

Vergleich von FDR und FWER Kontrolle bei genetischen Fragestellungen unter Berücksichtigung verschiedener Prozeduren

A. Victor und G. Hommel
IMBEI Mainz



Motivation

- ❖ Multiples Testen in genetischen Studien als Problem erkannt
- ❖ FWER (Familywise Error Rate) oft zu strikt
- ❖ FDR (False Discovery Rate) oft als "Allheilmittel" gepriesen
- ❖ Aber:
 - Wenn alle Hypothesen wahr sind, gilt $FDR = FWER$
 - Unterschied zwischen FDR und FWER Rate steigt mit steigendem Anteil falscher Hypothesen zunächst an

Motivation

- ❖ Bei genetischen Assoziationsstudien (auf Assoziation zwischen genetischen Markern und einer Erkrankung) können auftreten:
 - Nicht stetige Teststatistiken
 - Abhängigkeiten zwischen untersuchten Markern
 - Situationen mit höherem Anteil an wahren Alternativen (z.B. Untersuchung von Kandidatengenomen)
 - Situationen mit niedrigem Anteil wahrer Alternativen ('Scan')

Ziele

- ❖ Wie verhalten sich 'klassische' FWER und FDR Prozeduren und solche mit Beachtung der Diskretheit der Teststatistiken bei
 - Nicht stetigen Teststatistiken
 - Abhängigen Markern
 - Unterschiedlichem Anteil wahrer Hypothesen

Prozeduren im Vergleich

- ❖ Zur Kontrolle der FWER:
 - Bonferroni
 - Bonferroni-Holm
 - Tarone
 - Verbesserung Bonferroni bei diskreten Daten
 - Hommel & Krummenauer
 - Verbesserung der Tarone Prozedur

Prozeduren im Vergleich

- ❖ Zur Kontrolle der FDR
 - Explorative Simes Prozedur
 - “Benjamini-Hochberg”
 - Ein “adaptierter” Benjamini Hochberg
 - Benjamini-Yekutieli
 - Strafterm für Abhängigkeit
 - Tarone für FDR nach Gilbert (2005)

Prozeduren für diskrete Daten

❖ Tarone (1990)

- p-Werte bei diskreten Daten nicht gleichverteilt (kleinstes erreichbares Niveau)
- Auf eine Untermenge der Hypothesen beschränkte Korrektur
- Prozedur nicht monoton im Signifikanzniveau

❖ Hommel & Krummenauer (1998)

- Verbesserung zum Erreichen der Monotonie
- Führt damit auch zu mächtigerer Prozedur

❖ Gilbert (2005)

- Übertragung von Tarone auf FDR
- Anwendung der explorativen Simes Prozedur auf mittels Tarone ermittelte Untermenge

Adaptierter Benjamini-Hochberg

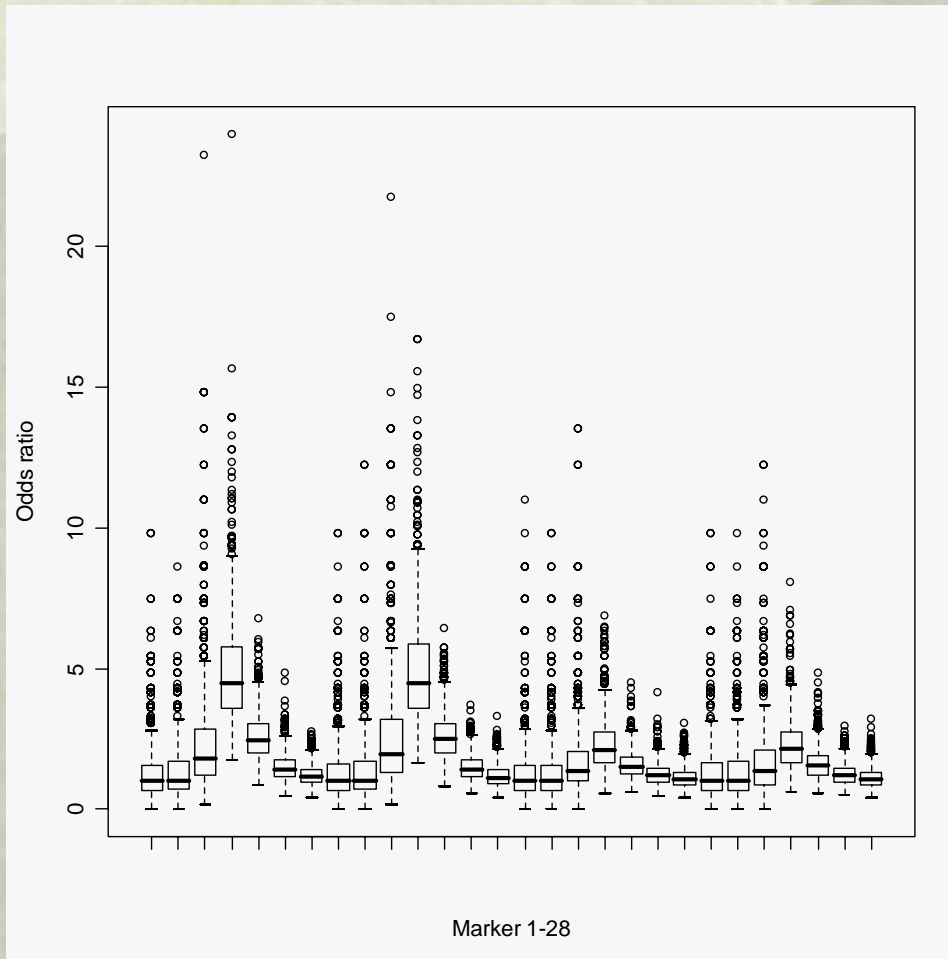
- ❖ Explorative Simes Prozedur kontrolliert die FDR (im Falle PRDS) zum Niveau m_0/m
- ❖ Zahlreiche Verbesserungsvorschläge für diese Prozedur zielen auf Schätzung m_0 und anschließende Korrektur durch Verwendung des Niveaus $m/m_0 * \alpha$ ab (vgl. z.B. Vortrag Walther; Hemmelmann und Vollandt)
- ❖ Hier „Adaptierter Benjamini-Hochberg“ aus (Benjamini & Hochberg 2000)

Simulationen Typ 1

- ❖ Komplett simulierte Daten
 - dichotome Marker, dichotomes Outcome
 - In Blocks zu je 7 Markern
 - Mittlerer Marker in Block zufällig simuliert
 - Benachbarte Marker mit Wahrscheinlichkeiten bedingt auf den Nachbarmarker simuliert
 - d.h. wenn Nachbarmarker=1, hat der benachbarte Marker eine 6- bzw. 3 fach höhere Wahrscheinlichkeit auch 1 zu sein
 - 3 Marker links des mittleren Markers mit geringerer Häufigkeit des abweichenden Allels
 - 3 Marker rechts mit höherer Häufigkeit
 - Falls Assoziation, so jeweils zu mittlerem Marker simuliert

Simulationen Typ 1

Beispiel sich ergebendes OR (Auszug)



Simulationen Typ 2

- ❖ Reale SNP-Daten mit simuliertem Outcome
 - Verschiedene Modelle bestimmen die Wahrscheinlichkeit für das Outcome (D) unter bestimmter Genotypkonstellation

$P(D \mid \text{Genotypkonstellation})$

$$= \frac{\exp(-2 + \vec{\text{coeff}} \cdot \text{Genotypkonstellation})}{1 + \exp(-2 + \vec{\text{coeff}} \cdot \text{Genotypkonstellation})}$$

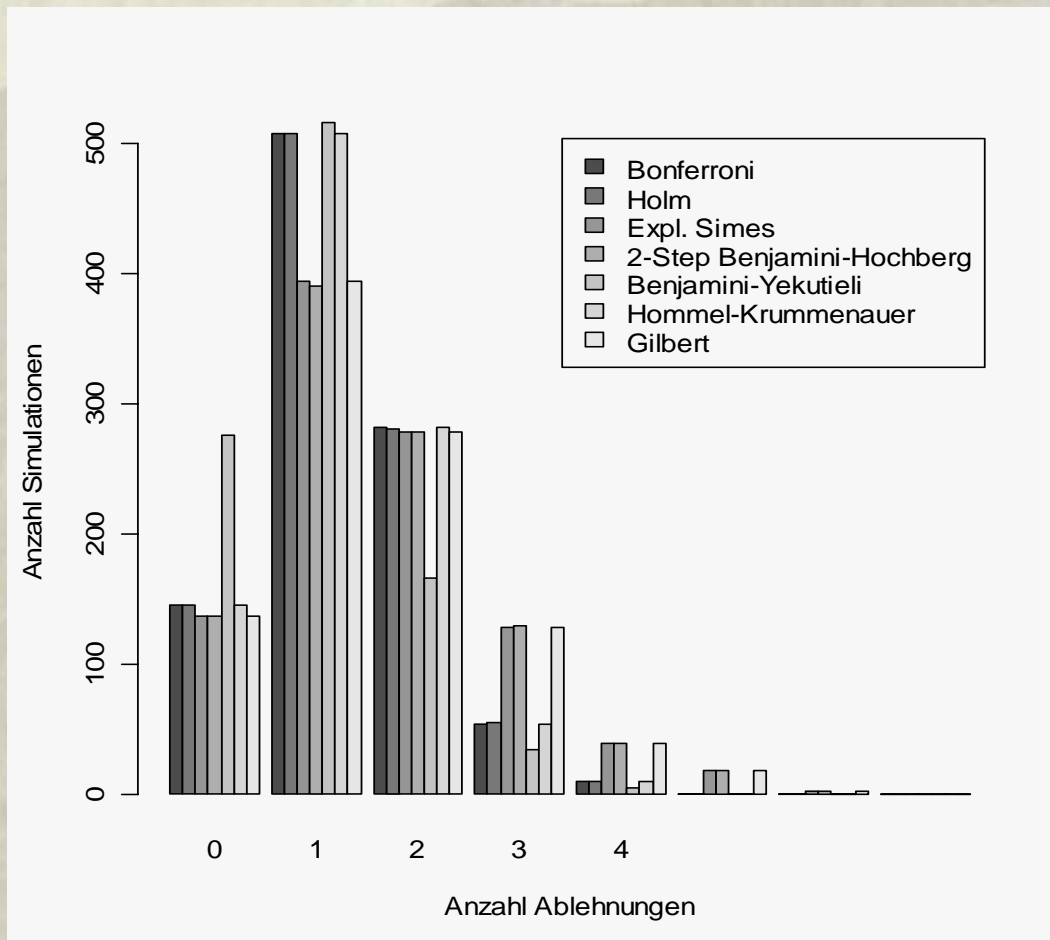
- Outcome für jedes Individuum binomial simuliert mit Wahrscheinlichkeit jeweils nach individueller Genotypkonstellation und gewähltem Modell

Reale SNP Daten

- ❖ Studie aus I.Med (bereits etwas älter)
- ❖ Ursprünglich 122 Individuen
 - verdoppelt für Simulationen
- ❖ SNPs aus der V- β -Region des T-Zell-Rezeptors
 - 26 davon für Simulationen ausgewählt
- ❖ Daten nicht perfekt
 - z.T. kaum polymorph
 - Manche SNPs zeigen Abweichung von HWE
 - Fehlende Werte
- ❖ 2 "Haplotypblocks" mit dazwischenliegenden "unabhängigeren" SNPs
 - SNPs 1-15 (Lewontin's D' zwischen 0.3-0.9)
 - SNPs 18-22 (Lewontin's D' zwischen 0.3-0.8)

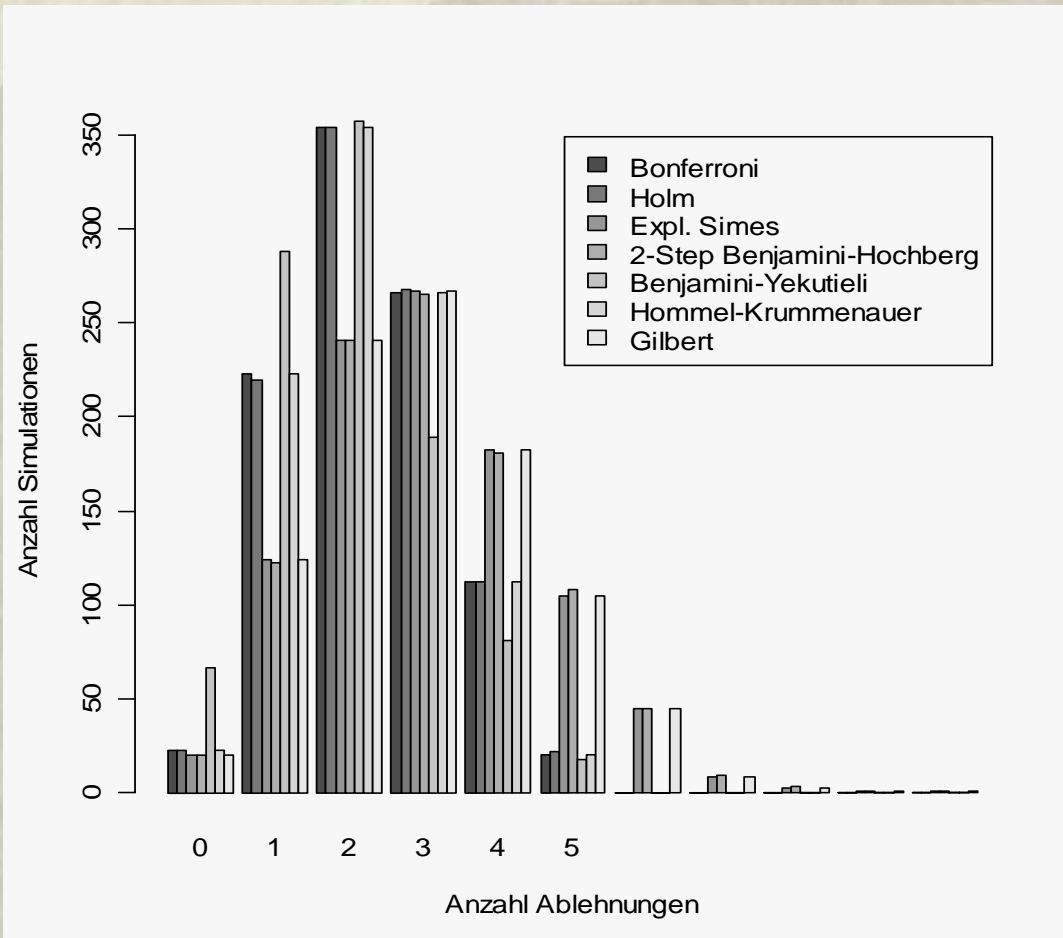
Typ 1 – Situation 1

30 Blocks=210 Marker, 200 Fälle und 200 Kontrollen,
geringere Häufigkeit=5%, höhere Häufigkeit=20%,
1 Block RR=3, 2 Blöcke RR=2 (RR für zentralen Marker)



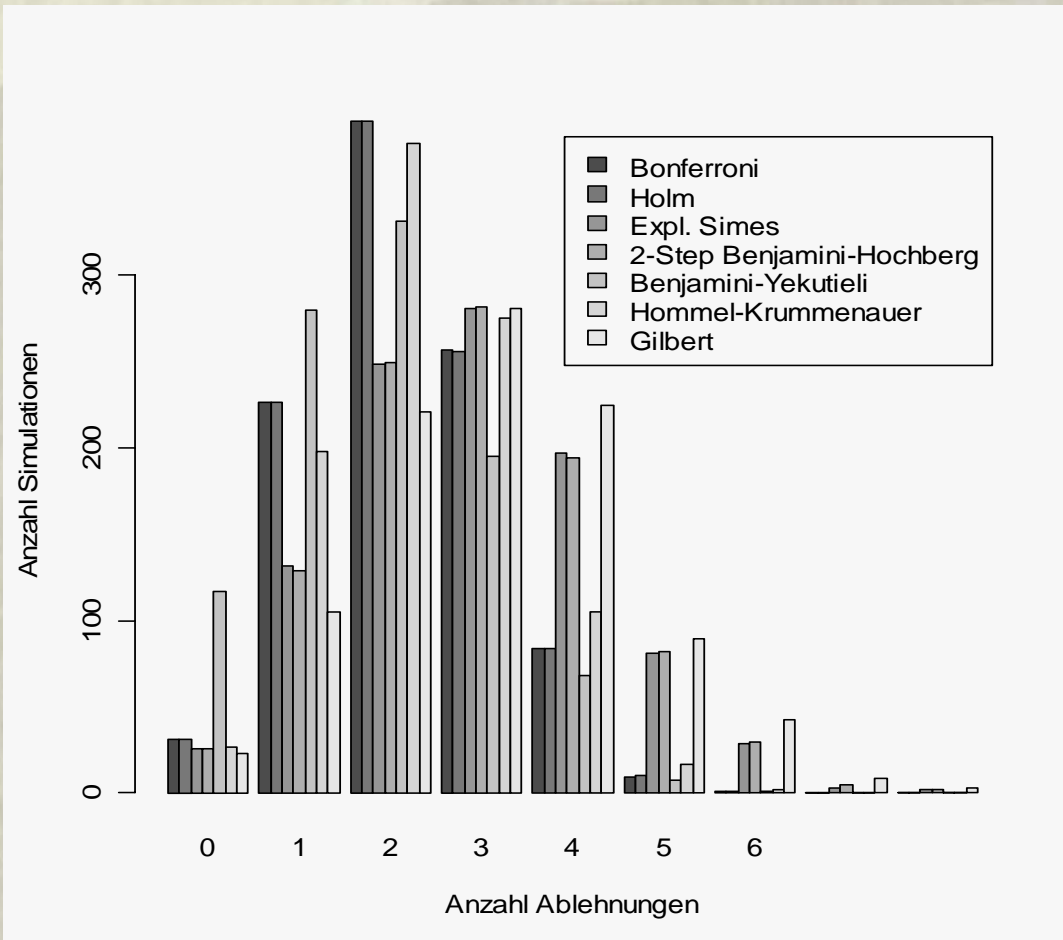
Typ 1 – Situation 2

30 Blocks=210 Marker, 200 Fälle und 200 Kontrollen
geringere Häufigkeit=10%, höhere Häufigkeit=25%,
1 Block RR=3, 2 Blöcke RR=2 (RR für zentralen Marker)



Typ 1 – Situation 3

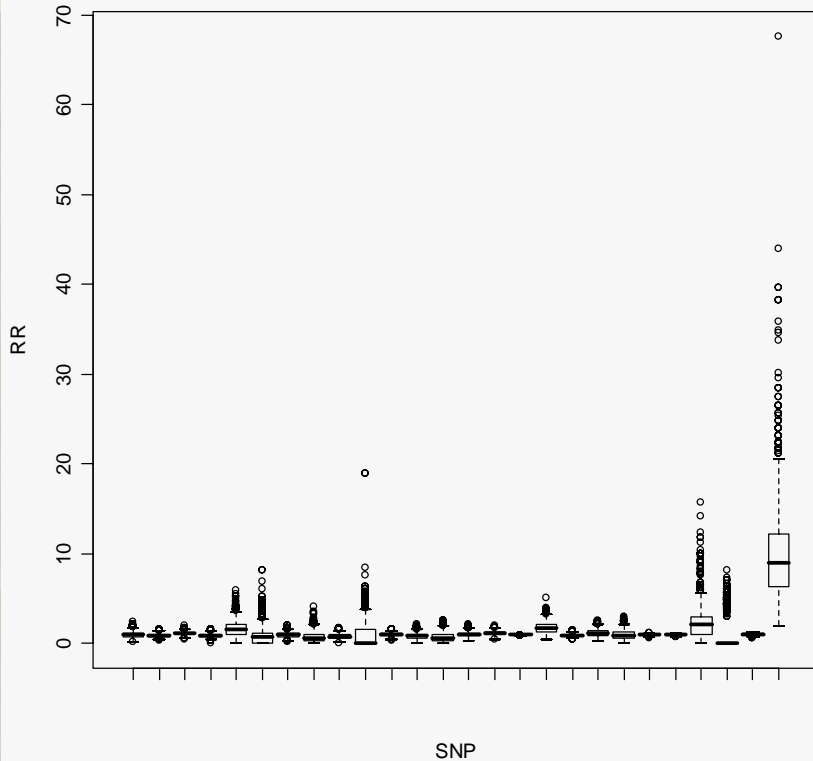
20 Blocks=140 Marker, 100 Fälle und 100 Kontrollen
geringere Häufigkeit=10%, höhere Häufigkeit=25%,
2 Blöcke RR=4, 3 Blöcke RR=2 (RR für zentralen Marker)



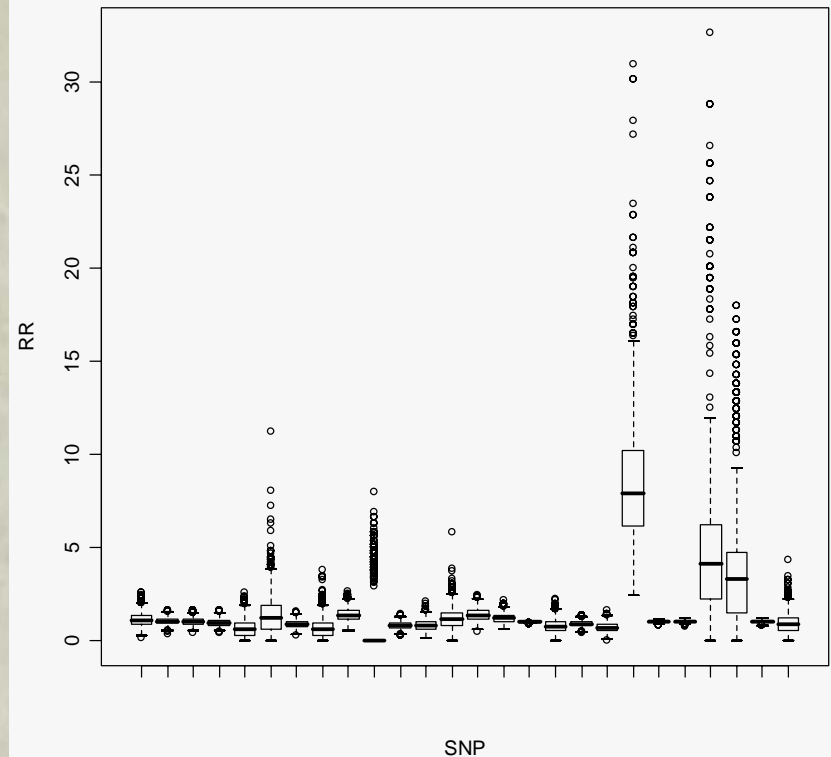
Simulationen von Typ 2

Sich ergebendes RR

SNP mit Risiko am Rand



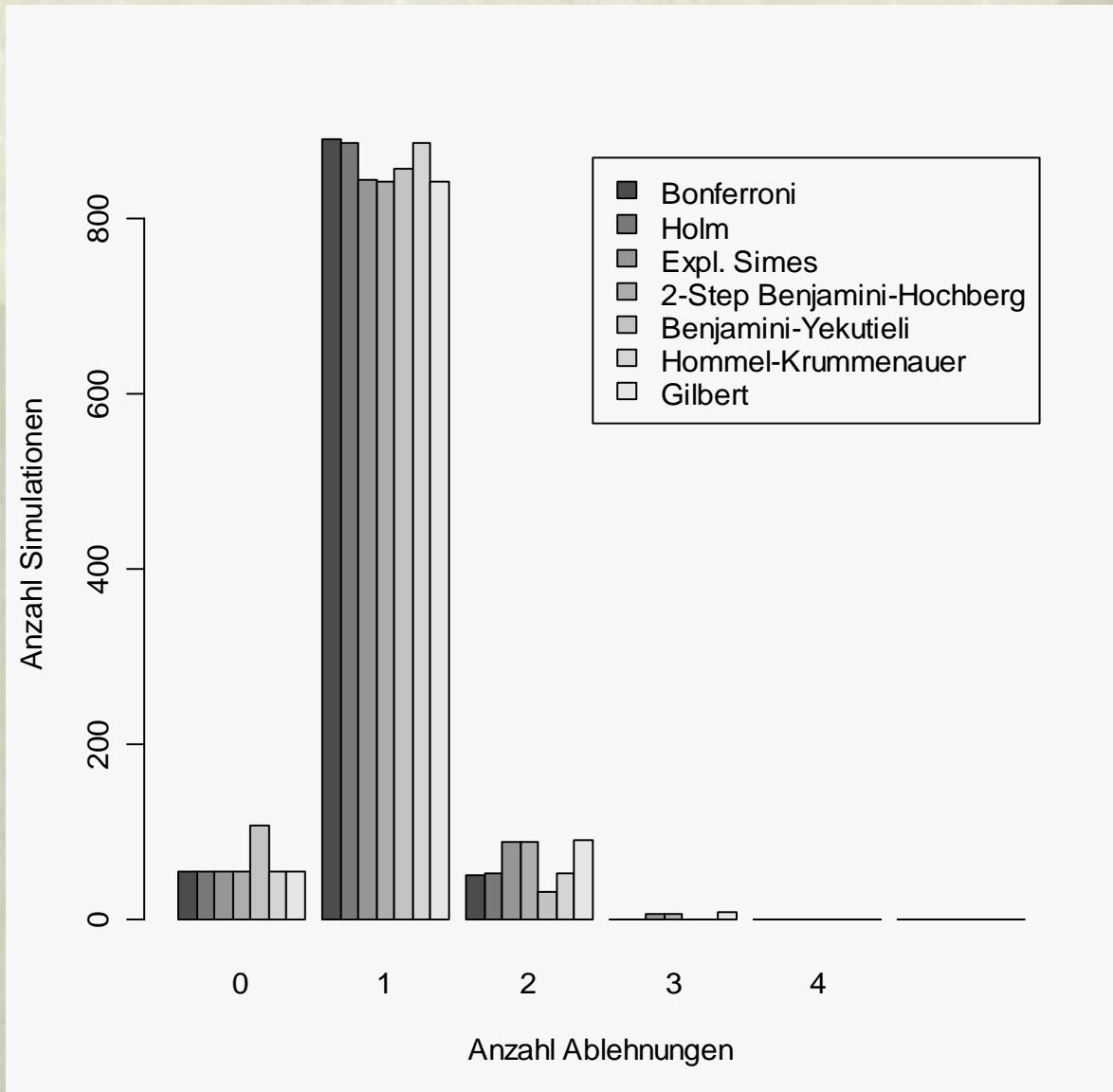
SNP mit Risiko in einem Block



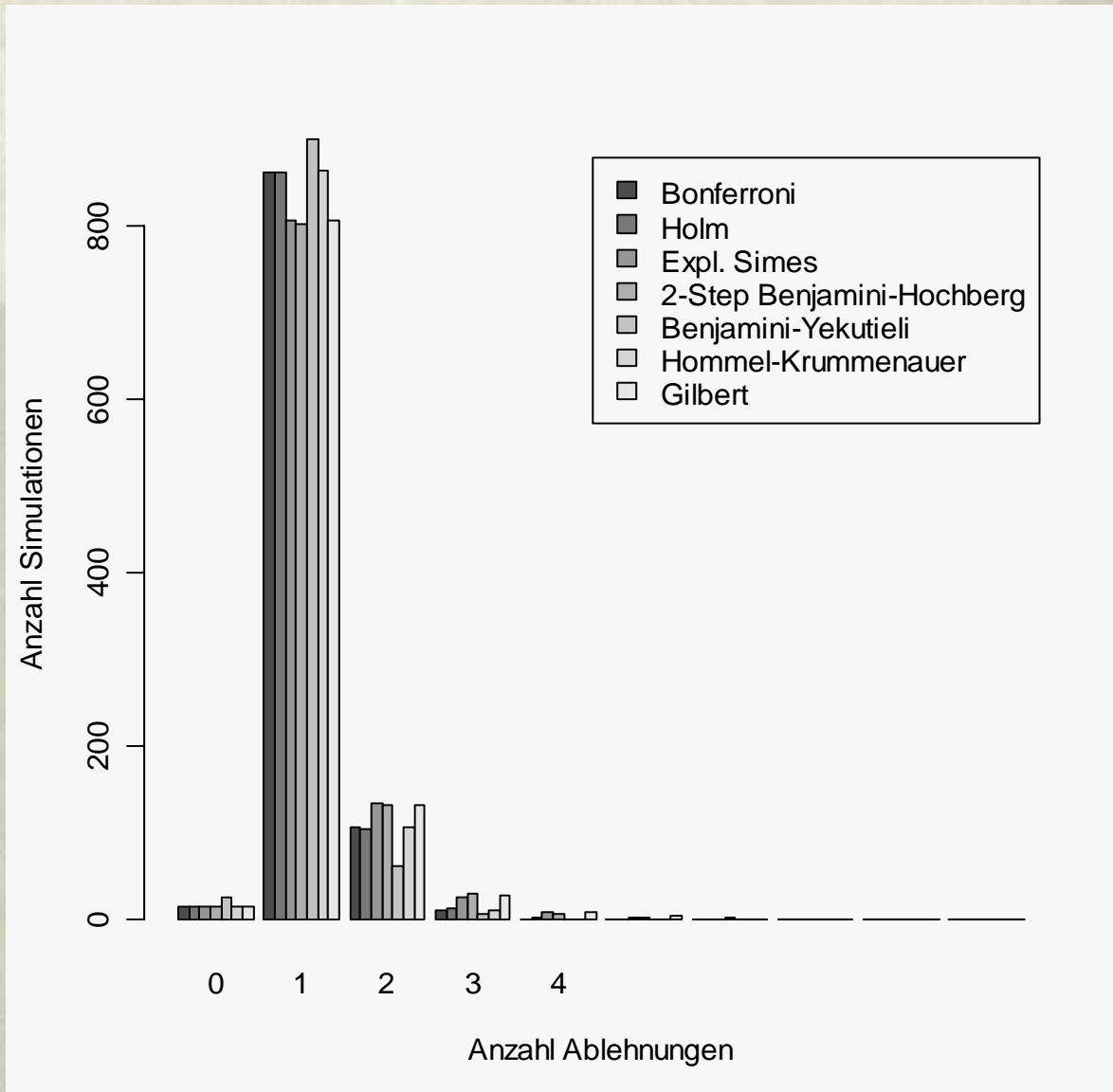
Simulationen von Typ 2

- ❖ Ohne Assoziation
 - Outcome binomial simuliert mit Wahrscheinlichkeit 50% ohne Assoziation zu den SNP Daten
- ❖ So gut wie keine Unterschiede zwischen den Prozeduren
- ❖ alle Prozeduren zeigen in weniger als 5% Simulationen eine oder mehr Ablehnungen
 - 1% der Simulationen bei Benjamini-Yekutieli
 - Sonst 4-5% der Simulationen
 - Also von allen die FWER eingehalten
 - Ohne α -Korrektur jedoch bei >50% der Simulationen mindestens 1 Ablehnung

Typ 2 – Rand SNP assoziiert



Typ 2 – SNP in Block assoziiert



Zusammenfassung

- ❖ Methoden zur Beachtung der kategorialen Daten zeigen nur bei kleiner Fallzahl mehr Power
- ❖ Adaptierter Benjamini-Hochberg scheint in Einzelfällen stärker antikonservativ zu werden
- ❖ Kein exorbitant hoher Gewinn durch FDR, steigt wie erwartet mit dem Anteil an falschen Hypothesen
 - Kaum Gewinn bzgl. mindestens einer Ablehnung
 - Auffälliger ist zumeist eine Steigerung der Anzahl der Ablehnungen

Diskussion

- ❖ Ergebnisse bei Chen et al. (2006)
 - Vergleich explorativer Simes zu min P - Prozedur bei SNPs mit hohem LD (Abhängigkeiten): ergab leicht höhere Power bei min P Test
- ❖ Kropf/Läuter (2002), Kropf et al. (2004)
 - FWER-Kontrolle manchmal besser als FDR-Kontrolle
- ❖ Work in Progress
 - Z.B. Wie verhalten sich Prozeduren, wenn Risiko nicht aufgrund eines SNPs (Markers) entsteht, sondern aufgrund von bestimmten Haplotypen (also bestimmten Kombinationen)?

Literatur

- ❖ Tarone, RE (1990) A modified Bonferroni method for discrete data. *Biometrics* 46: 515-522
- ❖ Hommel, G & Krummenauer, F (1998) Improvements and modifications of Tarone's multiple test procedure for discrete data. *Biometrics* 54:673-681.
- ❖ Gilbert, PB (2005) A modified false discovery rate multiple-comparisons procedure for discrete data, applied to human immunodeficiency virus genetics. *Journal of Applied Statistics (JRSS C)* 44:143-158 (Danke fuer Hinweis von Prof. Hothorn)
- ❖ 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
- ❖ Chen, BE, Sakoda, LC, Hsing, AW and Rosenberg, PS (2006) Resampling-based multiple hypotheses procedures for genetic case-control association studies. *Genetic Epidemiology* 30:495-507