

Schätzung des Genotyp-und Haplotypfehlers und deren Einfluss auf die Haplotypassoziationsanalyse

Claudia Lamina

Institut für Epidemiologie

Helmholtz Zentrum München

German Research Center for Environmental Health (GmbH)

Haplotypen stellen eine bestimmte Form von genetischen Varianten dar, bestehend aus den Kombinationen mehrerer Allele auf einem Chromosomstrang, und spielen eine große Rolle in genetischen Assoziationsstudien. Da Methoden zur Haplotypisierung in großen epidemiologischen Studien nicht effizient durchführbar sind, werden sie mit statistischen Methoden aus den Genotypen geschätzt. Das geschieht mittels Bayesianischen Methoden oder dem EM-Algorithmus.

Durch die Schätzung aus Genotypen kann es einerseits durch Genotypisierungsfehler im Labor als auch durch Fehler in der statistischen Haplotyprekonstruktion zu fehlerhaften Haplotypen kommen, was wiederum zu einer verzerrten Schätzung in der Assoziationsanalyse führen kann.

Der Genotypisierungsfehler wurde mittels Maximum Likelihood Verfahren aus Wiederholungsmessungen geschätzt. Zur Quantifizierung des Haplotypfehlers wurden Simulationsstudien durchgeführt und Fehlermasse wie Sensitivität, Spezifität und Fehlklassifikationsmatrix geschätzt, als auch deren Einfluss auf genetische Assoziationsanalysen. Als Beispieldatensatz diente eine Haplotypassoziationsanalyse des *APM1*-Gens zu Adiponectin-Plasma-Werten. Hier zeigte sich, dass der Haplotypfehler vernachlässigbar ist, wenn kein oder nur ein geringer Genotypisierungsfehler angenommen werden kann. Ist dies nicht der Fall, kann die Haplotypmisklassifikation approximiert werden, um anschließend dafür in der Assoziationsanalyse mit geeigneten statistischen Methoden zu korrigieren. Für diese Misklassifikationskorrektur bei der Haplotypassoziationsanalyse wurde die MC-SIMEX-Methode (Misclassification-SIMEX) verwendet.