

Vergleich von False Discovery Rate (FDR) und Familywise Error Rate (FWER) Kontrolle bei genetischen Fragestellungen unter Berücksichtigung verschiedener Prozeduren

Victor A, Hommel G

Institut für Medizinische Biometrie, Epidemiologie und Informatik, Universität Mainz, Deutschland
victor@imsd.uni-mainz.de

Einleitung und Fragestellung Die Kontrolle der False Discovery Rate (FDR) ist ein im Augenblick sehr populärer Ansatz bei explorativen Fragestellungen mit zahlreichen Endpunkten. Viele modifizierte Prozeduren zur Kontrolle der FDR sind bereits und werden derzeit noch vorgeschlagen, besonders im Zusammenhang mit Microarray-Untersuchungen. Für die Anwendung im Microarray-Bereich sind erste Vergleiche neuerer Prozeduren zur „klassischen“ explorativen Simes Prozedur, welche von Benjamini & Hochberg [1] für die FDR-Kontrolle vorgeschlagen wurde, publiziert worden [2,3,4]. In Zusammenhang mit genetischen Assoziationsstudien mit Single Nucleotide Polymorphisms sind Untersuchungen zu Vorteilen durch Kontrolle der FDR anstatt der FWER und Vergleiche verschiedener FDR-Kontrollprozeduren jedoch selten. Bei der Auswertung von Assoziationsstudien mit SNPs werden andere Teststatistiken als in Microarray-Untersuchungen verwendet. Der Anteil der wahren Nullhypothesen, welcher einen Einfluss auf die Unterschiede zwischen FDR und FWER hat, hängt von der betrachteten Situation ab, und kann bei genomweiten Assoziationsscans sehr klein sein. Es stellt sich daher die Frage, welche Vorteile die Kontrolle der FDR in solchen Studien bietet, und ob und in welcher Form sich verschiedene FDR-Kontrollprozeduren unterscheiden.

Material und Methoden Es werden Simulationen in Anlehnung an genetische Assoziationsstudien mit SNPs durchgeführt. Im Rahmen dieser Simulationen werden verschiedene einfache FWER-Kontrollprozeduren mit der klassischen explorativen Simes Prozedur und neuerer FDR-Kontrollprozeduren und diese FDR-Kontrollprozeduren untereinander verglichen (u.a. [5,6]). In Zusammenhang mit FDR-Kontrolle bei diskret verteilten Teststatistiken, wie sie zum Beispiel bei Verwendung des exakten Fisher Tests auftreten, soll auch der Ansatz von Gilbert [7] berücksichtigt werden.

Ergebnisse Wie zu erwarten, zeigt sich eine Abhängigkeit der Ergebnisse von der jeweils simulierten Situation. Dies betrifft sowohl die Unterschiede zwischen den einzelnen Testprozeduren sowie die Unterschiede zwischen den FDR-Kontrollprozeduren und der FWER-Kontrolle mit einfachen Prozeduren. Deutlichen Einfluss zeigt der Anteil der wahren Nullhypothesen. Betrachtet man eher eine Scansituation mit einem sehr hohen Anteil wahrer Nullhypothesen ($>95\%$), so ist zum einen nur eine gering höhere Ablehnhäufigkeit der FDR-kontrollierenden Prozeduren im Vergleich zur Bonferroni-Prozedur zu erkennen. Zum anderen sind auch bei Verwendung neuerer FDR-Kontrollprozeduren, deren Verbesserung auf einer Korrektur für den Anteil wahrer Nullhypothesen beruhen, kaum Unterschiede zur explorativen Simes-Prozedur zu erkennen. Bei Simulation von Situationen, bei denen man relativ wenige Kandidatengenloki untersucht und der Anteil wahrer Nullhypothesen nicht so hoch ist (zum Beispiel eine optimistische Situation mit 60% wahren Nullhypothesen), treten deutlichere Unterschiede zur Bonferroni-Prozedur auf, und leichtere zwischen den FDR-Kontrollprozeduren.

Diskussion Es wird in explorativen Untersuchungen derzeit erwartet, dass deutlich mehr Ablehnungen möglich sind, wenn man die FDR anstatt der FWER kontrolliert. Es muss jedoch berücksichtigt werden, dass das Verhalten der FDR-Kontrollprozeduren stark von der Studien-Situation abhängt.

Literatur

- [1] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *JRSS B* 1995; 57: 289-300.
- [2] Li SS, Bigler J, Lampe JW, Potter JD, Feng Z. FDR-controlling testing procedures and sample size determination for microarrays. *Statistics in medicine* 2005; 24: 2267-2280.
- [3] Qian HR, Huang S. Comparison of false discovery rate methods in identifying genes with differential expression. *Genomics* 2005; 86: 495-503.
- [4] Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* 2003; 19: 368-375.
- [5] Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Annals of statistics* 2001; 29: 1165-1188.
- [6] Benjamini Y, Krieger AM, Yekutieli D. Adaptive linear step up procedures that control the FDR. Department of Statistics and OR. Tel Aviv University, Tel Aviv, 2004
- [7] Gilbert PB. A modified false discovery rate multiple-comparisons procedure for discrete data, applied to human immunodeficiency virus genetics. *Applied Statistics* 2005; 44:143-158