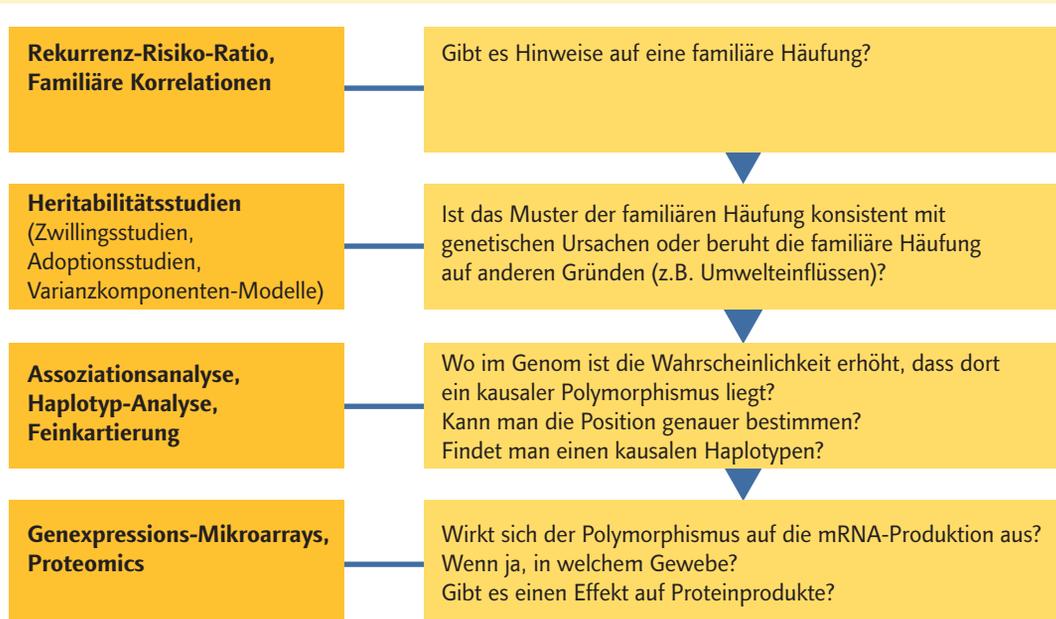
2003; Lohmüller et al. 2003). Die Umsetzung der genetischen Grundlagenforschung in klinischen Nutzen zieht sich bedeutend länger hin als erwartet oder blieb sogar ganz aus. Ebenso bilden auf dem Gebiet der komplexen Krankheiten (die nicht durch einen, sondern durch das Zusammenspiel mehrere Gendefekte und Umwelteinflüsse verursacht werden) wie beispielsweise Herz-Kreislauf-Erkrankungen bisher große Erfolge eher die Ausnahme. Deshalb achtet man inzwischen verstärkt auf die Güte von Studien beziehungsweise Aussagefähigkeit von Publikationen. Als Grund für die mangelnde Reproduzierbarkeit von Ergebnissen rückt neben der Qualität der Durchführung von biologischen Tests zunehmend das Studiendesign und die damit verbundenen statistischen Auswertungen ins Blickfeld. Viele Publikationsergebnisse zu genetischen Untersuchungen halten einer statistisch validen Überprüfung nicht stand.

■ 2 Studientypen in der genetischen Epidemiologie

Idealtypisch betrachtet, erfolgt die Aufdeckung genetischer - Krankheitsursachen in mehreren Schritten (Abbildung 1). Zunächst wird untersucht, ob eine Erkrankung erblich ist. Daran kann sich eine genauere Analyse der Form der familiären Häufung beziehungsweise Vererbung anschließen. Der nächste wichtige Schritt besteht in Untersuchungen zur Grobkartierung auf

ABBILDUNG 1

Genetische Epidemiologie: Ablaufschema zur Aufdeckung genetischer Ursachen



Quelle: nach Burton, Tobin und Hopper 2005

dem Genom, zum Beispiel mittels Assoziationsstudien. Die letzten Schritte sind der Feinkartierung und der genauen Charakterisierung des genetischen Effekts sowie der Aufdeckung von Wechselwirkungen des Gens mit anderen Loci oder Umwelteinflüssen gewidmet.

2.1 Hinweise auf familiäre Häufung

Bevor zur Suche nach spezifischen Ursachen einer Erkrankung eine genetische Assoziationsstudie mit Markern geplant wird, sollte gesichert sein, dass genetische Faktoren an der Entstehung der Erkrankung beteiligt sind. Ein erster Hinweis dazu ist eine Häufung der Erkrankung in Familien. Eine familiäre Häufung kann, muss aber nicht auf genetischen Ursachen beruhen. Auch Umwelteinflüsse oder Verhaltensmuster, die auf dem Umfeld oder der Erziehung basieren, können beispielsweise eine familiäre Häufung bewirken.

„Familiäre Häufung“ bedeutet, dass spezielle Krankheiten gehäuft in Familien auftreten, dass also eine Person mit höherer Wahrscheinlichkeit erkrankt, wenn bereits ein Angehöriger an der Krankheit leidet oder in der Vergangenheit gelitten hat. Eine wichtige Maßzahl stellt die Rekurrenzzisiko-Ratio λ (Risch 1990) dar, die die Erkrankungswahrscheinlichkeit für eine Person mit einem erkrankten Angehörigen ins Verhältnis zur Prävalenz, also der Häufigkeit einer Krankheit in der Allgemeinbevölkerung setzt. In der Regel betrachtet man bestimmte Verwandtschaftsbeziehungen, zum Beispiel Geschwister. Ist λ deutlich größer als 1, so spricht dies dafür, dass Personen eher erkranken, wenn gleichartige Krankheitsfälle unter den Angehörigen vorliegen.

Zunächst ist bei der Bestimmung der Rekurrenzzisiko-Ratio zu beachten, dass die Möglichkeit zur zufälligen Beobachtung aller erkrankten Personen in einer Population gewährleistet sein muss. Ansonsten kann es durch einen als „Ascertainment Bias“ bezeichneten Fehler zu verzerrten Werten bei Schätzungen kommen. Aber auch die Abhängigkeit des Schätzwerts λ von der Prävalenz, also der Häufigkeit einer Erkrankung in einer Population, ist bei der Interpretation des Rekurrenzzisikos zu beachten. Zu berücksichtigen ist auch, dass die Schätzungen von λ oft nur sehr grob sind. Damit man die Genauigkeit der Schätzung der Rekurrenzzisiko-Ratio erkennen kann, sollte deshalb der Wert mit Konfidenzintervall angegeben werden. Dieser Bereich um den Schätzwert enthält mit einer zuvor festgelegten Wahrscheinlichkeit die wahre (aber unbekannt) Lage der Ratio. Gebräuchlich ist ein 95-Prozent-Konfidenzintervall. Das heißt, dass der tatsächliche Wert im Schnitt in lediglich 5 von 100 Fällen nicht innerhalb der errechneten Intervallgrenzen liegt. So weist eine erhebliche Spannbreite des Konfidenzintervalls auf eine zu geringe Stichprobe hin. Ein erhöhtes erblich bedingtes Krankheitsrisiko für Familienangehörige kann nicht als erwiesen gelten, wenn die 1 im 95-Prozent-Konfidenzintervall enthalten ist.

2.2 Ausschluss von Umwelteinflüssen bei familiärer Krankheitshäufung

Gilt eine familiäre Häufung auf Grundlage der ermittelten Rekurrenzzisiko-Ratio als wahrscheinlich, so bleibt immer noch fraglich, ob dieses Muster tatsächlich konsistent mit genetischen Ursachen ist. Die familiäre Häufung einer Krankheit könnte auch andere Gründe wie beispielsweise Umwelteinflüsse haben. Zum Ausschluss der Umwelteinflüsse als Ursache familiärer Häufigkeiten lassen sich unterschiedliche Verfahren anwenden. Zu nennen sind Zwillingsstudien, Adoptionsstudien sowie Untersuchungen von Familien mittels Varianzkomponenten-Modellen. Auf Zwillingsstudien sowie auf Varianzkomponenten-Modelle wird im Folgenden kurz eingegangen.

2.2.1 Zwillingsstudien

Bei Zwillingsstudien werden die sogenannten Konkordanzen oder Übereinstimmungen von Merkmalen (z.B. das Auftreten von Krankheiten oder stetige Merkmale wie der Body-Mass-Index oder Blutfettwerte) zwischen einerseits dizygoten, das heißt mehreiigen Zwillingen und andererseits monozygoten, also eineiigen Zwillingen bestimmt und anschließend die Konkordanzen beider Zwillingsstypen verglichen. Um einen Einfluss des Geschlechts auf Unterschiede zwischen dizygoten Paaren auszuschließen, dürfen nur gleichgeschlechtliche dizygoten Zwillingspaare in Studien eingeschlossen werden, was bei eineiigen Paaren von Natur aus der Fall ist. Auf der Grundlage dieses Designs ist es möglich herauszufinden, ob genetische Ursachen zu einer familiären Häufung führen oder eher Umwelteinflüsse hierfür verantwortlich sind. Dies basiert auf einer Annahme: Es wird vorausgesetzt, dass bei dizygoten und monozygoten Zwillingen die Übereinstimmungen in Bezug auf Umwelteinflüsse jeweils in etwa gleich sind. Zum anderen ist von Bedeutung, dass monozygote Zwillinge dieselbe Erbinformation tragen, dizygoten Zwillinge aber nur so viel gemeinsame Erbinformation aufweisen wie auch andere Geschwister.

Aus den Unterschieden zwischen den Übereinstimmungen lässt sich eine Schätzung der Erblichkeit ableiten, die als Heritabilität (h^2) bezeichnet wird. Eine höhere Übereinstimmung zwischen monozygoten Zwillingspaaren als zwischen dizygoten Zwillingspaaren (h^2 nahe 1) weist auf genetische Ursachen der Krankheit hin. Eine ähnlich starke Übereinstimmung zwischen monozygoten Paaren wie zwischen dizygoten Paaren (h^2 nahe 0) hingegen deutet auf gemeinsame Umwelteinflüsse als Grund familiärer Häufung hin.

2.2.2 Varianzkomponenten-Modelle

Eine weitere Möglichkeit, die Größe des genetischen Einflusses unter Ausschluss von Umweltfaktoren abzuschätzen, bieten Vari-

anzkomponenten-Modelle. Varianz bezeichnet in der Statistik vereinfacht die Abweichung eines gemessenen Werts vom Mittelwert aller untersuchten Fälle. In Varianzkomponenten-Modellen wird die Varianz in drei Komponenten zerlegt. Es wird untersucht, welche Anteile der Varianz jeweils der gemeinsamen Umweltkomponente, der unabhängigen Umweltkomponente (der interindividuellen Varianz und dem Messfehler) oder aber der genetischen Komponente zugeordnet werden können. Die Heritabilität wird dann aus dem Anteil der genetischen Komponente an der gesamten Varianz geschätzt. Möchte man genauer wissen, welche genetischen Modelle die Beobachtungen erklären, kann man die genetische Varianzkomponente noch weiter unterteilen.

2.3 Genetische Assoziationsstudien

Hat man ausreichende Beweise für eine genetische Komponente einer Erkrankung gesammelt, so versucht man im dritten Schritt diese genetische Komponente oder, bei komplexen Erkrankungen, die genetischen Komponenten auf dem Genom zu lokalisieren. Dazu führt man Assoziationsstudien durch, um eine eventuelle Verknüpfung zwischen dem Phänotyp einer Krankheit und einem oder mehreren genetischen Markern aufzudecken. Bei Markern auf einem Kandidatengen ist entweder aus der Literatur, früheren Untersuchungen oder aus einem Experiment bekannt oder wird vermutet, dass das Gen an einem regulatorischen Mechanismus beteiligt und/oder krankheitsassoziiert ist.

Von Assoziation eines genetischen Markers mit einer Erkrankung spricht man, wenn die relative Häufigkeit des Markers bei Erkrankten gegenüber gesunden Kontrollen unterschiedlich ist.

Genetische Assoziationsstudien ähneln insofern klassischen epidemiologischen Fall-Kontroll-Studien. Es gibt jedoch viele Erweiterungen und Designs, die speziell für die genetische Forschung entwickelt wurden. Heutzutage werden unzählige sogenannte genetische Assoziationsstudien durchgeführt. Verwendete Marker sind die bereits erwähnten Single Nucleotide Polymorphisms (SNPs) sowie sogenannte Mikrosatelliten, kurze sich wiederholende DNA-Sequenzen. Unterschieden werden kann bei solchen Studien zwischen zwei Komplexitätsgraden. Es gibt einfache Assoziationsstudien und darüber hinaus Haplotyp-Analysen, die im Rahmen von Assoziationsstudien durchgeführt werden. Aus Kostengründen haben verschiedene Autoren vorgeschlagen, die DNA der Kontrollen und Fälle jeweils zu poolen, so dass nicht für jedes Individuum eine getrennte Sequenzierung durchgeführt wird. Die Qualität der Resultate ist jedoch vor allem wegen möglicher Verzerrungen umstritten.

2.3.1 Einfache Assoziationsstudien

Im einfachen Fall handelt es sich bei diesen Untersuchungen um Fall-Kontroll-Studien, bei denen zwischen einer Gruppe erkrank-

ter Personen (Fall) und einer weiteren Gruppe, die aus gesunden Personen besteht (Kontrolle), unterschieden wird. Es wird untersucht, ob ein Allel, also die Ausprägung eines untersuchten Markers, in der einen Gruppe häufiger vorkommt als in der anderen.

Wenn man einen additiven Effekt vermutet, also einen Zusammenhang derart, dass homozygote Träger des Risikoallels ein höheres Erkrankungsrisiko haben als heterozygote Träger, so ist die Verwendung von Trendtests sinnvoll. Trendtests untersuchen, ob mit der Anzahl der Risikoallele das Erkrankungsrisiko steigt.

HARDY-WEINBERG-GESETZ

Unter dem Hardy-Weinberg-Gleichgewicht oder Hardy-Weinberg-Equilibrium (HWE) versteht man den Zustand der konstanten Genotypverteilung über verschiedene Generationen hinweg in einer idealen Population (ohne Selektion, Mutation, Migration und unter zufälliger Paarung). Es wird angenommen, dass die Genotyphäufigkeiten einfache Funktionen der Allelfrequenzen sind. Im Fall von zwei Allelen A und B gilt bei Hardy-Weinberg-Gleichgewicht: Wahrscheinlichkeit für den homozygoten Genotyp AA (beziehungsweise BB) ist das Quadrat der Allelfrequenz von A (beziehungsweise B) und die Wahrscheinlichkeit des heterozygoten Genotyps AB ist zweimal das Produkt der Allelfrequenz von A mal der von B:

$$P(AA)=P(A)^2, P(BB)=P(B)^2 \text{ und } P(AB)=2 \cdot P(A) \cdot P(B)$$

Damit gilt:

$$P(AA)+P(AB)+P(BB)=P(A)^2+2 \cdot P(A) \cdot P(B)+P(B)^2=[P(A)+P(B)]^2=1$$

Besonders wenn bestimmte theoretische Voraussetzungen wie beispielsweise das sogenannte Hardy-Weinberg-Gleichgewicht (Hardy-Weinberg-Equilibrium, HWE) nicht erfüllt sind, kann es bei allelbasierten Tests zu Verfälschungen kommen (Sasieni 1997).

Eine gezeigte Assoziation zwischen Marker und Krankheit ist noch kein Beweis für eine kausale Beziehung. Es kann auch durch das sogenannte Linkage Disequilibrium (LD) des Markers mit einem kausalen Polymorphismus zu einer Assoziation kommen. LD beschreibt die Verbindung von Allelen an zwei Genorten. Das heißt, dass das gemeinsame Auftreten von zwei Allelen häufiger ist, als zufällig erwartet werden kann.

Entsprechend sollte sich an den Nachweis einer Assoziation zwischen Krankheit und Gen die sogenannte Feinkartierung anschließen, bei welcher nun der genaue Polymorphismus auf dem Genom lokalisiert wird.

Die rasante Entwicklung der Labortechniken ermöglicht es heute, in einer Assoziationsstudie viele Marker gleichzeitig zu

analysieren. Man untersucht zum Beispiel möglichst viele SNPs in einer ausgewählten Kandidatenregion eines Genoms. Alternativ kann eine Gruppe von Markern aus Kandidatengenomen an verschiedenen Orten im Chromosom untersucht oder auch ein genomweiter Scan durchgeführt werden. Das Ziel eines genomweiten Scans beziehungsweise einer genomweiten Assoziationsstudie ist es, mit Markern in geringen Abständen das gesamte Genom abzudecken, um so systematisch alle möglichen assoziierten Regionen auffinden zu können.

Derzeit werden die Ergebnisse der ersten genomweiten Assoziationsstudien publiziert, zum Beispiel zu Diabetes mellitus Typ 2 von Sladek und Mitarbeitern (*Sladek et al. 2007*).

2.3.2 Haplotyp-Analysen im Rahmen von Assoziationsstudien

Genetische Assoziationsstudien beschränken sich häufig nicht nur auf den einfachen Fall-Kontroll-Vergleich einzelner genetischer Marker. Die Studien werden vielmehr neben der Untersuchung von Umwelteinflüssen auch auf die Untersuchung sogenannter Haplotypen ausgeweitet. Bei Haplotypen handelt es sich um die Anordnung von Allelen aufeinanderfolgender SNPs auf einem Chromosom. In der Regel werden die Allele eines Haplotyps (da sie auf einem Chromosom liegen) miteinander weitervererbt, der Haplotyp bleibt so erhalten. Im Genom gibt es Gebiete, die ein hohes Linkage Disequilibrium aufweisen und in denen die Anzahl möglicher Haplotypen daher gering ist (haplotype blocks), sowie Gebiete, mit geringem Linkage Disequilibrium und entsprechend sehr variantenreichen Haplotypen. In solchen Haplotyp-Blöcken reicht die Kenntnis weniger Marker aus, um den ganzen Haplotyp zu bestimmen.

Durch die Analyse von Haplotypen in Assoziationsstudien verspricht man sich eine höhere statistische Aussagekraft (Power), um eventuelle Assoziationen aufzudecken (*Cardon und Abecasis 2003; Clark 2004*). Zum einen kann es sein, dass ein Haplotyp mehrere Krankheitsloki auf einem Genom umfasst und man so durch Haplotyp-Analyse einen kombinierten Effekt aufdecken kann. Zum anderen könnte der Haplotyp auch einen Krankheitsloкус mit umfassen, der einzeln nicht sequenziert wurde und aufgrund Linkage Disequilibrium zu einem untersuchten Marker mit aufgedeckt wird. Weiterhin kann es sein, dass nicht ein einziger SNP einen Effekt hat, sondern dass der Effekt erst in Kombination mit bestimmten benachbarten Allelen auftritt. Man nimmt zum Beispiel an, dass ein bestimmtes Allel eines SNPs möglicherweise nicht allein krankheitsverursachend ist, sondern nur in Kombination mit einem anderen Allel eines auf dem Chromosom benachbarten SNP, also nur in einem speziellen Haplotypen.

Noch ist es technisch aufwendig, Haplotypen als gesamtes Chromosomenstück zu sequenzieren. Vielmehr werden jeweils nur die einzelnen SNP-Positionen sequenziert. Aus diesen Einzel-

daten muss anschließend rekonstruiert werden, welche Allele gemeinsam auf demselben Chromosom lagen, also welcher Haplotyp vorlag.

Der Entwicklung von Methoden zur Haplotyp-Rekonstruktion und der Haplotyp-Analyse wurde in den letzten Jahren viel Aufmerksamkeit gewidmet. Es gibt generell zwei Vorgehensweisen: Man kann zunächst die Haplotypen schätzen und mit diesen Schätzungen anschließend Assoziationsanalysen durchführen. Andernfalls führt man die Schätzung nicht explizit durch, sondern nimmt stattdessen eine kombinierte Analyse vor.

Eine neuere Methode genetischer Assoziationsuntersuchungen bilden in diesem Zusammenhang die Haplotype-Sharing-Ansätze. Untersucht wird hierbei, ob es Regionen im Genom gibt, bei denen die Haplotypen in der Fallgruppe eine größere Übereinstimmung aufweisen als die Haplotypen in der Kontrollgruppe. In einer solchen Region könnte ein Krankheitsloкус liegen. Angenommen wird, dass eine Mutation, die zu einer Erkrankung führt, zunächst wegen bestehendem Linkage Disequilibrium immer mit ihrer Nachbarregion (also in Form eines bestimmten Haplotypen) weitervererbt wird (*Te Meerman und van der Meulen 1997; McPeck und Strahs 1999; Zhang und Zhao 2000; Qian und Thomas 2001*).

■ 3 Weitere statistische Schwierigkeiten

Generell lässt sich sagen, dass Berichte über Assoziationen häufig überinterpretiert werden (*Cardon und Palmer 2003; Colhoun et al. 2003*). Wie schon angedeutet, lassen sich zahlreiche Berichte über gefundene Gen-Krankheits-Assoziationen nicht bestätigen. Als Gründe werden sogenannte Populationsstratifikationen sowie die Nichtbeachtung des multiplen Testproblems im Zusammenhang mit Publikationsbias genannt (u.a. *Cardon und Bell 2001*).

3.1 Populationsstratifikation oder ethnisches Confounding

Wenn sich Populationen unterschiedlicher ethnischer Herkunft und mit unterschiedlichen Erkrankungsraten mischen, können sich Verzerrungen bei der statistischen Auswertung ergeben. Die ethnisch verschiedenen Populationen werden sich bei einer Fall-Kontroll-Studie nicht gleichmäßig auf Fälle und Kontrollen verteilen. Die Population mit der höheren Prävalenz oder Erkrankungsrate bezogen auf die Gesamtbevölkerung kann in der Fallgruppe überrepräsentiert sein. Dadurch werden im Fall-Kontroll-Vergleich genetische Unterschiede aufgedeckt, die auf der unterschiedlichen Ethnizität beruhen und nicht für die Erkrankung verantwortlich sind. Zum ersten Mal wurde dieses Problem am Beispiel der Pima- und Papago-Indianer erläutert (*Knowler et al. 1988*), die häufiger an Diabetes mellitus Typ 2 leiden als Kau-

KLEINES GLOSSAR DER GENETISCHEN EPIDEMIOLOGIE

Allel = Allele sind alternative Formen eines Gens oder Markers, die an demselben Genlocus auftreten. Bei SNPs (s.u.) sind Allele die jeweils an dem SNP auftretenden Nukleotide. Beispielsweise kann ein Individuum das Allel C auf dem einen und das Allel G auf dem anderen Chromosom tragen, sein Genotyp ist dann heterozygot CG.

Assoziation = Von Assoziation eines genetischen Markers mit einer Erkrankung spricht man, wenn die relative Häufigkeit des Markers bei Erkrankten und Kontrollen unterschiedlich groß ist. Aus einer Assoziation kann jedoch nicht auf eine Kausalität geschlossen werden. Studien zur Suche von Assoziationen sind in der Regel Fall-Kontroll-Studien, wobei speziell in genetischen Studien auch familienbasierte Kontrollen verwendet werden, um ethnische Confounding zu vermeiden.

Genetische Assoziationsstudie = Studie, um eine eventuelle Assoziation zwischen einem Phänotyp (z.B. einer Krankheit) und einem oder mehreren Kandidatengen-Loci (s.u.) oder Markern aufzudecken. Speziell in genomweiten Assoziationsstudien wird nach möglichen Assoziationen zu Markern, die über das gesamte Genom verteilt sind, gesucht, um so Krankheitsgenorte zu entdecken.

Genetischer Marker = Polymorphismus im Genom, dessen Lage exakt bekannt ist. Zum Beispiel sind SNPs genetische Marker.

Haplotyp = Anordnung von Allelen auf einem Chromosom. In der Regel werden die Allele eines Haplotyps (da sie auf einem Chromosom liegen) miteinander weitervererbt, der Haplotyp bleibt so erhalten.

Haplotyp-Block = Im Genom gibt es Gebiete, die ein hohes Linkage Disequilibrium aufweisen und in denen die Anzahl möglicher Haplotypen daher gering ist (haplotype blocks), sowie Gebiete, mit geringem Linkage Disequilibrium und sehr variantenreichen Haplotypen.

Kandidatengen = Gen, von dem vermutet wird oder bekannt ist, dass es an einem regulatorischen Mechanismus beteiligt und/oder krankheitsassoziiert ist. Dieses Vorwissen kann aus der Literatur, früheren Untersuchungen oder aus

einem Screening (zum Beispiel mittels eines Microarray-Experiments) stammen.

Komplexe Erkrankung = Krankheiten mit einer genetischen Komponente, deren Vererbung nicht den Mendel'schen Regeln folgt. Das Erkrankungsrisiko für Verwandte von Kranken ist aber höher als für die Allgemeinbevölkerung. Komplexe Erkrankungen können durch Interaktionen verschiedener Gene oder Gen-Umwelt-Interaktionen beeinflusst werden.

Linkage (auch Koppelung genannt) = Zwei Genorte auf einem Chromosom werden gekoppelt („linked“) genannt, wenn sie aufgrund räumlicher Nähe nicht unabhängig voneinander vererbt werden.

Linkage Disequilibrium = Die überzufällige Assoziation (häufigeres gemeinsames Auftreten als zufällig erwartet) von Allelen an zwei oder mehr Genorten.

Polymorphismus = Wenn an einem Genlocus mehr als ein Allel in der Bevölkerung auftritt, spricht man von einem Polymorphismus. Um zufällig einmalig aufgetretene Abweichungen auszuschließen, stellt man an einen Polymorphismus die Zusatzbedingung, dass mindestens zwei verschiedene Allele mit einer Häufigkeit von jeweils mehr als einem Prozent auftreten. Polymorphismen können überall im Genom auftreten. Es gibt unter anderem kodierende Polymorphismen (wenn die Abweichungen Auswirkungen auf das Genprodukt haben, etwa eine andere Aminosäure entsteht), regulatorische Polymorphismen (wenn der Polymorphismus Auswirkungen auf die Genregulation hat) und stille Polymorphismen (wenn durch die Abweichung keine Änderung auftritt).

SNP = Kurzform für Single Nucleotide Polymorphism. Ein SNP ist eine DNA-Sequenzvariation, die sich auf ein Nukleotid beschränkt. Diese Variation bedeutet, dass an bestimmten Stellen im Genom bei manchen Menschen ein abweichendes Nukleotid vorkommt. So kann bei einem SNP zum Beispiel an Stelle eines Guanins ein Cytosin auftreten. Um als SNP anerkannt zu werden, muss das abweichende Nukleotid bei mindestens einem Prozent der betrachteten Population auftreten.

kasier. Zusätzlich kann es vermehrt zu unbekanntem Verwandtschaftsbeziehungen zwischen den Fällen und somit zu Abhängigkeiten kommen, die berücksichtigt werden müssten. Gerade bei genetisch bedingten Erkrankungen ist die Wahrscheinlichkeit einer Verwandtschaft zwischen erkrankten Personen erhöht.

Populationsstratifikationen werden von einigen Autoren als ein möglicher Grund dafür angesehen, dass sich genetische Assoziationen nicht replizieren lassen (Cardon und Palmer 2003; Thomas und Witte 2002). Andere Autoren argumentieren, dass Populationsstratifikation bei großen Studien mit sorgfältig ausgewählten Fällen und Kontrollen keine großen Auswirkungen

haben (Wacholder et al. 2002). Bei weniger sorgfältig geplanten Studien, bei denen gerade verfügbare Personen (zum Beispiel Laborpersonal) als Kontrollen herangezogen werden, ist solch ein Einfluss jedoch durchaus denkbar, und auch solche Studien werden immer wieder publiziert. Ein Nachweis des Problems ist leider kaum möglich. Ergebnisse solcher Studien sollten deshalb vorsichtig bewertet werden.

Um auf Populationsstratifikation beruhende falsche Assoziationen zu vermeiden beziehungsweise zu korrigieren, wurden verschiedene Strategien für die Analyse der Daten vorgeschlagen. Zunächst ist es möglich, die von den Studienteilnehmern erfragte

ABBILDUNG 2

Überprüfung der statistischen Aussagekraft von genetischen epidemiologischen Studien

- ▼ Ist die Fragestellung der Studie klar definiert ?

- ▼ Sind die verwendeten statistischen Methoden ausreichend beschrieben?
Sind sie angemessen?
Sind die Methoden der Fallzahl angepasst?
– Exakte Tests oder Permutations-p-Werte bei kleinen Fallzahlen? (+)
– In Regressionsmodellen sollte die Fallzahl ein Vielfaches (je nach Art der Einflussgrößen z.B. 10-faches) der Anzahl potenzieller Einflussgrößen sein.

- ▼ Multiples Testen beachtet?
– Signifikanzaussagen entsprechend korrigiert? (+)
– Waren Subgruppenanalysen geplant? (+)
– Erscheinen Ergebnisse selektiert? (-)

- ▼ Confounder berücksichtigt?
– Analyse stratifiziert durchgeführt oder mit Berücksichtigung von Kovariaten?
– Speziell bei genetischen Assoziationsstudien: Mögliche Populationsstratifikation berücksichtigt? ▶ Genomic Control, Stratifizierte Analyse

- ▼ Streuungsmaße oder Konfidenzintervalle angegeben?

- ▼ Bei Heritabilitätsuntersuchungen:
– Keine Überinterpretation der Rekurrenzzisiko-Ratio auf Erblichkeit (Berücksichtigung möglicher geteilter Umwelt)?
– Korrekte Erfassung der Verwandten (+) oder Hinweise auf Ascertainment-Bias (-)?
– Erfassung durch Untersuchung (+) oder nur Befragung (-)?

- ▼ Bei Assoziationsstudien:
– Wurde auf Abweichungen vom Hardy-Weinberg-Equilibrium (HWE) untersucht?
– Korrekte Anzahl von Freiheitsgraden beim Test auf HWE? (korrekt bei SNPs ist 1)
– Exakter Test (oder Permutations-p-Wert) bei kleinen Fallzahlen/seltenen Allelen?
– Keine Überinterpretation der Ergebnisse? ($p > 0,05$ oder $0,1$ bedeutet nicht, dass HWE nachgewiesen ist)

- ▼ Zeigen grafische Darstellungen die Ergebnisse verzerrt (beschönigt)? (-)

- ▼ Werden Ergebnisse beschönigt oder übertrieben? (-)
Werden Ergebnisse kritisch diskutiert? (+)

- ▼ Unterstützen biologische Modelle die Ergebnisse?
Zeigt sich eine Dosis-Wirkungs-Beziehung?

- ▼ Ist das Ergebnis klinisch relevant?
Dies kann nicht anhand des p-Werts, sondern muss anhand der Effektgröße (z.B. Risikomaß) entschieden werden.

- ▼ Ergebnisse validiert?
– intern
– Wurden die Daten in Trainings- und Testdatensatz getrennt; mit Validierung der Ergebnisse am Testdatensatz?
– extern (neue Studie)

Quelle: Victor 2007

ethnische Herkunft als möglichen Confounder (Störvariable) in der Analyse zu berücksichtigen. Dabei stellt sich allerdings das Problem, dass Selbstangaben von Teilnehmern nicht immer korrekt sind. Für eine nachträgliche Korrektur auf der Basis der gesammelten genetischen Daten sind die zwei bekanntesten Methoden die so genannte Genomic Control (Devlin und Roeder

1999; Devlin, Roeder, Bacanu 2001; Devlin, Roeder, Wasserman 2001) und die etwas komplexere Structured Association (Pritchard und Donnelly 2001). Während bei der Genomic Control ein Korrekturfaktor für die Teststatistik eingesetzt wird, erfolgt bei der Structured Association eine Schätzung der Subpopulationszugehörigkeit und eine entsprechend stratifizierte Auswertung.

Eine weitere Möglichkeit, das Problem des ethnischen Confoundings anzugehen, sind spezielle Studiendesigns mit einer Rekrutierung der Kontrollen aus der Familie, also Studien mit familienbasierter Kontrolle. Beispiele sind die Rekrutierung von Eltern-Kind-Trios oder von diskordanten Geschwistern. Diskordant bedeutet zum Beispiel, dass ein Geschwister krank, das andere gesund ist, oder dass ein Geschwister einen extrem hohen Body-Mass-Index, das andere einen normalen Wert aufweist. Mit dem sogenannten Transmission-Disequilibrium-Test (TDT) (Spielman *et al.* 1993; Ewens und Spielman 1995) wird untersucht, inwieweit die Häufigkeit der an das kranke Kind vererbten Allele heterozygoter Eltern von der bei Unabhängigkeit von der Krankheit erwarteten Häufigkeit von 0,5 abweicht. Ist ein Elternteil heterozygot, so liegt die Wahrscheinlichkeit, dass er eines der beiden Allele vererbt, im Idealfall jeweils bei 0,5. Der TDT untersucht, ob ein Allel davon abweichend gehäuft an Erkrankte vererbt wird. Im klassischen Fall werden dafür zusätzlich zu erkrankten Personen beide Elternteile untersucht, um festzustellen, welche elterlichen Allele vererbt beziehungsweise nicht vererbt wurden. Sind beide Elternteile verfügbar, so ist klar bestimmbar, wer welches Allel vererbt hat. Wenn mindestens ein Elternteil nicht verfügbar ist, kann man bei der Typisierung von Geschwistern in manchen genetischen Konstellationen trotzdem rekonstruieren, welchen Genotyp der fehlende Elternteil hatte. Da diese Rekonstruktion jedoch von der genetischen Konstellation abhängt, würde das Ergebnis des einfachen TDT-Tests verzerrt. Es ist daher nötig, eine korrigierende Version anzuwenden (Knapp 1999).

3.2 Multiples Testen in genetischen Studien

Ein weiteres generelles Problem genetischer Studien – unabhängig vom Design – ist das multiple Testen. Bei genetischen Studien werden sehr viele Hypothesen (über Marker, Interaktionen oder Subgruppen) an einem Datensatz (häufig von wenigen Individuen) untersucht. Hypothesen können zum Beispiel sein (zunächst nur auf einen Marker bezogen):

- Es besteht eine Assoziation zwischen der Krankheit und dem Marker 1.
- Es besteht eine Interaktion der Marker 1 und 2 bezüglich des Erkrankungsrisikos.
- Es besteht eine Assoziation zwischen der Krankheit und dem Marker 1 in der Subgruppe der Frauen.

Die Anzahl der untersuchten Marker reicht von unter 100 für die Untersuchung einiger Kandidatengen-Loki über größere Assoziationsstudien mit mehr als 10.000 untersuchten Markern bis in den Bereich der genomweiten Assoziationsstudien mit um die 500.000 untersuchten Markern. Die Anzahl der untersuchten Hypothesen und damit der durchgeführten statistischen Tests ist

somit sehr groß. Bei der statistischen Auswertung solcher genetischen Studien ist man daher immer wieder mit dem Problem des multiplen Testens konfrontiert. Die Wahrscheinlichkeit, mindestens eine dieser Hypothesen fälschlich anzunehmen (also eine falsch positive Aussage zu machen), steigt mit der Anzahl der untersuchten Hypothesen. Je höher die Anzahl der durchgeführten Tests, desto mehr rein zufällige Auffälligkeiten kann man beobachten, die zu falsch positiven Entscheidungen führen. Das bedeutet, dass die Gefahr von falsch positiven Aussagen in der Genetik hoch ist. Da man nicht weiß, welche Ergebnisse zufällig falsch positiv und welche tatsächlich richtig sind, haben die Resultate von Studien mit vielen Tests (bei denen das multiple Testproblem nicht berücksichtigt wurde) wenig Aussagekraft. Eine striktere Definition von „signifikant“ wäre nötig, oder die Aussagen müssten entsprechend relativiert werden.

Leider nennen viele Autoren von Publikationen nicht die genaue Anzahl untersuchter Marker, Subgruppen usw., so dass die Beurteilung der Qualität der Ergebnisse bezüglich multiplem Testen schwierig ist. Häufig kommt man jedoch schon bei Berücksichtigung nur der Anzahl der in der Publikation genannten statistischen Tests zu einer vorsichtigeren Interpretation der Ergebnisse als die Autoren.

4 Schlusswort

Die statistischen Designs und Verfahren der genetischen Epidemiologie sind sehr komplex. Deshalb sind auch Ex-post-Bewertungen zu der Qualität von Untersuchungen und damit der Validität der Ergebnisse nur schwer durchführbar. Immer wieder kommt es vor, dass mithilfe statistischer Überprüfungen Schlüsse aus genetischen Untersuchungen gezogen werden, die sehr viel schwächer oder gar anders ausfallen als in den entsprechenden Publikationen zunächst dargestellt. Es ist daher beim Studium genetischer Publikationen dringend angeraten, die Ergebnisse kritisch zu hinterfragen; zumindest in Bezug darauf, ob das Studiendesign und die verwendeten statistischen Methoden angemessen sind und die Schlussfolgerungen gerechtfertigt sind. Im schlimmsten Fall glaubt der Leser lange Zeit einer Falschmeinung. Für Forscher ist die korrekte Anwendung der verschiedenen Methoden und die korrekte Interpretation der Ergebnisse dringend zu empfehlen. Ein Biometriker sollte in das Forschungsprojekt einbezogen werden. Editoren und Reviewer wissenschaftlicher Journals sollten Beiträge einem qualifizierten statistischen Review unterziehen und vor allem auch die Publikation von Negativergebnissen zulassen, damit falsch positive Ergebnisse relativiert werden. ♦

Literatur

- Burton PR, Tobin MD, Hopper JL (2005):** Key concepts in genetic epidemiology. *The Lancet*, Band 366, 941–951
- Cardon LR, Bell JI (2001):** Association study designs for complex diseases. *Nature Reviews Genetics*, Band 2, 91–99
- Cardon LR, Abecasis GR (2003):** Using haplotype blocks to map human complex trait loci. *Trends in Genetics*, Band 19, 135–140
- Cardon LR, Palmer LJ (2003):** Population stratification and spurious allelic association. *The Lancet*, Heft 361, 598–604
- Clark AG (2004):** The role of haplotypes in candidate gene studies. *Genetic Epidemiology*, Band 27, 321–333
- Colhoun HM, McKeigue PM, Smith GD (2003):** Problems of reporting genetic association with complex outcomes. *The Lancet*, Band 361, 865–872
- Devlin B, Roeder K (1999):** Genomic control for association studies. *Biometrics*, Band 55, 997–1004
- Devlin B, Roeder K, Wasserman L (2001):** Genomic control, a new approach to genetic-based association studies. *Theoretical Population Biology*, Band 60, 155–166
- Devlin B, Roeder K, Bacanu SA (2001):** Unbiased methods for population-based association studies. *Genetic Epidemiology*, Band 21, 273–284
- Ewens WJ, Spielman RS (1995):** Transmission/disequilibrium test: history, subdivision and admixture. *American Journal of Human Genetics*, Band 57, 455–464
- Hirschhorn JN, Lohmüller K, Byrne E, Hirschhorn K (2002):** A comprehensive review of genetic association studies. *Genetics in Medicine*, Band 4, 45–61
- Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG (2001):** Replication validity of genetic association studies. *Nature Genetics*, Band 29, 306–309
- Ioannidis JP (2003):** Genetic associations: false or true? *Trends in Molecular Medicine*, Band 9, 135–138
- Knapp M (1999):** The transmission/disequilibrium test and parental-genotype reconstruction: The reconstruction-combined transmission/disequilibrium test. *American Journal of Human Genetics*, Band 64, 861–870
- Knowler WC, Williams RC, Pettitt DJ, Steinberg AG (1988):** Gm3;5;13;14 and type 2 diabetes mellitus: An association in American Indians with genetic admixture. *American Journal of Human Genetics*, Band 43, 520–526
- Lohmüller KE, Pearce CL, Pike M, et al. (2003):** Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nature Genetics*, Band 33, 177–182
- McPeck MS, Strahs A (1999):** Assessment of linkage disequilibrium by the decay of haplotype sharing, with application to fine-scale genetic mapping. *American Journal of Human Genetics*, Band 65, 858–875
- Pritchard JK, Donnelly P (2001):** Case-control studies of association in structured or admixed populations. *Theoretical Population Biology*, Band 60, 227–237
- Qian D, Thomas DC (2001):** Genome scan of complex traits by haplotype sharing correlation. *Genetic Epidemiology*, Band 21 (Suppl. 1), 582–587
- Risch N (1990):** Linkage strategies for genetically complex traits. I. Multilocus models. *American Journal of Human Genetics*, Band 46, 222–228
- Sasieni PD (1997):** From genotypes to genes: Doubling the sample size. *Biometrics*, Band 53, 1253–1261
- Sladek R, Rocheleau G, Rung J et al. (2007):** A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*, Band 454, 881–885
- Spielman RS, McGinnis RE, Ewens WJ (1993):** Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *American Journal of Human Genetics*, Band 52, 506–516
- Te Meerman GJ, van der Meulen MA (1997):** Genomic sharing surrounding alleles identical by descent: effects of genetic drift and population growth. *Genetic Epidemiology*, Band 14, 1125–1130
- Thomas DC, Witte JS (2002):** Point: Population stratification: A problem for case-control studies of candidate-gene associations. *Cancer Epidemiology Biomarkers and Prevention*, Band 11, 505–512
- Wacholder S, Rothman N, Caprosa N (2002):** Counterpoint: Bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. *Cancer Epidemiology Biomarkers and Prevention*, Band 11, 513–520
- Zhang S, Zhao H (2000):** Linkage disequilibrium mapping in populations of variable size using the decay of haplotype sharing and stepwise-mutation model. *Genetic Epidemiology*, Band 19 (Suppl. 1), 99–105

DIE AUTORIN

**Dr. rer. physiol. Anja Victor**

Von 1994 bis 1999 Studium der Mathematik mit Nebenfach Biologie an der Universität Marburg. Seit 2000 wissenschaftliche Mitarbeiterin am Institut für medizinische Biometrie, Epidemiologie und Informatik der Universität Mainz mit

Schwerpunkt Statistik in der Genetik tätig. 2005 Promotion an der Universität Mainz zum Thema „Kontrolle der False Discovery Rate in adaptiven Designs“.