

Assoziationsstudien bei Familiendaten in der Genetischen Statistik

Hans Jörg Baurecht
Institut für Medizinische Statistik
und Epidemiologie / TU-München
Ismaninger Str. 22
81675 München
hansjoerg.baurecht@tum.de

Zusammenfassung

Familiendaten zu Vater, Mutter und Kinder finden im Rahmen von Assoziationsstudien zur Untersuchung von Genotyp- Phänotyp-Beziehungen in der genetischen Statistik und Epidemiologie breite Anwendung. Für die biostatistische Analyse ist der sogenannte Transmission-Disequilibrium-Test (TDT) (1) und die daraus hervorgegangene Klasse von Verfahren wie beispielsweise der S-TDT, der die Geschwisterinformation mit einbezieht, von essentieller Bedeutung. Der TDT ist attraktiv aufgrund seiner Einfachheit und Robustheit gegenüber Scheinassoziation, die bei Heterogenität in der Population auftreten kann (4).

Schlüsselwörter: Transmission-Disequilibrium-Test, TDT, Genetik, Trios, Familien, Assoziationsanalyse

Einleitung

Beim TDT werden Kernfamilien (Trios), das sind Eltern mit erkranktem Kind, analysiert. Der beobachtete Phänotyp ist die Erkrankung (ja/nein) des Kindes. Der Genotyp ist bei unseren Analysen bestimmt durch die vor der Untersuchung festgelegten SNPs. SNP ist eine Abkürzung für Single Nucleotid Polymorphismus und bezeichnet Stellen im Gen, bei denen ein Austausch von Basenpaaren stattfinden kann. Die alternativen Ausprägungen an einem SNP nennt man Allele.

1 Transmission-Disequilibrium-Test (TDT)

1.1 Theoretischer Hintergrund

Mit Hilfe des TDT wird untersucht, welches Allel häufiger von einem heterozygoten Elternteil an ein erkranktes Kind übertragen wird. Dieser Sachverhalt lässt sich leicht mit Hilfe einer Vierfeldertafel aufstellen. Damit wird eine Assoziationsstudie mit internen Kontrollen und daher höchstmöglicher Homogenität durchgeführt, weil die beiden nicht auf das erkrankte Kind übertragenen elterlichen Allele automatisch als interne Kontrollgruppe fungieren (2). Der TDT entspricht dem aus der Statistik bekannten McNemar-Test für dichotome Merkmale in verbundenen Stichproben.

Ein wesentlicher Vorteil des TDT gegenüber klassischen Fall-Kontroll-Studien ist seine Robustheit gegenüber Stratifizierungen innerhalb der untersuchten Population. Es werden nämlich falsch-positive Resultate aus Fall-Kontroll-Studien, die auf eine nicht spezifizierte Populationsstratifizierung zurückzuführen sind, durch die Verwendung interner Kontrollen vermieden.

1.2 Anforderungen und Probleme in der Praxis

Für Publikationen in der genetischen Epidemiologie werden zusätzlich zum Testergebnis die Übertragungshäufigkeiten des assoziierten und nicht assoziierten Allels angegeben, welche aus der Vierfeldertafel des zugrundeliegenden McNemar-Tests hervorgehen. Dies entspricht dem Verhältnis der diskordanten Paare in der Tabelle.

Leider wird diese Information von PROC FAMILY im Modul SAS/Genetics (3) nicht bereitgestellt. Es gibt alternative, freie Software im Internet, die dieses leistet. Jedoch ist das zugrundeliegende Testverfahren oft unzulänglich beschrieben. Es wird nicht angegeben, ob für den Chi-Quadrat-Test eine Stetigkeitskorrektur vorgenommen wird oder ob das exakte Verfahren verwendet wird. Zudem werden oft nur vollständig genotypisierte Trios verwendet, wenn mehrere SNPs untersucht werden.

1.3 Individuelle Lösung mit SAS

In unserer Arbeitsgruppe haben wir mit SAS die zugehörige Vierfeldertafel selbst programmiert, um die Übertragungshäufigkeiten angeben zu können. Dabei wird jeder SNP separat ausgewertet und nicht in einem vorhergehenden Schritt die vollständigen Familien selektiert, wie dies z.B. von Haploview (9) gemacht wird. Schließlich lassen sich über die asymptotische und exakte Variante des McNemar-Tests die p-Werte für den TDT berechnen. Des weiteren lassen sich Odds Ratios und die dazugehörigen Konfidenzintervalle angeben, wie sie in (7) beschrieben sind. Zudem lässt sich die beobachtete Power und das attributable Risiko (8) angeben. Schließlich erhält man mit Hilfe des attributablen Risikos und der Prävalenz des Expositionsallels das populationsbasierte attributable Risiko (7,8).

Anhand der Untersuchung des Filaggrin-Gens (10) wird gezeigt, wie die Resultate aufbereitet werden können.

Grenzen des TDT und Ausblick

Ein Nachteil des klassischen TDT ist, dass weitere Kovariablen wie im Regressionskontext nicht berücksichtigt werden können. Bei qualitativen Kovariablen können verschiedene Strata betrachtet werden, jedoch lässt sich kein Effekt schätzen. Dazu sei auf die Arbeiten von Lange und Laird (5,6) verwiesen, die eine generelle Klasse von familienbasierten Assoziationstests (FBAT) entwickelten.

Literatur

- [1] Spielman RS, McGinnis RE, Ewens JW. Transmission test for linkage disequilibrium: the insulin gene region and insulin-independent diabetes mellitus (IDDM). *Am. J. Hum. Genet* 1993;52:506-516
- [2] Falk CT, Rubinstein P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann. Hum. Genet* 1987;51:227-233
- [3] SAS/Genetics User's Guide
- [4] Lange C, DeMeo DL, Laird NM . Power and design considerations for a general class of family-based association tests: quantitative traits. *Am J Hum Genet.* 2002 Dec;71(6):1330-41
- [5] Lange C, Laird NM. On a general class of conditional tests for family-based association studies in genetics: the asymptotic distribution, the conditional power, and optimality considerations. *Genet Epidemiol.* 2002 Aug;23(2):165-80.
- [6] Lange C, Laird NM. Power calculations for a general class of family-based association tests: dichotomous traits. *Am J Hum Genet.* 2002 Sep;71(3):575-84.
- [7] Morris JA, Gardner MJ, Epidemiological Studies. In: Altman DG, Machin D, Bryant TN ed., "Statistics with Confidence", volume 2, BMJ book, Bristol 2005, pp.57-72
- [8] Kreienbrock L, Schach S, Epidemiologische Methoden, volume 4, Elsevier GmbH, München 2005
- [9] Broad Institute of MIT and Harvard 2003-2006, www.broad.mit.edu/mpg/haploview/
- [10] Weidinger, S et al. (2006), 'Loss-of-function variations within the filaggrin gene predispose for atopic dermatitis with allergic sensitizations.', *J Allergy Clin Immunol* 118(1), 214-219.