

Aus dem Institut für Medizinische Biometrie und Statistik  
der Universität zu Lübeck  
Direktor: Univ.-Prof. Dr. rer. nat. Andreas Ziegler

# **Das Meta-Analyse Modell nach DerSimonian & Laird mit exakten Gewichten**

Inauguraldissertation  
zur  
Erlangung der Doktorwürde  
der Universität zu Lübeck

Aus der Sektion Informatik / Technik

vorgelegt von

**Michael Preuß**

aus Würselen

Lübeck 2013

1. Berichterstatter: Prof. Dr. Ziegler

2. Berichterstatter: Prof. Dr. Schuster

Tag der mündlichen Prüfung: 04.03.2014

Zum Druck genehmigt.

Lübeck, den

# Inhaltsverzeichnis

<b>Tabellenverzeichnis</b>	<b>6</b>
<b>Abbildungsverzeichnis</b>	<b>7</b>
<b>1 Einleitung</b>	<b>8</b>
1.1 Hintergrund . . . . .	8
1.2 Typen von meta-analytischen Modellen . . . . .	10
1.2.1 Modell mit festen Effekten . . . . .	10
1.2.2 Modell mit zufälligen Effekten . . . . .	12
1.2.3 Regressionsmodelle für Meta-Analysen . . . . .	12
1.2.4 Bayesianische Methoden für Meta-Analysen . . . . .	15
1.2.5 Weitere Typen von meta-analytischen Modellen . . . . .	16
1.3 Variabilität . . . . .	17
1.3.1 Test auf Heterogenität zwischen den Studien . . . . .	19
1.3.2 Das $I^2$ Maß für Heterogenität zwischen den Studien . . . . .	20
1.4 Ziel der Arbeit . . . . .	22
1.5 Aufbau der Arbeit . . . . .	23
<b>2 Das meta-analytische Modell für zufällige Effekte mit exakten Gewichten</b>	<b>25</b>
2.1 Die exakte Verteilung von Cochran's $Q$ . . . . .	26
2.2 Iterative Bestimmung der exakten Verteilung von Cochran's $Q$ . . . . .	27
2.3 Berechnung des Modells für zufällige Effekte mit exakten Gewichten	28
2.4 Herleitung von $I^2$ in Abhängigkeit von $\tau^2$ . . . . .	29
2.5 Berechnung der Konfidenzintervalle von $\tau^2$ und $I^2$ mit Cochran's $Q$	30
2.6 Das R-Paket metaxa . . . . .	30
<b>3 Simulationsstudie I – Charakteristik des Meta-Analyse Modells für zufällige Effekte mit exakten Gewichten</b>	<b>34</b>
3.1 Einführung . . . . .	34
3.2 Simulationsaufbau . . . . .	34
3.3 Ergebnisse . . . . .	37
3.3.1 $\alpha$ -Fehler . . . . .	37
3.3.2 Teststärke . . . . .	40
3.3.3 Das $\tau^2$ mit Konfidenzintervalle . . . . .	41

3.3.4	Vergleich zwischen $I^2$ und dem exaktem $I^2$ und deren Konfidenzintervalle . . . . .	42
3.4	Interpretation der Ergebnisse . . . . .	44
<b>4</b>	<b>Simulationsstudie II – Einfluss der Gewichte des Modells für zufällige Effekte mit exakten Gewichten</b>	<b>46</b>
4.1	Einführung . . . . .	46
4.2	Simulationsaufbau . . . . .	46
4.3	Ergebnisse . . . . .	47
4.3.1	$\alpha$ -Fehler . . . . .	47
4.3.2	Teststärke . . . . .	49
4.4	Interpretation der Ergebnisse . . . . .	51
<b>5</b>	<b>Simulationsstudie III – Meta-Analysen zur Beurteilung von Abweichungen vom Hardy-Weinberg Gleichgewicht</b>	<b>52</b>
5.1	Einführung . . . . .	52
5.2	Maße für das Hardy-Weinberg Gleichgewicht . . . . .	53
5.2.1	Der Inzucht-Koeffizient $f$ . . . . .	54
5.2.2	Der Disequilibrium-Koeffizient $D$ . . . . .	55
5.2.3	Das REH $\omega$ . . . . .	55
5.3	Simulationsaufbau . . . . .	57
5.4	Ergebnisse . . . . .	58
5.4.1	$\alpha$ -Fehler und Teststärke . . . . .	58
5.4.2	Vergleich zwischen dem Modell mit zufälligen Effekten und dem Modell mit zufälligen Effekten mit exakten Gewichten . . . . .	64
5.5	Interpretation der Ergebnisse . . . . .	66
<b>6</b>	<b>Illustration anhand von Realdaten</b>	<b>67</b>
6.1	Diuretika und Präeklampsie Meta-Analyse . . . . .	68
6.2	Unterschiede bei der Signifikanz bezüglich des zufällige Effekte Modells und des exakten zufällige Effekte Modells . . . . .	71
6.3	Multizentrische Studie für den <i>CASP8</i> -652 6N del Promoter Polymorphismus . . . . .	73
<b>7</b>	<b>Diskussion</b>	<b>76</b>

<b>8 Zusammenfassung</b>	<b>79</b>
<b>Literaturverzeichnis</b>	<b>81</b>
<b>Danksagung</b>	<b>89</b>
<b>Lebenslauf</b>	<b>90</b>
<b>Publikationsverzeichnis</b>	<b>91</b>

# Tabellenverzeichnis

3.1	Vierfelder-Tafel Beispiel . . . . .	35
3.2	Parameterzusammenstellung der drei Simulationsszenarien . . . . .	36
3.3	Beispiel für studienspezifische Stichprobengröße . . . . .	37
3.4	$\alpha$ -Fehler für Simulationsstudie I . . . . .	39
3.5	Teststärke für Simulationsstudie I . . . . .	40
4.1	$\alpha$ -Fehler für Simulationsstudie II . . . . .	48
4.2	Teststärke für Simulationsstudie II . . . . .	50
5.1	Teststärke für Simulationsszenario 1 bis 3 . . . . .	61
6.1	Diuretika und Präeklampsie Datensatz . . . . .	69
6.2	Ergebnisse für Diuretika und Präeklampsie Datensatz . . . . .	70
6.3	$\tau^2$ Schätzungen für Diuretika und Präeklampsie Datensatz . . . . .	71
6.4	Einzelstudieneffekte für Hospital- und Inkubator-Datensatz . . . . .	72
6.5	Ergebnisse für den Hospital und den Inkubator Datensatz . . . . .	73
6.6	Ergebnisse für <i>CASP8</i> Promoter Polymorphismus Datensatz . . . . .	74

# Abbildungsverzeichnis

3.1	$\tau^2$ mit Konfidenzintervalle . . . . .	42
3.2	Das (exakte) $I^2$ mit Konfidenzintervalle . . . . .	44
5.1	$\alpha$ -Fehler für Simulationsstudie III . . . . .	60
5.2	$\alpha$ -Fehler für Simulationsstudie I bei Allelfrequenz zwischen 0,1 und 0,2 . . . . .	63
5.3	Vergleich von RE-Modelle für Simulationsstudie III . . . . .	65

# 1 Einleitung

## 1.1 Hintergrund

Meta-Analysen sind Methoden, um zu einer zu einer Informationsverdichtung über viele wissenschaftlich-empirische Einzeluntersuchungen in einem Fachgebiet zu gelangen. Mit den erzielten Ergebnissen erfolgt dann eine Informationsbewertung, deren Ziel eine evidenzbasierte Konsensbildung ist.

Im Vordergrund steht, eine Erklärung der Variabilität zwischen den Ergebnissen von Einzeluntersuchungen vorzunehmen und zu analysieren, falls mehrere Einzeluntersuchungen zuvor keine eindeutigen oder sogar gegensätzliche Ergebnisse hinsichtlich einer spezifischen Fragestellung erzielt haben.

Des Weiteren kann beispielsweise umfangreiches Datenmaterial mit gleicher Fragestellung erst explorativ ausgewertet werden. Darauf aufbauend können später Hypothesen für weitere spezifische Fragestellungen generiert werden.

Insbesondere in der evidenzbasierten Medizin haben Meta-Analysen große Bedeutung erlangt, da sie eine methodische Grundlage darstellen, um Behandlungsempfehlungen bzw. Behandlungsrichtlinien auszusprechen. Sie spielen eine immer größere Rolle, um als Instrument für patientenorientierte Entscheidungen bezüglich empirisch nachgewiesener Wirksamkeit von medizinischen Behandlungen eingesetzt zu werden.

Erste mathematisch fundierte Ansätze einer Meta-Analyse zur systematischen und quantitativen Zusammenführung von Ergebnissen einzelner, ähnlicher Beobachtungen finden sich bereits Anfang des 19. Jahrhunderts.

Carl Friedrich Gauß und Adrien-Marie Legendre entwickelten etwa zur gleichen Zeit die Methode der kleinsten Fehlerquadrate, um Daten von verschiedenen astronomischen Observatorien zusammenzufassen. Dadurch konnte Gauß erstmals eine korrekte Bahnbestimmung des Zwergplaneten Ceres vornehmen [5].

Im Jahr 1904 lieferte Karl Pearson weitere statistische Grundlagen hinsichtlich der Fragestellung, ob ein Impfstoff gegen Typhus wirksam ist. Er sammelte dazu Datenmaterial von 11 relevanten Studien und verglich die Gruppensterblichkeit und Immunität in Bezug auf eine Impfung mithilfe der von ihm entwickelten Methode der Korrelationskoeffizienten.

Im engeren Sinn sind Meta-Analysen ein Bestandteil von systematischen Übersichtsarbeiten. Diese sind ein Instrument der Forschung für die methodische



Suche nach neuen Erkenntnissen. Man leistet also mithilfe dieser Forschung einen Beitrag zur Wissenschaft.

Eine systematische Übersichtsarbeit beginnt mit einer spezifischen Fragestellung, die aus früherer Forschung mit Ergebnissen systematischer Einzeluntersuchungen einhergeht. Mithilfe von geeigneten Methoden wird versucht, verfügbares Wissen zu aquirieren, zu verdichten und zu bewerten. Liegen die Ergebnisse der Einzeluntersuchungen in einer systematischen und quantisierten Form vor, kann eine systematische Übersichtsarbeit um eine Meta-Analyse ergänzt werden. Die Meta-Analyse ist also stets ein Teil einer systematischen Übersichtsarbeit.

Die Haupteigenschaften einer systematischen Übersichtsarbeit umfassen nach Higgins et al. [6]:

- eine klar definierte Zielsetzung mit vordefinierten Auswahlkriterien der Studien,
- eine eindeutige, reproduzierbare Methodologie,
- eine systematische Suche, die darauf abzielt alle Studien zu identifizieren, die die Auswahlkriterien erfüllen,
- eine Beurteilung der Validität der Ergebnisse der enthaltenen Studien,
- eine systematische Darstellung und Zusammenfassung der Studieneigenschaften sowie der Studienergebnisse.

Die statistische Zusammenfassung der Resultate von Einzelstudien wird Meta-Analyse genannt [7]. Nach der Definition von Glass et al. [8] und Higgins et al. [6] ist eine Meta-Analyse die Anwendung von statistischen Methoden, um die ermittelten Effekte von unabhängigen Studien zusammenzufassen. Ziel dabei ist es, durch die Zusammenführung der Information aller relevanten Studien, eine präzisere Schätzung eines Gesamteffekts zu erhalten. Mit der Meta-Analyse erfolgt also mit einer vordefinierten Methodik, eine gewichtete Zusammenfassung von Effektgrößen einzelner Studien, um eine spezifische wissenschaftliche Fragestellung zu beantworten [6].

Betrachtet man die Zusammenführung von Informationen hinsichtlich der Methodik in einem Fachgebiet kommen Blettner et al. [9] zu folgender Klassifizierung:

1. eine qualitative Zusammenfassung bzw. in Form eines Berichts,
2. eine quantitative Zusammenfassung von publizierten Daten; normalerweise als Meta-Analyse bezeichnet,
3. eine, teils wiederholende, Analyse auf Individualebene von Studien; oft auch als gepoolte Analyse oder Meta-Analyse bezeichnet,
4. eine prospektiv geplante und gepoolte Analyse von mehreren Studien. Datenakquirierung, Hypothesen, Vorgehensweise und gepoolte Analyse sind dabei bereits Bestandteil der Studienprotokolle.

In dieser Reihenfolge nimmt der Aufwand hinsichtlich Planung und Zeit zu, da zu jedem nachfolgenden Punkt alle vorherigen Punkte mit berücksichtigt werden müssen.

Mittlerweile ist zur Durchführung einer Meta-Analyse ein sehr breites Spektrum an Auswertungsstrategien bzw. Methoden entwickelt worden, die sich je nach Fragestellung und Forschungsgebiet unterscheiden können [10–12]. Seit Anfang der 1990er Jahre finden deshalb verstärkt Kollaborationen zwischen führenden Wissenschaftlern, Ärzten und Herausgebern statt [13–15], um eine Qualitätssicherung bei der Durchführung, Berichterstattung sowie Registrierung von Meta-Analysen mithilfe einheitlicher Richtlinien und Standards zu erzielen.

## 1.2 Typen von meta-analytischen Modellen

### 1.2.1 Modell mit festen Effekten

Ein weit verbreitetes Modell zur Schätzung einer Meta-Analyse ist das Modell mit festen Effekten (Fixed-Effects Model, FE-Model). Es wird angenommen, dass die wahre Effektgröße  $\theta$  in allen zugrunde liegenden Studien  $i = 1, \dots, k$  immer gleich ist und Unterschiede nur durch zufällige Abweichung zustande kommen. Am wohl bekanntesten ist die Inverse-Varianz-Methode. Die Methode kommt lediglich mit Studieneffekten und deren Varianzen aus, die z.B. aus Kontingenz-

tafeln leicht ermittelt werden können. Die Inverse-Varianz-Methode wurde von Birge [16] und Cochran [17] in den 1930er Jahre erstmals beschrieben: Dabei wird dem geschätzten Effekt einer Studie  $Y_i$  ein Gewicht zugeordnet, welches direkt proportional zur Präzision steht. Dies gilt für das Inverse der Varianzen der Studieneffekte, so dass diese direkt als Gewichte übernommen werden können. Da die Varianzen der Studien in der Praxis zumeist unbekannt sind, werden diese üblicherweise geschätzt.

Mit der Annahme, dass die Effekte der Studien normalverteilt sind und Homogenität des Effekts für alle Studien gilt, ergibt sich der gepoolte Effektschätzer aus dem Verhältnis der gewichteten Effekten zu der Summe aller Gewichte:

$$\hat{\theta}_{FE} = \frac{\sum_{i=1}^k w_i Y_i}{\sum_{i=1}^k w_i}$$

Die Gewichte  $w_i$  der entsprechenden Varianzen  $\sigma_i^2$  der Studien werden für gewöhnlich durch ihre Schätzung  $\hat{\sigma}_i^2$  ersetzt:

$$w_i = \frac{1}{\hat{\sigma}_i^2}$$

Der inverse Varianz gewichtete Effektschätzer  $\hat{\theta}_{FE}$  folgt approximativ einer Normalverteilung. Mithilfe des Standardfehlers  $SE(\hat{\theta}_{FE}) = \sqrt{\sum w_i}^{-1}$  lässt sich eine Teststatistik konstruieren:

$$Z_{FE} = \frac{\hat{\theta}_{FE}}{SE(\hat{\theta}_{FE})} \stackrel{a}{\underset{H_0}{\sim}} N(0,1)$$

Die FE-Methode lässt sich grundsätzlich für Messgrößen nutzen, die approximativ einer Normalverteilung folgen. Neben der Inverse-Varianz-Methode gibt es weitere Methoden, wie die Peto Odds-Ratio-Methode und die Mantel-Haenszel Odds-Ratio-Methode [6], die sich für Studiendaten in Form von Kontingenztafeln gut eignen und einen Inverse-Varianz-Ansatz verfolgen bzw. ihnen das FE-Modell unterliegt.

### 1.2.2 Modell mit zufälligen Effekten

Im Gegensatz zum FE-Modell, bezieht das Modell mit zufälligen Effekten (Random-Effects Model, RE-Modell) Heterogenität zwischen den Studien mit ein. Diese Art von Heterogenität kann viele Ursachen haben. Hauptgründe hierfür sind Unterschiede in der Studienpopulation z.B. kulturelle Unterschiede, im Aufbau bzw. Design der Studie oder durch den Stichprobenfehler. Übersteigt die Variabilität zwischen den Schätzern der Studienergebnisse das Ausmaß der Zufallsvariabilität, würde eine Analyse mit dem FE-Modell zu einem Anstieg des Fehlers 1. Art führen.

Das RE-Modell ist eine Erweiterung des FE-Modells, bei der die Heterogenität zwischen den Studien mit  $\tau^2$  quantifiziert wird. Es gibt diverse Methoden um  $\tau^2$  zu schätzen. Häufig verwendeten Methoden sind die momentenbasierte Schätzung nach DerSimonian und Laird [2] (Gleichung (2.3)), die Maximum-Likelihood Methode [18] und die Restricted-Maximum-Likelihood Methode (REML) [19]. Der Effektschätzer unterliegt einer ähnlich Berechnung wie im FE-Modell 1.2.1 nur wird der zusätzliche Varianzterm  $\tau$  für die Gewichtung berücksichtigt:

$$\hat{\theta}_{RE} = \frac{\sum_{i=1}^k w_i^* Y_i}{\sum_{i=1}^k w_i^*} \text{ mit } w_i^* = w_i + \tau^2$$

Daraus folgernd ist der Standardfehler definiert mit  $SE(\hat{\theta}_{RE}) = \sqrt{(\sum w_i^{-1} + \tau^2)^{-1}}$ . Die Teststatistik lässt sich ähnlich wie im FE-Modell konstruieren mit:

$$Z_{RE} = \frac{\hat{\theta}_{RE}}{SE(\hat{\theta}_{RE})} \stackrel{a}{H_0} \sim N(0, 1)$$

Eine Besonderheit des Modells ergibt sich wenn die wahre Effektgröße  $\theta$  in allen unterliegenden Studien  $i = 1, \dots, k$  gleich ist. Dann geht das RE-Modell in das FE-Modell über, da sich der zusätzliche Varianzterm aufhebt ( $\tau^2 = 0$ ).

### 1.2.3 Regressionsmodelle für Meta-Analysen

Grundsätzlich zielen Regressionsmodelle im Vergleich zu herkömmlichen Meta-Analysen darauf ab, eine Beziehung zu der Größe des Effekts mit bestimmten

Charakteristika (Kovariaten) von Studien zu beschreiben. So kann beispielsweise das Einbeziehen einer studienspezifischen Kovariate, wie z.B. Dosierungsmenge von Medikamenten, das Ziel haben, eine Quantifizierung der Abhängigkeit zwischen relativem Risiko und Medikamentendosis herzustellen. Weiterhin kann insbesondere die Meta-Regression genutzt werden, um Heterogenität (Abschn. 1.3) mit bestimmten Kovariaten zu untersuchen bzw. zu erklären.

Es lassen sich grundsätzlich zwei Regressionsmodelle unterscheiden. Das erste Regressionsmodell für Meta-Analysen wurde von Greenland beschrieben [20]. Dabei handelt es sich um eine Erweiterung des FE-Modells, genannt Meta-Regression. Als zweites Regressionsmodell entwickelten Mitte der 1990er Jahre Berkey et al. [21] das gemischte Modell, welches als Regressionsäquivalent zum RE-Modell gilt.

Sollten Studien in Subgruppen ausgewertet werden, entspricht die herkömmliche Analyse, einer Analyse mit Regressionsmodellen mit einer Subgruppenzugehörigkeit als Kovariate. Jedoch bieten Regressionsmodelle den Vorteil, direkt Unterschiede zwischen den Subgruppen ermitteln zu können anstatt die Effekte separat pro Subgruppe zu ermitteln und dann auf Unterschiede zu verweisen [22].

### 1.2.3.1 Meta-Regression

Zur Berechnung einer Meta-Regression ist ein lineares Modell vonnöten. Der geschätzte Effekt einer Einzelstudie bzw. die abhängige Variable ist im folgenden definiert mit  $Y_i$  und die geschätzte Varianz mit  $\hat{\sigma}_i^2$  wobei  $i = 1, \dots, k$  den Index aller Studien kennzeichnet. Der Effekt einer Studie setzt sich zusammen aus:

$$Y_i = \theta_i + \epsilon_i \text{ mit } \epsilon_i \sim N(0, \sigma_i^2)$$

mit  $\theta_i$  als dem wahren Effekt pro Studie.

Im Gegensatz zum FE-Modell kann  $\theta_i$  variieren. Für  $p$  Kovariaten hat das lineare Modell folgende Form:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}$$

wobei  $x_{i1}, \dots, x_{ip}$  die Werte der Kovariaten und  $\beta_0, \beta_1, \dots, \beta_p$  die unbekannt-ten Regressionskoeffizienten sind. Die Koeffizienten lassen sich mit einer ge-wichteten Methode der kleinsten Quadrate berechnen. Dabei sind die Gewichte (wie im FE-Modell) definiert mit  $w_i = 1/\sigma_i^2$ . Das Ergebnis des Koeffizienten  $\beta_0$  ohne zusätzliche Kovariaten entspricht dabei dem FE-Modell. Approxima-tive Konfidenzintervalle für die Koeffizienten lassen sich auf übliche Weise, bei-spielsweise mit zweiseitigen kritischen Grenzen einer Standardnormalverteilung sowie den Standardfehlern der Koeffizienten, konstruieren:

$$b_j - Z_{\alpha/2} * S_{\beta_j} \leq \beta_j \leq b_j + Z_{\alpha/2} * S_{\beta_j}$$

wobei  $Z$  die Standardnormalverteilung darstellt und  $S_{\beta_j}$  der korrigierte Stan-dardfehler von  $\beta_j$  entspricht [12]. Ein zweiseitiger Signifikanztest  $H_0 : \beta_j = 0$  ist dann mithilfe der Konfidenzintervalle entscheidbar, wenn das Konfidenzin-tervall die 0 nicht enthält.

### 1.2.3.2 Gemischtes Modell

Das gemischte Modell eignet sich vor allem dann, wenn die Heterogenität mit dem Modell der Meta-Regression über die Kovariaten nicht vollständig erklärt werden kann [22]. Es ist eine Erweiterung der Meta-Regression und berück-sichtigt im Besonderen die Heterogenität zwischen den Studien. Es stellt somit eine Erweiterung des RE-Modells dar. Das gemischte Modell besitzt folgende Form:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + E_i$$

Der Unterschied zur Meta-Regression ist die Erweiterung  $E_i$ , welches sowohl den studienspezifischen zufälligen Effekt als auch den Stichprobenfehler enthält:

$$E_i = u_i + e_i \text{ mit } u_i \sim N(0, \tau^2) \text{ und } e_i \sim N(0, \sigma_i^2)$$

Wobei  $\tau^2$  die Heterogenität zwischen den Studien beschreibt und  $e_i$  den Stich-probenfehler pro Studie widerspiegelt.

Die Gleichung enthält also zwei Fehlerterme. Die Varianzen dieser Fehlerterme

sind formuliert mit:

$$v_i^* = \text{Var}(u_i + e_i) = \tau^2 + \sigma^2$$

Die Gewichte ergeben sich wiederum aus der inversen Varianz  $w_i^* = 1/v_i^*$ . Da eine Abhängigkeit zu den Regressionskoeffizienten und  $\tau^2$  besteht, können sowohl iterative Methoden als auch die Maximum-Likelihood Methode (unter der Bedingung einer Normalverteilung von  $Y_i$ ) genutzt werden, um die Gewichte zu berechnen. Herrscht Unsicherheit hinsichtlich einer vollständigen Erklärung der Heterogenität über die Kovariaten, sollte auf das gemischte Modell zurückgegriffen werden, da die Meta-Regression auf einen gemeinsamen Effekt abzielt während das gemischte Modell den Mittelwert einer Verteilung der Effekte über alle Studien berücksichtigt [22].

#### 1.2.4 Bayesianische Methoden für Meta-Analysen

Der bayesianische Ansatz unterscheidet sich konzeptuell vom frequentistischen Ansatz. Während beim frequentistischen Ansatz ausschließlich auf die erhobenen Daten zurückgegriffen wird, kann mit dem bayesianischen Ansatz noch zusätzliche *Evidenz* bzw. Expertenwissen in Form von a priori festgelegten Verteilungen (A-priori-Verteilungen) eingebracht werden. Beispielsweise lassen sich unter Berücksichtigung von Richtlinien oder Kosten Entscheidungshilfen konstruieren. Sowohl die Datenbasis als auch die Modellparameter dieser Verteilungen werden als zufällige Größen angesehen. Eine Likelihood-Funktion dient dann als Konnektor, der die Plausibilität der datengegebenen Werte für die Modellparameter erklärt [23]. Mit dem Theorem von Bayes wird unter Zuhilfenahme der Likelihood-Funktion  $P(\text{Data} | \theta)$  und der A-priori-Verteilung  $P(\theta)$  auf die sogenannte A-posteriori-Verteilung geschlossen:

$$P(\theta | \text{Data}) \propto P(\theta)P(\text{Data} | \theta)$$

Eine weitere Gruppe von Methoden lässt sich dem empirisch-bayesianischen Ansatz zuordnen, bei dem Schlüsselemente des vollständigen bayesianischen Ansatzes wiederum mithilfe der erhobenen Daten geschätzt werden und somit die zwangsläufig subjektive Komponente von Expertenwissen abgeschwächt wird.

Bayes Methoden wurden in den vergangenen Jahren immer häufiger eingesetzt, da es nicht zuletzt große Fortschritte bei den rechenintensiven Methoden wie Markov-Chain-Monte-Carlo (MCMC) und Gibbs-Sampling [24] gibt, die insbesondere benötigt werden, um die A-posteriori-Verteilung zu berechnen.

### 1.2.5 Weitere Typen von meta-analytischen Modellen

Zu den bisher genannten Meta-Analysen existieren noch zahlreiche Erweiterungen, Abwandlungen sowie datenspezifische Meta-Analysen. Nachfolgend kann nur ein Ausschnitt von weiteren Typen von meta-analytischen Modellen präsentiert werden. Für eine vollständige und ausführliche Beschreibung wird auf folgende Gesamtwerke verwiesen [10–12, 25–28].

Die *kumulative Meta-Analyse* ist eine Erweiterung hinsichtlich der Zeitachse. Neue Studien werden mit den bisherig gepoolten Studien mit den gleichen Methoden gepoolt. Zusätzliche Auswertungsstrategien können dann genutzt werden: z.B. Beobachtung der Entwicklung des Effekts in der Zeit oder Anordnung der Studien nach bestimmten Kriterien (z.B. Ordnung der Kontrollgruppen nach Grundrisiko) [29].

Für *Meta-Analysen von Ereigniszeitanalysen* ist der Effekt in Abhängigkeit von der Zeit vom besonderen Interesse. Zur Modellierung von Ereigniszeiten spielen insbesondere der Kaplan-Meier-Schätzer sowie die Cox-Regression eine zentrale Rolle. Man benötigt spezielle Methoden, um Studien dieser Art zu kombinieren, da sich mit zensierten Daten die Auswertung komplex gestaltet. Je nach Datengrundlage ergibt sich eine Auswahl an Methoden. Sie erstrecken sich von einfacher Gewichtung und Kombination von Differenzen von Überlebensraten, über die Berechnung eines Gesamtschätzers des Effekts, bis hin zu einer Kombination auf Individualebene der Daten sämtlicher Studien [30].

Die *Meta-Analyse von individuellen Patienten Daten (IPD)* hebt sich von den bisher genannten Meta-Analysen ab. Anstatt auf aggregierte Daten zurückzugreifen, wie es in den vorherigen Meta-Analysen der Fall war, werden in diesem Fall die Individualdaten der einzelnen Studien betrachtet. Dies hat mehrere Vorteile [31, 32]:

1. Daten der Studien lassen sich besser auf ihre Qualität überprüfen



2. Subgruppen können feiner granuliert werden, um beispielsweise weitere Hypothesen bzw. Auswertungsmöglichkeiten definieren zu können
3. Ein größerer Konsens der Studienprotokolle kann hervorgebracht werden z.B. gemeinschaftliche Ein-/Ausschlusskriterien für alle Studien oder auch die Überprüfung von Interaktionen zwischen Behandlungseffekt und Patientenprofile

Eine Auswertung von Individualdaten in Meta-Analysen ist häufig kosten- und zeitintensiv gegenüber alternativer Methoden. Allerdings mit dem Vorteil einer hohen Qualität und Evidenz der Daten, da eine direkte Kooperation und Kommunikation mit Projektpartnern der einzelnen Studien zumeist unumgänglich ist.

*Netzwerk-Meta-Analysen* bieten die Möglichkeit indirekte Vergleiche von Behandlungseffekten vorzunehmen, falls keine (aussagekräftigen) direkten Vergleiche vorliegen (beispielsweise der Vergleich von Behandlungseffekten aus zwei oder mehr Behandlungsarmen mit gleicher Ausgangsbehandlung). Dabei ist von den gleichen Grundannahmen wie für paarweise Vergleiche auszugehen.

Des Weiteren muss eine besondere Sorgfalt bei der Einschätzung und Kontrolle von Heterogenität und Inkonsistenzen von Meta-Analysen geltend gemacht werden. Da u.a. hierfür bisher keine klar akzeptierten Regeln vorliegen, haben das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), die Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS), und die Deutsche Region der Internationalen Biometrischen Gesellschaft (IBS-DR) in einer gemeinsamen Stellungnahme verkündet, dass Netzwerk-Meta-Analysen in der Regel zu Aussagen mit einer geringeren Ergebnissicherheit führen als Meta-Analysen mit direkten Studienvergleichen [33].

## 1.3 Variabilität

Von großer und nicht zu unterschätzender Bedeutung ist die Betrachtung der Variabilität in Meta-Analysen. Intuitiv ist es verständlich, dass bei gepoolten Studien praktisch immer eine Abweichung zur wahren Effektgröße besteht. Sei es

durch abweichende Definition des Studienprotokolls, unterschiedliche Skalenvorwendung zur Messung des Behandlungseffekts, der Variabilität bei der Definition von Erkrankungen oder auch unerklärlichen sonstigen Gründen.

Bleibt die Variabilität bei der Durchführung einer Meta-Analyse außer Betracht, kann dies zu unerwünschten Effekten führen, wie z.B. eine Erniedrigung der Spezifität. Insbesondere können diese Ergebnisse einer Meta-Analyse dann auch als fraglich eingestuft werden [34]. Um Variabilität im statistischen Sinne einzubeziehen, muss zuvor eine Quantifizierung dieser vorgenommen werden. Zur Unterscheidung des allgemein gehaltenen Begriffs der Variabilität und der Quantifizierung im statistischen Sinne spricht man auch von der statistischen Heterogenität bzw. einfach nur Heterogenität. Eine Unterteilung dieser in folgenden Kategorien erleichtert die Unterscheidung [6]:

**Klinische Heterogenität** bezieht sich auf Heterogenität die beispielsweise in Studiendaten auftritt. Im Folgenden können drei Ebenen unterschieden werden:

- auf der Patientenebene kann es unterschiedliche Anteile in Bezug auf Alter, Geschlecht und Krankheitsausprägung geben, die zu einer Schwankung des Effekts führen,
- auf der Interventionsebene kommen z.B. Dosierung eines Medikaments oder die Dauer und der Zeitpunkt einer Therapie für eine Heterogenität zwischen den Studien in Frage,
- auf der Endpunktebene können Definitionen von Krankheiten sowie der Zeitpunkt der Erhebung ausschlaggebend sein.

**Methodologische Heterogenität** kann gegeben sein, wenn sich die Methoden bei der Auswertung von Studien unterscheiden. Eine weitere Quelle ist der Publikationsbias, der durch nicht publizierte, meist nicht signifikante Ergebnisse ebenfalls zu einer Heterogenität führt. Auch Ausreißerstudien bzw. Studien mit geringer Qualität können zu Heterogenität zwischen den Studien beitragen, wenn diese in die Meta-Analyse einfließen.

**Statistische Heterogenität** bedeutet, dass die wahren Effekte in jeder Studie nicht identisch sind. Klinische und methodologische Heterogenität führen beispielsweise zu statistischer Heterogenität. Anders formuliert, ergibt sich diese Art der Heterogenität, wenn die Abweichung zu den beobachteten Effekten größer ist als man es durch Ziehen von zufälligen Stichproben erwarten würde [35].

### 1.3.1 Test auf Heterogenität zwischen den Studien

Sehr häufig wird die Heterogenität zwischen den Studien durch eine modifizierte  $\chi^2$ -Teststatistik nachgewiesen.

Die Nullhypothese besagt, dass die wahren Effekte  $\theta_i$  aller Studien  $i = 1, \dots, k$  als gleich angenommen werden. Gegeben der Notation wie im Abschnitt 1.2.1, lautet die Nullhypothese entsprechend  $H_0 : \theta_1 = \theta_2 = \dots = \theta_k$  gegen die Alternativhypothese, dass sich mindestens eine Effektgröße  $\theta_i$  von den anderen unterscheidet. Für die klassische Cochran's Q-Teststatistik [36, 37] ergibt sich:

$$Q = \sum_{i=1}^k w_i Y_i^2 - \frac{\left(\sum_{i=1}^k w_i Y_i\right)^2}{\sum_{i=1}^k w_i}$$

Dabei folgt  $Q$  unter  $H_0$  einer  $\chi^2$ -Verteilung mit  $k - 1$  Freiheitsgraden unter  $H_0$  wodurch der zugehörige  $p$ -Wert ermittelt werden kann.

Die Teststärke kann bei einer geringen Anzahl von Studien ( $k < 10$ ) sehr gering sein [38].

Des Weiteren ergibt sich das Problem, dass nur der Nachweis der Heterogenität erzielt wird, nicht aber die Größe der Heterogenität selbst. Also nicht „Wie viel Heterogenität ist vorhanden?“ sondern nur „Ist Heterogenität vorhanden?“. Sollten „zu viele“ Studien in den Test einfließen, kann der Test in seiner Mächtigkeit übermäßig groß werden, so dass selbst kleinste Heterogenität, obwohl unbedeutend, nachgewiesen wird.

Aufgrund der Abhängigkeit der Studienanzahl und der damit einhergehenden starken Schwankung der Mächtigkeit des Tests, ergibt sich das Problem ein passendes festes Signifikanzniveau zu wählen [35, 39]. Higgins et al. schlagen als Alternative das Heterogenitätsmaß  $I^2$  vor, welches eine Unabhängigkeit

gegenüber der Studienanzahl bietet (Abschn. 1.3.2).

Zur Aufdeckung von Heterogenitätsquellen besteht die Möglichkeit zur Durchführung einer Meta-Regression (Abschn. 1.2.3.1) oder einer Sensitivitätsanalyse. Für letzteres wird die primäre Analyse mit alternativen Auswahlkriterien (beispielsweise Kriterien zum Einschluss von Studien (Kap. 1)) durchgeführt, um der Frage nachzugehen, ob die Ergebnisse der primären Analyse robust gegenüber dem Prozess zur Findung von Auswahlkriterien sind [6, Abschn. 9.7].

### 1.3.2 Das Heterogenitätsmaß $I^2$ für Heterogenität zwischen den Studien

Ein Heterogenitätsmaß für Meta-Analysen beschreibt den Anteil der Variabilität der nicht durch Stichprobenfehler entstanden ist [40]. Dieses Maß beruht dementsprechend auf der Reproduzierbarkeit (Präzision) der einzelnen geschätzten Varianzen der Studieneffekte [39]. Da Meta-Analysen meist eine unterschiedliche Studienanzahl besitzen, muss das Heterogenitätsmaß entsprechend der Anforderung einer Unabhängigkeit gegenüber der Anzahl der Studien erfüllen. Dann erst ist es überhaupt möglich, das Ausmaß der Heterogenität zwischen mehreren Meta-Analysen vergleichbar zu machen [41].

Ein Test auf Heterogenität (z.B. Cochran's Q-Teststatistik; siehe Abschn. 1.3.1) ist für ein Heterogenitätsmaß nicht geeignet, da es lediglich die Nullhypothese überprüft, ob alle Studien einen gleich großen Behandlungseffekt aufweisen. Dieser Test ist jedoch von der Studienanzahl abhängig, da die Heterogenität mit wachsender Studienanzahl steigt.

Das gleiche gilt für die Interstudienvarianz  $\tau^2$ , die abhängig von der Anzahl sowie der Größe von Studien in Meta-Analysen ist, deren Wert aber nicht unbedingt systematisch ansteigen muss [40]. Jedoch kann  $\tau^2$  nur das Ausmaß, nicht aber den Anteil der Variabilität erklären, der durch die Zwischenstudienvarianz verursacht ist. Somit scheidet  $\tau^2$  als brauchbares Heterogenitätsmaß aus [39].

Das von Higgins und Thompson [39] entwickelte  $I^2$  berücksichtigt die Unabhängigkeit gegenüber der Studienanzahl und erfüllt so die Anforderung für ein

sinnvolles Heterogenitätsmaß. Das  $I^2$  ist definiert mit [39]:

$$I^2 = \begin{cases} 100\% \times \frac{Q-(k-1)}{Q}; & \text{für } Q > (k-1) \\ 0; & \text{für } Q \leq (k-1) \end{cases}$$

wobei  $Q$  der Wert aus Cochran's  $Q$ -Teststatistik und  $k$  die Anzahl der Studien darstellt.

Das  $I^2$  bietet zudem folgende Eigenschaften [41]:

- es unterstützt eine intuitive Interpretation durch die Beschreibung des prozentualen Anteils der totalen Varianz,
- es lässt sich leicht ein Konfidenzintervall für das  $I^2$  konstruieren,
- es ist einfach herzuleiten und kann oft aus bereits publizierten Meta-Analysen berechnet werden,
- es ist grundsätzlich nicht abhängig von der Studienanzahl,
- es lässt ähnliche Interpretationen für unterschiedliche Maßeinheiten (dichotom, quantitativ, Odds-Ratio usw.) zu.

Jedoch sind gewisse Punkte bei der Interpretation des  $I^2$  zu beachten. Das  $I^2$  ist (wie auch die  $Q$ -Statistik) abhängig von der Präzision, die wiederum proportional zur Studiengröße ist. D.h. bei steigender Anzahl von Studiengrößen (und fixem  $\tau^2$ ) in einer gegebenen Meta-Analyse, werden die Konfidenzintervalle der Studieneffekte kleiner, aber damit auch die Heterogenität größer. Das wiederum hat zur Folge, dass auch das  $I^2$  größer wird [42]. Des Weiteren darf das  $I^2$  keinesfalls als Teststatistik betrachtet werden, sondern nur als Maß, Index oder Punktschätzer [40].

Es sollte zudem nicht als Basis zum Ein- oder Ausschluss von Studien verwendet werden. Denn Higgins [43] kommentierte, dass erstaunlicherweise gerade der Ausschluss von Studien mit einem Effekt nahe des gepoolten Schätzers der effizienteste Weg ist, das  $I^2$  zu reduzieren.

Zur Unterstützung der Interpretation von  $I^2$  schlagen Higgins und Thompson [41] vor folgende Attribute zu verwenden:

- $I^2 \approx 25\%$ : niedrige Heterogenität,
- $I^2 \approx 50\%$ : moderate Heterogenität,
- $I^2 \approx 75\%$ : hohe Heterogenität.

Higgins und Thompson betonen jedoch dazu, dass es zu vereinfachend ist, feste Grenzen bzw. Kategorien für Werte von  $I^2$  festzulegen [41]. Jedoch kann mithilfe der Attribute auf eine Tendenz aus den  $I^2$  Werten geschlossen werden.

In der Literatur existieren abweichende Definitionen der Grenzen von der ursprünglichen Definition von Higgins und Thompson [6, 42, 44]. Beispielsweise empfiehlt die Cochrane Collaboration folgende alternative Klassifizierung für das  $I^2$  [6] :

- 0% – 40% möglicherweise unbedeutende Heterogenität,
- 30% – 60% kann moderate Heterogenität repräsentieren,
- 50% – 90% kann erhebliche Heterogenität repräsentieren,
- 75% – 100% beträchtliche Heterogenität.

Die Klassengrenzen sind dabei bewußt nicht eindeutig gewählt worden, da die Bedeutung der  $I^2$  Werte sowohl von der Stärke der Evidenz, als auch vom Ausmaß und Richtung der Studieneffekte abhängt.

## 1.4 Ziel der Arbeit

Biggerstaff und Tweedie [45] entwickelten für Meta-Analysen einen Ansatz zur Abschätzung der Variabilität der Zwischenstudienvarianz auf Grundlage der approximativen Cochran's  $Q$ -Verteilung. Später leiteten Biggerstaff und Jackson [1] auch die exakte Cochran's  $Q$ -Verteilung her.

Sie schätzten ab, dass sich die exakte Verteilung besonders bei Meta-Analysen mit geringer Anzahl von Einzelstudien sowie hoher Heterogenität eignen könnte,

wie schon die asymptotischen Verteilungen für moderate Heterogenität. Weiterhin empfohlen sie, die exakte Verteilung für routinemäßig eingesetzte Methoden hinsichtlich Meta-Analysen aufzunehmen, sowie die daraus neu gewonnenen Ansätzen einer umfassenden Untersuchung zu unterziehen.

Auf Grundlage dieser Vorarbeiten wird in dieser Arbeit die Variabilität von  $\tau^2$  mithilfe der exakten Verteilung von Cochran's  $Q$  geschätzt und erstmalig im DerSimonian und Laird (DSL) zufällige Effekte Modell für Meta-Analysen [2] berücksichtigt. Daraus ergibt sich das neue zufällige Effekte Modell mit exakten Gewichten.

Es werden Simulationsstudien zur Bestätigung der Validität des neuen Ansatzes durchgeführt. Hierfür wird der Ansatz mit dem etablierten DSL zufällige Effekte Modell verglichen.

Des Weiteren werden umfangreiche Simulationsszenarien und Beispiele anhand von Realdaten betrachtet, um die Leistung des neuen Ansatzes, bei hoher Heterogenität, im Vergleich zum herkömmlichen DSL-Ansatz beurteilen zu können.

Darüber hinaus wird ein neues R-Paket vorgestellt, welches im Zuge dieser Arbeit entwickelt wurde, in der die neue Methode für Meta-Analysen im vollen Umfang implementiert und leicht zu bedienen ist.

## 1.5 Aufbau der Arbeit

In Kapitel 2 erfolgt die Herleitung sowie eine Vereinfachung für das zufällige Effekte Modell mit exakten Gewichten auf Grundlage der Vorarbeiten von Biggerstaff und Tweedie [45] sowie Biggerstaff und Jackson [1], die bereits den asymptotischen Ansatz entwickelt haben.

Weiterhin werden in diesem Kapitel Maße zur Beurteilung der Heterogenität im Kontext zur exakten Verteilung von  $Q$  unter Berücksichtigung der Variabilität des Schätzers für die Heterogenität zwischen den Studien betrachtet.

In Kapitel 3 wird für das neue Gewichtungmodell, das DerSimonian und Laird zufällige Effekte Modell mit exakten Gewichten, mithilfe umfangreicher Simulationsstudien auf Validität geprüft.

In einer zweiten Simulationsstudie (Kap. 4) werden Unterschiede zwischen dem klassischen Ansatz und dem Ansatz mit exakten Gewichten ausgearbeitet.

Kapitel 5 beschäftigt sich mit einem Anwendungsfeld für Meta-Analysen in der Populationsgenetik. Hier werden Maße zur Beurteilung der Abweichung vom Hardy-Weinberg Gleichgewicht, mit denen auch eine Quantifizierung erfolgen kann, mithilfe von Simulationsstudien sowohl im Hinblick auf die Maße selbst, als auch im Hinblick zwischen den meta-analytischen Ansätzen verglichen.

Für Anschauungszwecke werden in Kapitel 6 drei reale Datensätzen aus klinischen Studien betrachtet. Anhand dieser Datensätze werden u.a. Unterschiede hinsichtlich der Gewichtung einzelner Studien, sowie Unterschiede hinsichtlich der Signifikanz zwischen dem klassischen DerSimonian und Laird zufällige Effekte Modell und dem Modell für zufällige Effekte mit exakten Gewichten betrachtet.

Eine Interpretation und Diskussion der Ergebnisse und Aussichten für weitere Arbeiten erfolgt im Kapitel 7.



## 2 Das meta-analytische Modell für zufällige Effekte mit exakten Gewichten

Die hier beschriebenen Herleitungen basieren auf die Verteilung der Cochran's  $Q$ -Statistik. Die Bedeutung der Verteilung von Cochran's  $Q$  ist gut abzulesen, wenn diese im Kontext zum „Test auf Heterogenität zwischen Studien“ (Abschn. 1.3.1) betrachtet wird. Deswegen wird im ersten Abschnitt zunächst auf die Cochran's  $Q$ -Statistik eingegangen und darauf aufbauend die exakte Verteilung von  $Q$  hergeleitet.

Mithilfe der exakten Verteilung von  $Q$  erhält auch  $\tau^2$  direkt Verteilungseigenschaften womit schließlich die Varianz von  $\tau^2$  ermittelt werden kann. Mittels der nun bekannten Verteilungsfunktion von  $\tau^2$  ist es nun auch möglich die Variabilität von  $\tau^2$  für die Gewichte des zufälligen Effekte Meta-Analyse Modells zu berücksichtigen. Zur Berechnung der Cochran's  $Q$ -Verteilung existieren sowohl asymptotische Methoden als auch die exakte Methode (Kap. 2.1).

Im Hinblick auf das RE-Modell wurde die Herleitung von Biggerstaff und Tweedie in allgemeiner Form und explizit nur unter Verwendung der asymptotischen Cochran's  $Q$ -Verteilung [45] vorgenommen. In dieser Arbeit erfolgt die Herleitung für den exakten Ansatz und zusätzlich eine Vereinfachung, so dass die ermittelte Variabilität von  $\tau^2$  mithilfe der exakten Verteilung von  $Q$  in ein neues Modell einfließt: das „zufällige Effekte Meta-Analyse Modell mit exakten Gewichten“. In der weiteren Ausführung werden auch synonyme Bezeichnung benutzt: das „exakte RE-Modell“ bzw. das „RE-Modell mit exakten Gewichten“. Die nachfolgenden Abschnitt beschreiben die Herleitung sowie weitere Verwendungsmöglichkeiten für die exakte Verteilung von Cochran's  $Q$ . Insbesondere das exakte  $I^2$ , genauer das  $I^2$  unter Berücksichtigung der Variabilität von  $\tau^2$  mithilfe der exakten Verteilung von Cochran's  $Q$ , sowie die exakten Konfidenzintervalle für das exakte  $I^2$  und das  $\tau^2$ .

## 2.1 Die exakte Verteilung von Cochran's Q

Wie schon in 1.3.1 beschrieben, lässt sich die Cochran's Q-Statistik auch als Test auf Heterogenität zwischen den Studien nutzen. Weiterhin ist es möglich, die Abhängigkeit zwischen der Cochran's Q-Statistik und der Heterogenität mithilfe der exakten Verteilung von Q zu betrachten. Zur Herleitung der exakten Verteilung wird mit der klassischen Cochran's Q-Teststatistik begonnen. Diese ist definiert durch:

$$Q = \sum w_i (Y_i - \hat{\theta}_{FE})^2 \quad (2.1)$$

wobei  $w_i$  die Gewichte aus der inversen Varianz  $1/\sigma_i$  der Studieneffekte  $Y_i$  sind und  $\hat{\theta}_{FE}$  der FE-Schätzer darstellt (Abschn. 1.2.1).

Eine äquivalente Schreibweise von Q in Matrixnotation ergibt:

$$Q = \mathbf{Y}^t \mathbf{A} \mathbf{Y} \quad (2.2)$$

wobei  $\mathbf{Y}^t$  den transponierten Vektor von Y darstellt. Die Zentrierungsmatrix  $\mathbf{A}$  lässt sich als Diagonalmatrix  $\mathbf{W}$  mit den Werten  $w_i = 1/\sigma_i^2$  sowie dem Spaltenvektor  $\mathbf{w}$  und dem Zeilenvektor  $\mathbf{w}^t$  mit den Gewichten  $w_i$  definieren:

$$\mathbf{A} = \mathbf{W} - \left( \frac{1}{\sum w_i} \right) \mathbf{w} \mathbf{w}^t$$

Sei  $\Sigma$  die diagonale Varianz-Kovarianz Matrix von Y mit  $\Sigma = \text{diag}(\sigma_i^2 + \tau^2)$ . Außerdem sei  $\mathbf{Z}$  ein  $k$ -dimensionaler multivariater standardnormalverteilter Zufallsvektor, dann ergibt sich:

$$Q = \mathbf{Y}^t \mathbf{A} \mathbf{Y} \stackrel{d}{=} \mathbf{Z}^t \Sigma^{1/2} \mathbf{A} \Sigma^{1/2} \mathbf{Z} = \mathbf{Z}^t \mathbf{S} \mathbf{Z}$$

Mithilfe des Spektralsatzes lässt sich  $\mathbf{S}$  nach Biggerstaff und Jackson umschreiben in [1]:

$$Q \stackrel{d}{=} \sum_{i=1}^k \lambda_i (\mathbf{v}_i^t \mathbf{Z})^2$$

wobei  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$  die geordneten Eigenwerte von  $\mathbf{S}$  und  $\mathbf{v}_i^t \mathbf{Z}$  die orthonormal unabhängigen Verteilungen mit  $\mathbf{v}_i^t \mathbf{Z} \sim N(0, 1)$  sind.

Dies lässt sich wiederum als unabhängige  $\chi^2$ -Zufallsvariable mit einem Freiheits-

grad ausdrücken:

$$Q \stackrel{d}{=} \sum_{i=1}^k \lambda_i \chi_i^2(1)$$

Die Eigenwerte sind mittels Standardmethoden einfach zu ermitteln. Jedoch ist die Bestimmung der kumulativen Verteilungsfunktion  $F_Q$  von (2.1), definiert mit:

$$F_Q(t; \tau^2)$$

einer positiven Linearkombination der  $\chi_i^2$  unabhängigen Zufallsvariablen, nur numerisch möglich (Abschn. 2.2).

Für den einfachen Fall mit  $\tau^2 = 0$  ergeben sich beispielsweise Eigenwerte mit  $\lambda_i = 1$  für Gleichung (2.1). In diesem Fall folgt  $Q$  einer  $\chi^2(k-1)$ -Verteilung [1]. Des Weiteren gibt es diverse asymptotische Methoden zur Berechnung der exakten  $Q$ -Statistik. Zu ihnen gehören u.a. eine zwei Momenten-Gamma Methode, eine Typ-III Momenten-Pearson Methode oder die Sattelpunktmethode [1].

## 2.2 Iterative Bestimmung der exakten Verteilung von Cochran's $Q$

Zur iterativen Bestimmung der exakten Verteilung von  $Q$  existieren u.a. Algorithmen von Farebrother [46] und Imhof [47]. Beide Algorithmen stehen beispielsweise im R-Paket *CompQuadForm* von P. Lafaye de Micheaux [48] zur Verfügung. Mit dem Paket lassen sich Gleichungssysteme mit Verteilungsfunktionen (Normalverteilungen in quadratischer Form) mit iterativen Algorithmen zu lösen.

Zur Berechnung der exakten Verteilung von Cochran's  $Q$  fließen die Eigenwerte von  $\mathbf{S}$  in die Algorithmen ein, die quasi eine andere Darstellung von  $\tau^2$  repräsentieren. Mittels der oben genannten Algorithmen ist die Berechnung der kumulativen Verteilungsfunktion von  $Q$ , repräsentiert durch  $F_Q(x, \tau^2)$ , möglich.

## 2.3 Berechnung des Meta-Analyse Modells für zufällige Effekte mit exakten Gewichten

Nach Biggerstaff und Tweedie [45] sind die Gewichte mit der Verteilung von  $\tau^2$  für das RE-Modell für den DSL-Ansatz [2] definiert mit:

$$\begin{aligned} w_i^*(\tau^2) &= E[w_i(\hat{\tau}_{DSL}^2)] \\ &= F_{DSL}(0; \tau^2) w_i(0) + \int_0^\infty f_{DSL}(t; \tau^2) w_i(t) dt \end{aligned}$$

wobei die Gewichte als Funktion  $w_i(t) = (\sigma_i^2 + t)^{-1}$  definiert sind.  $F_{DSL}(t; \tau^2)$  und  $f_{DSL}(t; \tau^2)$  stellen die kumulative Verteilungsfunktion und die Dichtefunktion von  $\tau^2$  dar. Mithilfe der partiellen Integration lässt sich dieser Term vereinfachen in:

$$\begin{aligned} w_i^*(\tau^2) &= F_{DSL}(0; \tau^2) w_i(0) + \left[ F_{DSL}(t; \tau^2) w_i(t) \right]_0^\infty - \int_0^\infty F_{DSL}(t; \tau^2) w_i'(t) dt \\ &= \underbrace{F_{DSL}(0; \tau^2) w_i(0) - F_{DSL}(0; \tau^2) w_i(0)}_0 + \underbrace{F_{DSL}(\infty; \tau^2) w_i(\infty)}_0 - \int_0^\infty F_{DSL}(t; \tau^2) w_i'(t) dt \\ &= - \int_0^\infty F_{DSL}(t; \tau^2) w_i'(t) dt = \int_0^\infty \frac{1}{(\sigma_i^2 + t)^2} F_{DSL}(t; \tau^2) dt \end{aligned}$$

mit  $w_i'(t) = \frac{-1}{(\sigma_i^2 + t)^2}$ . Bei der Herleitung wurde die gebräuchliche Annahme gemacht, dass die einzelnen Studienvarianzen bekannt sind. In der Praxis werden sie aus den Daten geschätzt. Da das zufällige Effekte Modell auf die Studienvarianzen bedingt, sind die hergeleiteten Gewichte in der Praxis als „Pseudo-Exakt“ anzusehen.

Die kumulative Verteilungsfunktion  $F_{DSL}(t; \tau^2)$  kann mittels einer linearen Transformation der exakten Verteilung von  $Q$  (Gleichung (2.1)) und der Berechnung von DSL  $\tau^2$  [2] (Abschn. 1.2.2) dargestellt werden:

$$\hat{\tau}_{DSL}^2 = \max \left\{ \frac{Q - (k - 1)}{c}, 0 \right\} \quad (2.3)$$

$$F_{DSL}(t; \tau^2) = F_Q(ct + k - 1, \tau^2) \quad (2.4)$$

Die Konstante  $c$  ist eine Zusammensetzung der Gewichte  $c = \sum w_i - (\sum w_i^2 / \sum w_i)$  wobei  $w_i = 1/\sigma_i$  die inverse Varianz der Studie  $i$  darstellt und  $k$  die Gesamtanzahl der Studien darstellt.

Mit der exakten Gewichtungsfunktion und des Schätzers  $\hat{\tau}_{DSL}^2$  für  $\tau^2$  ergibt sich der gepoolte Gesamteffekt  $Y(\hat{\tau}_{DSL}^2)$  zu:

$$Y(\hat{\tau}_{DSL}^2) = \frac{\sum w_i^* (\hat{\tau}_{DSL}^2) Y_i}{\sum w_i^* (\hat{\tau}_{DSL}^2)}$$

seine Varianz ist:

$$\text{Var} \left[ Y(\hat{\tau}_{DSL}^2) \right] = \frac{1}{(\sum w_i^* (\hat{\tau}_{DSL}^2))^2} \sum [w_i^* (\hat{\tau}_{DSL}^2)]^2 (\sigma_i^2 + \hat{\tau}_{DSL}^2)$$

## 2.4 Herleitung der Verteilung von $I^2$ in Abhängigkeit der Verteilung von $\tau^2$

Da  $I^2$  (Abschn. 1.3.2) eine Funktion der Cochran's  $Q$ -Verteilungsfunktion (Abschn. 1.3.1) ist, lässt sich auch die Verteilung von  $I^2$  unter Berücksichtigung der Variabilität von  $\tau^2$  entsprechend herleiten. Die Verteilung von  $I^2$  lautet:

$$I^2 = 100 \frac{Q - (k - 1)}{Q}$$

Das  $I^2$  als Verteilung in Abhängigkeit der Verteilung von  $\tau^2$  mit der exakten Verteilung von  $F_Q$  ergibt:

$$F_{I^2}(x; \tau^2) = F_Q \left( \frac{k - 1}{1 - \frac{x}{100}}; \tau^2 \right)$$

Da  $\hat{\tau}^2$  in der Praxis nur für positive Werte betrachtet wird, findet die Integration für den Erwartungswert und der Varianz von  $I^2$  nur im positiven Bereich statt, so dass das  $I^2$  nur in trunkierter Form betrachtet wird.

Der Erwartungswert für  $I^2$  ist dann:

$$\mathbb{E}[I^2] = \int_0^\infty [1 - F_{I^2}(x)] dx$$

und für die Varianz entsprechend:

$$\text{Var}[I^2] = \int_0^\infty 2x[1 - F_{I^2}(x)]dx - (\mathbb{E}[I^2])^2$$

Wie schon Mittlböck und Heinzl [42] sowie Biggerstaff und Jackson [1] anmerken, entspricht die nicht trunkierte exakte Version der  $I^2$ -Verteilung im Falle von Homogenität einer verschobenen  $\chi^2$ -Verteilung. Bei  $\tau^2 = 0$  ist der Erwartungswert  $\mathbb{E}(I^2) = \frac{-2}{k-3}$  für  $k > 3$ .

## 2.5 Berechnung der Konfidenzintervalle für $\tau^2$ und $I^2$ mittels der exakten kumulativen Verteilungsfunktion von Cochran's $Q$

Konfidenzintervalle für  $\tau^2$  und  $I^2$  gemäß Abschnitt 2.4 können numerisch konstruiert werden. Ein zweiseitiges 95%-Konfidenzintervall mit  $\alpha = 5\%$  kann konstruiert werden mit:

$$\begin{aligned} \overline{KI} &= x_1; \quad 1 - \alpha/2 = F_X(x_1; \tau^2) \\ \underline{KI} &= \begin{cases} x_2; & \alpha/2 = F_X(x_2; \tau^2) \\ 0; & \text{sonst} \end{cases} \end{aligned} \quad (2.5)$$

wobei  $F_X$  je nach Wahl entweder mit der Verteilung von  $\tau^2$  nach Gleichung (2.3) oder mit der Verteilung von  $I^2$  nach Gleichung (2.4) entspricht und  $x_i$  die numerische Lösung für die entsprechende Gleichung darstellt. Das Ergebnis ist das Konfidenzintervall  $[\underline{KI}; \overline{KI}]$  für  $I^2$  nach Abschnitt 2.4 bzw.  $\tau^2$ .

## 2.6 Das R-Paket `metaxa`

R [3] ist eine frei verfügbare Programmiersprache für statistisches Rechnen und besitzt derzeit eine steigende Bedeutung als Datenanalyse und Data Mining-

Software [49]. Zusätzliche Entwicklungen von statistischen Funktionen werden in Pakete organisiert, die in R in Form von Bibliotheken geladen werden können. Der Programmiersprache R steht zudem die Möglichkeit zur objektorientierten Programmierung zur Verfügung.

Die genannten Punkte kennzeichnen u.a. Anwenderfreundlichkeit bei Weiterentwicklungen und Benutzerfreundlichkeit bei der Bedienung der Software. Des Weiteren waren diese Faktoren ausschlaggebend, um im Zuge dieser Arbeit, das R-Paket `metaxa` (Meta-Analyse mit exakten Gewichten) [4] mit der Programmiersprache R zu entwickeln.

Folgende statistische Methoden wurden für das Paket `metaxa` objektorientiert implementiert:

- Das feste Effekte Modell
- Das zufällige Effekte Modell
- Das zufällige Effekte Modell mit exakten Gewichten
- Das  $I^2$  nach DerSimonian und Laird
- Iterative Konfidenzintervallberechnung für  $\tau^2$  mit exakter Verteilung von  $Q$
- Iterative Konfidenzintervallberechnung für  $I^2$  mit exakter Verteilung von  $Q$

Die Implementierung wurde mit dem S4-Klassensystem vorgenommen, welches eine formal strengere Programmierung erfordert als mit dem Vorgänger; das S3-Klassensystem.

Die Modelle entsprechen dabei den entwickelten Klassen in R. Als Basisklasse dient das FE-Modell, die davon abgeleitete Klasse das RE-Modell und davon wiederum abgeleitet, das RE-Modell mit exakten Gewichten.

Die Methoden für das  $I^2$  und das  $\tau^2$  wurden für das RE-Modell mit exakten Gewichten überlagert, um ausgehend vom RE-Modell, zusätzlich das  $I^2$  und das  $\tau^2$  mit der iterativen Konfidenzintervallbestimmung zu erhalten.

Zur iterativen Bestimmung der exakten Verteilung von  $Q$  existieren u.a. Algorithmen von Farebrother [46] und Imhof [47]. Beide Algorithmen stehen im R-Paket `CompQuadForm` von P. Lafaye de Micheaux [48] zur Verfügung, welches

grundsätzlich anbietet Gleichungssysteme mit Verteilungsfunktionen von Normalverteilungen in quadratischer Form mit iterativen Algorithmen zu lösen. Für `metaxa` wurde auf den Algorithmus von Farebrother aus `CompQuadForm` zurückgegriffen.

Nachfolgend zeigt der Beispielcode die Verwendung des R-Pakets `metaxa`. Zu Beginn steht im INPUT-Fenster wie aus der Programmbibliothek `metaxa` geladen wird. Im Anschluss erfolgt die Eingabe von Studieneffektschätzern und deren Standardabweichungen aus dem Diuretika und Präeklampsie Daten [50] in Form von Vektoren. Diese dienen als Initialisierung des `exact.obj`-Objekts der Klasse `ExactRE` (das RE-Modell mit exakten Gewichten). Im interaktiven Modus von R wird automatisch die `tt print`-Methode aufgerufen, wenn nur das Objekt in die Kommandozeile eingegeben wird.

Das OUTPUT-Fenster zeigt die Ausgabe des `exact.obj`-Objekts. Aufgelistet sind die Effekte und die Standardabweichungen der einzelnen Studien sowie der prozentuale Anteil der Gewichte vom FE-Modell, vom RE-Modell und vom RE-Modell mit exakten Gewichten.

Im Anschluss erfolgt der gepoolte Effektschätzer, das Konfidenzintervall, der  $z$ -Score und der  $p$ -Wert für jedes der genannten Modelle.

Für die Heterogenität sind zum einen die Teststatistik und der  $p$ -Wert von Cochran's  $Q$  angezeigt, zum anderen die  $I^2$  Werte mit Konfidenzintervalle für die jeweiligen RE-Modelle sowie Informationen über die Gewichtungsmethode und das Maß der Effekte.

Diese und weitere Informationen lassen sich komplett, in Form einer Liste, oder separat in Form von Methodenaufrufen, abrufen. So können z.B. Gewichte in Prozent oder in der ursprünglichen Berechnung abgerufen werden. Ergebnisse aus Listen können in R leicht weiterverarbeitet werden.



## INPUT

```

#load metaxa package
library(metaxa)

#diuretics and preeclampsia data
beta  <- c( 0.042, -0.924, -1.122, -1.473, -1.391, -0.297, -0.262, 1.089, 0.135)
sigma2 <- c(0.160, 0.118, 0.178, 0.299, 0.114, 0.015, 0.121, 0.686, 0.068)
sigma  <- sqrt(sigma2)

#initialization of S4 object
exact.obj <- ExactRE(beta, sigma, effectName = 'beta')

#command to print output
exact.obj

```

## OUTPUT

	Effects	Se	%W FE	%W RE	%W XCT RE
Trial1	0.042	0.4000	5.06	10.66	8.77
Trial2	-0.924	0.3435	6.86	11.94	10.46
Trial3	-1.122	0.4219	4.55	10.19	8.22
Trial4	-1.473	0.5468	2.71	7.86	5.87
Trial5	-1.391	0.3376	7.10	12.08	10.66
Trial6	-0.297	0.1225	53.96	16.95	28.54
Trial7	-0.262	0.3479	6.69	11.84	10.31
Trial8	1.089	0.8283	1.18	4.54	3.16
Trial9	0.135	0.2608	11.90	13.94	13.99

	Effect	CI	z.score	p.value
Fixed Effect:	-0.39976	[-0.576, -0.223]	-4.4435	8.8489e-06
Random Effect:	-0.51708	[-0.917, -0.117]	-2.5336	1.1290e-02
Exact Random Effect:	-0.47850	[-0.806, -0.151]	-2.8605	4.2293e-03

Heterogeneity:

qStatistic	df	p.value
27.25	8	0.0006406

I<sup>2</sup> Exact: 59.85 [0%, 88.47%]  
I<sup>2</sup>: 64.00 [40.27%, 85.57%]  
CI level: 0.95  
Weights method: inverse-variance  
Measure : beta

# 3 Simulationsstudie I – Charakteristik des Meta-Analyse Modells für zufällige Effekte mit exakten Gewichten

## 3.1 Einführung

Das Ziel dieser Simulationsstudie ist es, die Validität des RE-Modells mit exakten Gewichten zu zeigen. Hierzu wird das exakte Meta-Analyse Modell mit dem klassischen DSL für RE [2] hinsichtlich des  $\alpha$ -Fehlers und der Teststärke verglichen.

Des Weiteren werden Simulationen zum exakten  $I^2$  sowie zum  $\tau^2$  mit Konfidenzintervall betrachtet. Beide Simulationen beruhen auf dem Simulationsdesign von Mittlböck und Heinzl [42]. Diese wurde hinsichtlich der Zielsetzung angepasst und erweitert. Im Anschluss wird anhand eines realen Datensatzes auf Besonderheiten hinsichtlich der allgemeinen Charakteristik des exakten RE-Modells eingegangen.

## 3.2 Simulationsaufbau

Grundsätzlich liegt der Fokus des Simulationsaufbaus auf die Beeinflussung der Cochran's  $Q$ -Teststatistik. Sie dient u.a. der Abschätzung der Heterogenität zwischen den Studien (Abschn. 1.3.1) und ihre Varianz führt zu dem Momentenschätzer  $\tau^2$ , der beispielsweise mithilfe des DSL-Ansatzes geschätzt werden kann und in die Berechnung des RE-Modells und exakten RE-Modells einfließt.

Nach Hardy and Thompson [37] gibt es bei der Meta-Analyse drei Haupteigenschaften die die Cochran's  $Q$ -Teststatistik beeinflussen:

- Das Ausmaß der Heterogenität zwischen den Studien  $\tau^2$

- Die Anzahl der Studien  $k$  die in der Meta-Analyse eingeschlossen sind
- Die Gewichtung  $w_i$  für jede der  $i$  Studien

Basierend auf diesen drei Eigenschaften ist der Simulationsaufbau für eine (indirekte) Beeinflussung der  $Q$ -Teststatistik aufgebaut.

Mittels Vierfelder-Tafeln (Tab. 3.1) (entsprechend einer Fallgruppe und Kontrollgruppe) sind für jede Studie dichotome Studieneffekte, quantifiziert mit einem wahren log Odds-Ratio  $\beta$  und einer Zwischenstudienvarianz  $\tau^2$ , realisiert worden. Das log Odds-Ratio für jeden Studieneffekt  $\beta_i$  ist aus einer Normalverteilung mit  $N(\beta, \tau^2)$  gezogen worden. Unter Annahme einer binomialverteilten Zielwahrscheinlichkeit für (Fall-)Gruppe 1 mit  $p_1$  und (Kontroll-)Gruppe 2 mit  $p_2$  einer Studie, kann die Zielwahrscheinlichkeit  $p_2$  mithilfe der Gleichung  $p_2 = p_1 / (-\exp(\beta_i)p_1 + p_1 + \exp(\beta_i))$  modelliert werden.

Im Anschluss erfolgt die Festlegung der Zufallszahlen für ein Ereignis ( $a$  und  $b$ ) der jeweiligen Gruppen (Tab. 3.1). Hierfür wird mit festgelegten Gruppengrößen  $n_1$  und  $n_2$  aus den entsprechenden Binomialverteilungen mit  $B(n_1, p_1)$  und  $B(n_2, p_2)$  gezogen. Die Anzahl der ereignisfreien Beobachtungen ergeben sich somit aus  $c = (n_1 - a)$  und  $d = (n_2 - b)$ . Ist eine Zelle mit 0 besetzt, wurde 0,5 zu allen vier Zellen einer Tafel addiert um eine mögliche Division durch 0 zu vermeiden. Das log Odds-Ratio und die Varianz pro Tafel bzw. pro Studie wurde geschätzt mit  $\hat{\beta}_i = \ln((a_i d_i) / (b_i c_i))$  und  $\text{Var}(\hat{\beta}_i) = 1/a_i + 1/b_i + 1/c_i + 1/d_i$ .

**Tabelle 3.1:** Beispiel für die Zusammensetzung einer Vierfelder-Tafel pro Studie

	Fallgruppe	Kontrollgruppe
Ereignis	$a$	$b$
Ereignisfrei	$c$	$d$
$\Sigma$	$n_1$	$n_2$

Alle Simulationen wurden mit 2000 Iterationen durchgeführt. Die Initialisierungswerte aller Simulationen sind in folgender Tabelle aufgelistet:

**Tabelle 3.2:** Zusammenstellung der Parameter für die drei Simulationsszenarien

	Szenario I	Szenario II	Szenario III
Odds-Ratio	1; 1,1; 1,2; 1,5	1; 1,1; 1,2; 1,5	1; 1,1; 1,2; 1,5
Heterogenität $\tau^2$	0,00...0,50 mit* 0,01	0,00...0,50 mit* 0,01	0,00...0,50 mit* 0,01
Anzahl Studien $k$	6; 8; 10; 12; 15; 20; 50; 100	6; 8; 10; 12; 15	6; 8; 10; 12; 15; 20; 50; 100
Stichprobengröße $n$ pro Gruppe	20; 50; 100; 300; 500	20; 50	20; 50; 100; 300; 500

\*Schrittlänge

Die Stichprobengröße  $n$  pro Gruppe ist ein Initialwert mit unterschiedlicher Auswirkung auf die Gruppengröße in jedem Szenario. Für Szenario I findet einfach eine Wiederholung von  $n$  pro Studie statt. Für das Szenario II wird eine kontinuierliche Addition von  $n$  pro Studie vorgegeben und für Szenario III findet eine zusätzliche Gewichtung nach folgendem Algorithmus statt:

$$\begin{aligned}
 x &= \lfloor k/2 \rfloor \\
 y &= \lfloor k/6 \rfloor \\
 z &= k - x - y \\
 f(y, n) &= \begin{cases} n \cdot 5; & \text{wenn } y \geq 1 \\ 0; & \text{sonst} \end{cases} \\
 vec_1 &= \text{repliziere } n \quad x \text{ mal} \\
 vec_2 &= \text{repliziere } f(y, n) \quad y \text{ mal} \\
 vec_3 &= \text{repliziere } (n \cdot 10) \quad z \text{ mal} \\
 n_k &= (vec_1, vec_2, vec_3)
 \end{aligned}$$

um ein Ungleichgewicht bei der Stichprobengröße für alle Studien zu simulieren. Als Beispiel werden die Stichprobengrößen pro Gruppe für das jeweilige Szenario mit  $k = 8$  Studien aufgelistet (Tab. 3.3).

**Tabelle 3.3:** Studienspezifische Stichprobengröße pro Gruppe und den drei Simulationsszenarien mit  $n = 20$  und über insgesamt  $k = 8$  Studien für die drei Simulationsszenarien.

	Studie 1	Studie 2	Studie 3	Studie 4	Studie 5	Studie 6	Studie 7	Studie 8
Szenario I	20	20	20	20	20	20	20	20
Szenario II	20	40	60	80	100	120	140	160
Szenario III	20	20	20	20	100	200	200	200

Wegen der sich daraus ergebenden großen Anzahl von Simulationen werden die Ergebnisse im folgenden Kapitel nur ausschnittsweise präsentiert. Die Auswahl wurde dabei so vollzogen, dass Schlussfolgerungen auf die gesamte Simulationsstudie übertragen werden können und auftretende Besonderheiten geschildert werden.

## 3.3 Ergebnisse

### 3.3.1 $\alpha$ -Fehler

Da  $\tau^2$  einen direkten Einfluss auf die Cochran's  $Q$ -Teststatistik hat, ist es der Hauptparameter zum Testen des empirischen  $\alpha$ -Fehlers. Bei gleichbleibendem neutralen Odds-Ratio von 1 und steigendem  $\tau^2$  unter Einhaltung des  $\alpha$ -Fehlers mit  $\alpha = 5\%$  gibt die Simulation Aufschluss über die Reliabilität der beiden Modelle.

In Tabelle 3.4 zeigt den empirischen  $\alpha$ -Fehler mit steigender Studienanzahl und steigendem  $\tau^2$  und verschiedenen Simulationsszenarien auf. Für Szenario II gibt es im exakten RE-Modell fehlende Werte für große Initialisierungswerte hinsichtlich der Anzahl von Studien. Diese sind aufgrund von Konvergenzschwierigkeiten beim exakten RE-Modell entstanden.

Betrachtet man die Generierung der Stichproben pro Studie für diesen Fall (Abschn. 3.2) ergibt sich, mit wachsender Studienanzahl, ein starkes Ungleichgewicht der Stichprobengrößen im Bezug auf Studien, welches in der Realität in

diesem Ausmaß äußerst selten vorkommen sollte.

Grundsätzlich zeigt der Vergleich des empirischen  $\alpha$ -Fehlers keine nennenswert großen Differenzen zwischen dem RE-Modell und dem exakten RE-Modell. Bei einem  $\tau^2 = 0$  (keine Zwischenstudienvarianz), zeigen die beiden Modelle kein Überschreiten des  $\alpha$ -Niveaus. Jedoch schöpfen beide Modelle das Niveau nicht voll aus und es ist keine klare Überlegenheit eines bestimmten Modells zu erkennen.

Bei einer Zunahme von  $\tau^2$ , ist bei beiden Modellen ein Überschreiten des Niveaus in allen Simulationsszenarien feststellbar mit Ausnahme von Simulationsszenario I mit einer hohen Anzahl Studien von 100 bei dem das Niveau auch für große  $\tau^2$  Werte eingehalten wird.

Eine Zunahme hinsichtlich der Anzahl der Studien hat zumeist einen stabilisierenden Effekt auf die Einhaltung des  $\alpha$ -Niveaus. Jedoch wird, insbesondere bei einer hohen Studienanzahl, das Niveau für Simulationsszenario I und III teilweise wieder unterschritten.

Hinsichtlich der Simulationsszenarien gibt es Unterschiede beim maximalen Bereich des  $\alpha$ -Fehlers. Während er bei Szenario I (spaltenweise) von ca. 0,3 bis 0,6 liegt, ist der maximale Bereich für die beiden anderen Simulationen höher und liegt maximal zwischen 0,114 und 0,28 bei Simulationsszenario III.

**Tabelle 3.4:**  $\alpha$ -Fehler mit nominellem  $\alpha = 5\%$  des klassischen zufälligen Effekte Modells (RE) und des zufälligen Effekte Modells mit exakten Gewichten (exRE) für variierende Heterogenität  $\tau^2$  mit Stichprobengröße  $n = 20$  pro Gruppe und verschiedenen Simulationsszenarien. Wegen Konvergenzschwierigkeiten existieren keine Einträge für das Szenario II ab einer Anzahl von  $k = 50$  Studien.

Anzahl Studien		8		10		15		50		100	
Simulationsszenario	$\tau^2$	RE	exRE	RE	exRE	RE	exRE	RE	exRE	RE	exRE
Szenario I	0	0,036	0,036	0,036	0,038	0,038	0,038	0,042	0,042	0,043	0,043
	0,05	0,050	0,051	0,038	0,039	0,041	0,041	0,044	0,044	0,050	0,050
	0,1	0,052	0,052	0,049	0,049	0,044	0,042	0,047	0,046	0,047	0,048
	0,15	0,051	0,049	0,048	0,048	0,050	0,050	0,052	0,052	0,037	0,036
	0,2	0,058	0,057	0,059	0,059	0,052	0,051	0,042	0,042	0,042	0,041
Szenario II	0	0,046	0,047	0,038	0,038	0,034	0,034	—	—	—	—
	0,05	0,071	0,070	0,074	0,072	0,067	0,066	—	—	—	—
	0,1	0,089	0,089	0,082	0,078	0,086	0,086	—	—	—	—
	0,15	0,077	0,075	0,078	0,079	0,074	0,073	—	—	—	—
	0,2	0,102	0,099	0,086	0,084	0,080	0,077	—	—	—	—
Szenario III	0	0,028	0,029	0,038	0,037	0,040	0,036	0,040	0,042	0,036	0,036
	0,05	0,087	0,083	0,069	0,069	0,065	0,065	0,061	0,059	0,057	0,057
	0,1	0,101	0,103	0,092	0,090	0,092	0,091	0,069	0,065	0,059	0,059
	0,15	0,110	0,111	0,101	0,098	0,087	0,090	0,055	0,057	0,054	0,054
	0,2	0,114	0,113	0,108	0,112	0,080	0,082	0,066	0,066	0,051	0,051

### 3.3.2 Teststärke

Wie auch schon für das  $\alpha$ -Niveau (Abschn. 3.3.1) existieren keine nennenswerten Differenzen zwischen dem RE-Modell und dem exakten RE-Modell für die Teststärke. In Tabelle 3.5 ist zu erkennen, dass die Teststärke grundsätzlich bei einer steigenden Anzahl von Studien ansteigt. Dieses Resultat entspricht den Erwartungen für normalverteilte Studiendaten [42].

Des Weiteren erfolgt auch mit der Zunahme des Odds-Ratio in allen Simulationsszenarien ein erwarteter monotoner Anstieg der Teststärke. Dabei steigt die Teststärke in Szenario I langsamer an als in den beiden anderen Szenarien.

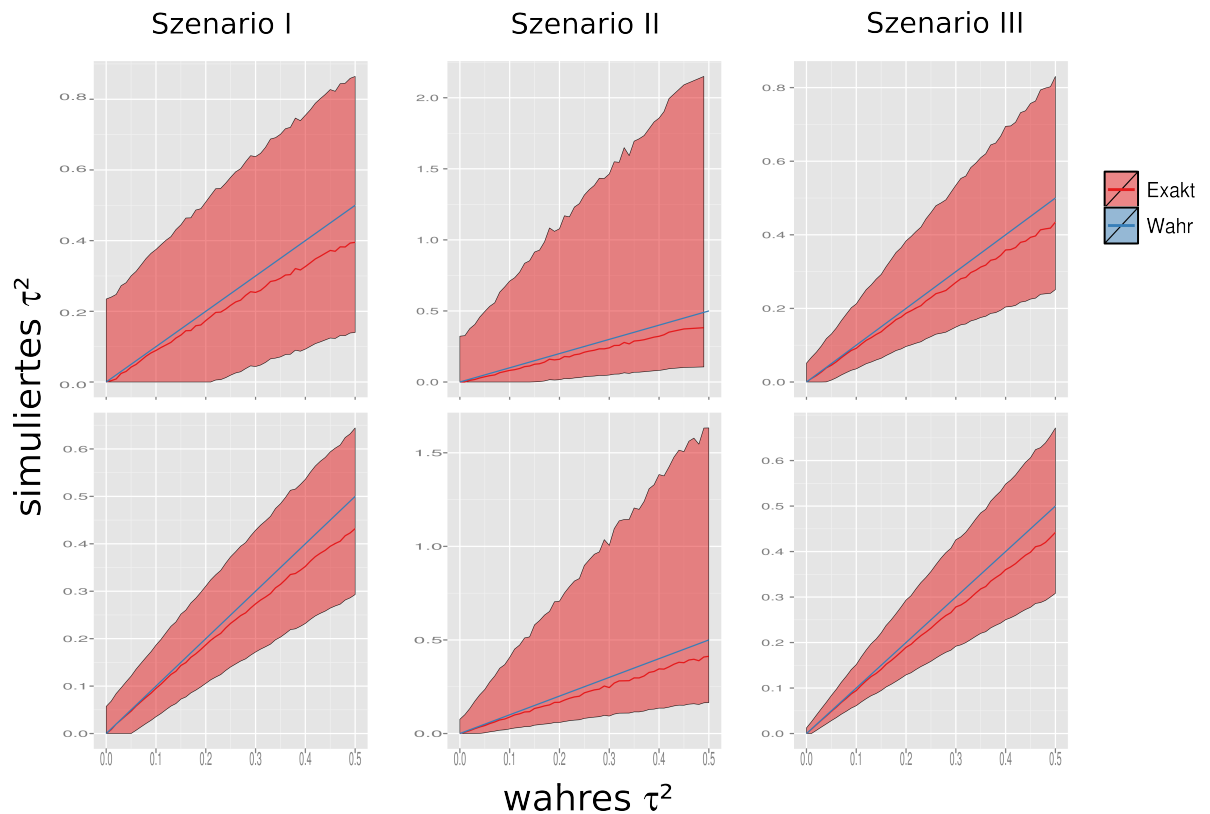
**Tabelle 3.5:** Teststärke für das zufälligen Effekte Modell (RE) und das zufällige Effekte Modell mit exakten Gewichten (exRE) für variierende Heterogenität  $\tau^2$  und Odds-Ratio (OR) mit Stichprobengröße  $n = 20$  pro Gruppe und verschiedenen Simulationsszenarien. Wegen Konvergenzschwierigkeiten existieren keine Einträge für das Szenario II ab einer Anzahl von  $k = 50$  Studien.

Anzahl Studien		8		10		15		50		100			
Simulationsszenario	$\tau^2$	OR	RE	exRE	RE	exRE	RE	exRE	RE	exRE	RE	exRE	
Szenario I	0,05	1,1	0,063	0,063	0,053	0,053	0,072	0,072	0,137	0,136	0,270	0,270	
		1,2	0,105	0,105	0,105	0,106	0,106	0,157	0,157	0,453	0,454	0,735	0,734
		1,5	0,344	0,343	0,421	0,414	0,596	0,594	0,982	0,982	1,000	1,000	1,000
	0,1	1,1	0,059	0,059	0,066	0,065	0,075	0,072	0,149	0,150	0,255	0,255	
		1,2	0,097	0,097	0,113	0,113	0,158	0,158	0,418	0,417	0,700	0,698	
		1,5	0,347	0,346	0,377	0,375	0,562	0,555	0,974	0,974	1,000	1,000	
Szenario II	0,05	1,1	0,135	0,136	0,145	0,143	0,234	0,232	—	—	—	—	
		1,2	0,290	0,288	0,368	0,364	0,563	0,562	—	—	—	—	
		1,5	0,808	0,802	0,921	0,920	0,995	0,995	—	—	—	—	
	0,1	1,1	0,136	0,139	0,147	0,141	0,181	0,180	—	—	—	—	
		1,2	0,251	0,250	0,294	0,278	0,420	0,410	—	—	—	—	
		1,5	0,710	0,704	0,829	0,826	0,967	0,963	—	—	—	—	
Szenario III	0,05	1,1	0,146	0,145	0,162	0,159	0,199	0,193	0,407	0,404	0,644	0,643	
		1,2	0,306	0,297	0,330	0,327	0,446	0,436	0,875	0,875	0,996	0,995	
		1,5	0,798	0,790	0,873	0,869	0,972	0,972	1,000	1,000	1,000	1,000	
	0,1	1,1	0,128	0,129	0,139	0,136	0,159	0,155	0,297	0,296	0,511	0,510	
		1,2	0,254	0,246	0,266	0,265	0,342	0,344	0,749	0,747	0,966	0,966	
		1,5	0,664	0,650	0,748	0,740	0,908	0,903	1,000	1,000	1,000	1,000	



### 3.3.3 Das $\tau^2$ mit Konfidenzintervalle auf Grundlage der exakten Cochran's $Q$ -Verteilung

Mit der exakten Cochran's  $Q$ -Verteilung [1] ergibt sich die Möglichkeit, direkt Konfidenzintervalle für das  $\tau^2$  nach DSL [2] zu konstruieren (Gleichung (2.5)). In Übereinstimmung mit dem Befund von Thorlund et al. [51] zeigt die Abbildung 3.1 auch für den RE-Schätzer mit exakter Gewichtung stets eine Unterschätzung der wahren Zwischenstudienvarianz  $\tau^2$  auf. Eine Erhöhung der Heterogenität hat zwei Effekte zur Folge: Zum einen nimmt die Stärke der Unterschätzung bezüglich der wahren Heterogenität  $\tau^2$  zu, zum anderen werden auch die Konfidenzintervalle breiter. Des Weiteren ist am häufigsten eine Trunkierung der Konfidenzintervalle in Szenario I zu verzeichnen, sowie im Gegensatz dazu am wenigsten in Szenario III.



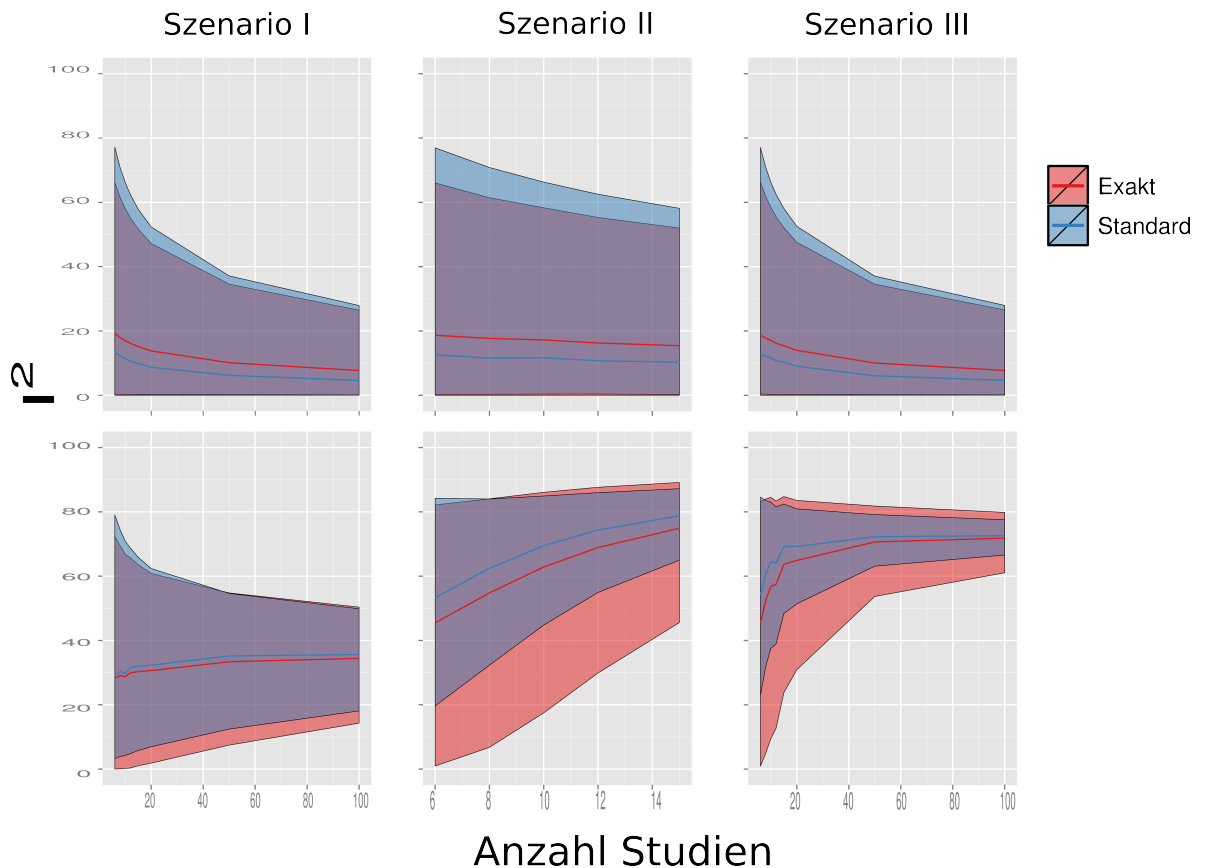
**Abbildung 3.1:** Simulationsergebnisse für  $\tau^2$  mit 95%-Konfidenzintervall für die drei Simulationsszenarien. Die erste Reihe zeigt die Ergebnisse für  $k = 50$  Studien ( $k = 8$  für Szenario II) und einer initialen Stichprobengröße von  $n = 20$  pro Gruppe. Die zweite Reihe zeigt die Ergebnisse für  $k = 100$  Studien ( $k = 10$  für Szenario II) und einer initialen Stichprobengröße von  $n = 50$  pro Gruppe. Die  $x$ -Achse und die  $y$ -Achse geben das wahre  $\tau^2$  auf der Hauptdiagonalen in blau und das simulierte  $\tau^2$  in rot wieder. Die rote Fläche stellt das entsprechende 95%-Konfidenzintervall für das exakte zufällige Effekte Modell dar.

### 3.3.4 Vergleich zwischen $I^2$ und das exakte $I^2$ mit Konfidenzintervallen mit der exakten Cochran's $Q$ -Verteilung

Die Konfidenzintervalle für das trunkeerte exakte  $I^2$  können, wie schon bei der Konstruktion der  $\tau^2$  Konfidenzintervalle (2.5), mithilfe der Cochran's  $Q$ -Verteilung hergeleitet werden.

Die Abbildung 3.2 in der ersten Reihe zeigt, dass mit dem  $I^2$  die wahre Heterogenität überschätzt wird, wenn keine Heterogenität vorhanden ist (beispielsweise bei einem  $\tau^2 = 0$ ).

Der  $I^2$  Punktschätzer für das exakte Modell ist stets höher als beim Standard-Ansatz. Die Konfidenzintervalle bleiben im Vergleich immer schmaler. Die Ergebnisse drehen sich für existierende Heterogenität in etwa um. Dies ist in der zweiten Reihe für Abbildung in 3.2 bei einem  $\tau^2 = 0,1$  feststellbar. Der  $I^2$  Punktschätzer ist für den Standard-Ansatz stets höher als der Punktschätzer für das exakte RE-Modell und grundsätzlich sind die Konfidenzintervalle für das exakte RE-Modell breiter als die berechneten Konfidenzintervalle für den Standard-Ansatz. Interessanterweise sind die oberen Grenzen der Konfidenzintervalle sehr ähnlich, jedoch unterscheiden sie sich grundlegend bei den unteren Grenzen.



**Abbildung 3.2:** Simulationsergebnisse für den Vergleich zwischen dem  $I^2$  (Standard) und dem exakten  $I^2$  (Exakt) mit 95%-Konfidenzintervall für die drei Simulationsszenarien mit variierender Anzahl der Studien  $k$  ( $x$ -Achse) sowie initialer Stichprobengröße  $n = 50$  pro Gruppe. Die erste Reihe der Grafik zeigt die Ergebnisse für ein  $\tau^2 = 0$ , und die zweite Reihe zeigt die Ergebnisse für ein  $\tau^2 = 0,1$ . Für das zufällige Effektmodell mit exakten Gewichten steht die rote Linie und die rot schattierte Fläche entsprechend für den Punktschätzer für das exakte  $I^2$  mit 95%-Konfidenzintervall. Für das DSL zufällige Effektmodell steht die blaue Linie und die blau schattierte Fläche entsprechend für den Punktschätzer für das Standard  $I^2$  mit 95%-Konfidenzintervall.

### 3.4 Interpretation der Ergebnisse

Simulationsstudie I zeigt auf, dass das exakte RE-Modell im Vergleich zum DSL RE-Modell im Rahmen der Initialisierungswerte  $i$ . Allg. valide Ergebnisse produziert. Große Abweichungen vom  $\alpha$ -Niveau und der Teststärke zwischen den

beiden Modellen treten nicht auf. Für bestimmte Initialisierungswerte (Tab. 3.4 und Tab. 3.5), die in der Realität i. Allg. nicht auftreten, ergeben sich Konvergenzschwierigkeiten für den Algorithmus zur Berechnung der exakten Verteilung von Cochran's  $Q$ . Damit kann das exakte RE-Modell zumindest als gleichwertig hinsichtlich des  $\alpha$ -Niveaus und der Teststärke gegenüber der etablierten Standardmethode, dem DSL RE-Modell, betrachtet werden.

Die Konstruktion der Konfidenzintervalle für das  $\tau^2$  bietet einen kohärenten Ansatz für die DSL Methode gegenüber anderen Verfahren [52], da der sowohl der gleiche Punktschätzer  $\tau^2$ , als auch die gleiche Variabilität von  $\tau^2$  zur Berechnung der Gewichte für das exakte RE-Modell und für die Konstruktion der Konfidenzintervalle von  $\tau^2$  benutzt werden.

Das exakte  $I^2$  zeigt gegenüber dem  $I^2$ , bei einem  $\tau^2 = 0$ , schmalere Konfidenzintervalle an. Hinzu kommt, dass die Konfidenzintervalle 0 wesentlich häufiger beinhalten bzw. die Konfidenzintervalle für  $I^2$  diese nur knapp verfehlen (in Abb. 3.2 nicht erkennbar). Jedoch zeigt das exakte  $I^2$  wesentlich breitere Konfidenzintervalle für ein  $\tau^2 > 0$ . Obwohl die Ergebnisse durch die Trunkierung von  $I^2$  auf positive Werte vergleichbar erscheinen hat das exakte  $I^2$  den Nachteil wenn vollständige Homogenität vorliegt. Denn die Verteilung von  $I^2$  hat bei einem  $\tau^2 = 0$  den Erwartungswert  $\mathbb{E}(I^2) = \frac{-2}{k-3}$  für  $k > 3$ . Das hat den Effekt, dass bei der Integration der trunkeerten Verteilung das  $I^2$  stets Werte größer 0 annimmt (diese stehen in Abhängigkeit zu  $k$ ), da der negative Teil der Verteilung nicht berücksichtigt wurde. Obwohl in solchen Fällen das Konfidenzintervall für  $I^2$  weiterhin die 0 beinhaltet, wäre es wünschenswert, das auch das  $I^2$  in Fällen von absoluter Homogenität 0 anzeigt.

# 4 Simulationsstudie II – Einfluss der Gewichte des Modells für zufällige Effekte mit exakten Gewichten

## 4.1 Einführung

Das Ziel dieser Simulationsstudie ist es, den Einfluss der unterschiedlichen Gewichtung des RE-Modells nach DSL und des RE-Modells mit exakten Gewichten Effekte herauszuarbeiten und zu vergleichen.

Hierzu wird der Fall simuliert, dass eine einzelne Studie mit geringer Varianz sowie einer Menge von Studien mit größer Varianz in einer Meta-Analyse auftreten. Dabei dominiert die einzelne „große“ Studie die „kleineren“ Studien hinsichtlich eines höheren Gewichts. Des Weiteren unterscheidet sich die Einzelstudie durch einen stark abweichenden Effektgröße gegenüber den anderen Studien. Mit der einzelnen großen Studie kann dann gut die Auswirkung der Variabilität von  $\tau^2$  auf die Gewichtung der Studie abgeschätzt werden.

## 4.2 Simulationsaufbau

Für die Simulation wurden Daten sowohl unter der Nullhypothese (kein Effekt) als auch unter der Alternativhypothese (Effekt sowohl in der Einzelstudie als auch in den kleinen Studien) generiert. Dabei gilt für die Alternativhypothese eine Effektgröße von  $\beta_1 = 0,2$  für die kleinen Studien und eine Effektgröße von  $\beta_2 = 2$  für die Einzelstudie.

Die Anzahl der kleinen Studien variiert zwischen 3 und 10. Studieneffekte für die kleinen Studien  $\hat{\beta}_1$  sind aus einer Normalverteilung mit  $N(\beta_1, \sigma_1^2)$  gezogen worden. Die Varianzen  $\sigma_1^2$  zum Erhalt der Studieneffekte wurden mit den Werten  $0,1^2; 0,2^2; \dots; 0,9^2; 1; 1,5^2; 2^2; 4^2$  variiert. Den Studieneffekten wurden wiederum Varianzen zugeordnet. Diese wurde aus einer Gleichverteilung mit  $U(0,1;0,4)$  gezogen.

Die Generierung der Einzelstudie erfolgt nach dem gleichen Schema. Der Studieneffekt wurde aus einer Normalverteilung mit  $N(\beta_2; 0, 2^2)$  gezogen. Die Varianz des Studieneffekts wurde aus einer Gleichverteilung mit  $U(0, 001; 0, 2)$  gezogen.

Alle Simulationen wurden mit 10000 Iterationen durchgeführt.

## 4.3 Ergebnisse

### 4.3.1 $\alpha$ -Fehler

In Tabelle 4.1 weist das exakte RE-Modell für alle kleinen Studienanzahlen ohne Ausnahme einen niedrigeren empirischen  $\alpha$ -Fehler auf als das klassische DSL RE-Modell (bei einem nominellen  $\alpha = 0, 05$ ).

Jedoch steigt für beide Ansätze der empirische Fehler mit steigender Varianz bezüglich der kleinen Studien, bei festgesetzter Varianz der großen Studie. Die kleinen Studien streuen in diesem Fall breiter um  $H_0 = 0$ , so dass der  $\alpha$ -Fehler ebenfalls ansteigt. Auch hier zeigt das exakte Modell eine Überlegenheit, da der  $\alpha$ -Fehler weniger stark ansteigt als für das RE-Modell. Dieser Effekt ist auch für eine steigende Anzahl von kleinen Studien zu verzeichnen.

Das Mittel des Gewichts für die einzelne große Studie sinkt für beide Modelle, sowohl bei wachsender Anzahl, als auch bei wachsender Varianz der kleinen Studien.

**Tabelle 4.1:** Empirischer  $\alpha$ -Fehler für das zufällige Effekte Modell ( $\alpha$  RE) und das zufällige Effekte Modell mit exakten Gewichten ( $\alpha$  exRE) für Simulation II. Anzahl kleiner Studien variiert zwischen 3 und 10. Effektgrößen  $\beta_1$  und  $\beta_2$  sind 0. Varianz kleiner Studien  $\sigma_1^2$  ist  $0,3^2; 0,5^2; 0,6^2$  und  $0,8^2$  und Varianz für große Studie ist  $\sigma_2^2 = 0,2^2$ . Varianzen für Effekte kleiner Studien sind aus Gleichverteilung mit  $U(0,1;0,4)$  gezogen. Varianz für große Studie aus Gleichverteilung mit  $U(0,001;0,2)$ . Mittlere Gewichte (Standardabweichung (SD)) sind für große Studie für das zufällige Effekte Modell (Gewicht RE) und das zufällige Effekte Modell mit exakten Gewichten (Gewicht exRE) angegeben.

$\sigma_1^2$	# Kleine Studien	Gewicht RE (SD)	Gewicht exRE (SD)	$\alpha$ RE	$\alpha$ exRE
0,3 <sup>2</sup>	3	47,59 (18,28)	46,49 (17,69)	0,0336	0,0323
	4	41,51 (19,14)	40,17 (18,32)	0,0370	0,0352
	5	37,02 (19,06)	35,60 (18,06)	0,0305	0,0283
	6	33,37 (18,68)	31,88 (17,59)	0,0314	0,0297
	7	30,56 (18,49)	29,01 (17,25)	0,0341	0,0317
	8	28,11 (18,27)	26,54 (16,85)	0,0336	0,0297
	9	26,57 (18,19)	24,93 (16,66)	0,0332	0,0310
	10	24,61 (17,48)	23,02 (15,89)	0,0299	0,0262
0,5 <sup>2</sup>	3	44,00 (17,09)	45,79 (17,76)	0,0483	0,0454
	4	37,65 (16,76)	39,36 (17,44)	0,0500	0,0487
	5	32,78 (16,12)	34,42 (16,98)	0,0493	0,0477
	6	29,36 (15,86)	30,95 (16,79)	0,0534	0,0506
	7	26,36 (15,31)	27,68 (16,06)	0,0507	0,0482
	8	24,30 (14,97)	25,60 (15,83)	0,0537	0,0529
	9	22,42 (14,44)	23,67 (15,28)	0,0551	0,0531
	10	20,61 (13,91)	21,76 (14,78)	0,0567	0,0554
0,6 <sup>2</sup>	3	41,74 (16,11)	45,37 (17,50)	0,0572	0,0542
	4	35,07 (15,44)	38,55 (17,02)	0,0575	0,0539
	5	30,10 (14,67)	33,42 (16,52)	0,0636	0,0589
	6	26,72 (14,12)	30,02 (16,09)	0,0679	0,0643
	7	23,57 (12,88)	26,80 (15,20)	0,0656	0,0624
	8	21,13 (12,16)	24,32 (14,85)	0,0643	0,0597
	9	19,36 (11,95)	22,30 (14,32)	0,0667	0,0619
	10	17,52 (10,75)	20,42 (13,61)	0,0672	0,0618
0,8 <sup>2</sup>	3	37,35 (13,56)	43,64 (16,30)	0,0566	0,0484
	4	30,30 (12,19)	36,53 (15,64)	0,0697	0,0591
	5	25,34 (10,83)	31,27 (14,98)	0,0783	0,0651
	6	21,85 (9,83)	27,29 (14,10)	0,0734	0,0627
	7	19,01 (8,84)	23,78 (12,92)	0,0816	0,0693
	8	16,63 (7,60)	20,93 (11,86)	0,0774	0,0696
	9	15,02 (7,07)	18,95 (11,24)	0,0762	0,0679
	10	13,43 (5,94)	16,99 (10,24)	0,0724	0,0646



### 4.3.2 Teststärke

In Tabelle 4.2 ist die Teststärke für den Fall aufgezeigt, bei der einer einzelnen großen Studie eine Effektgröße von  $\beta_2 = 2$  zugeordnet ist. Dabei werden zwei unterschiedliche Szenarien für die kleinen Studien unterschieden. Im oberen und unteren Teil der Tabelle sind die Ergebnisse für den Fall aufgelistet, wenn den kleinen Studien die Effekte  $\beta_1 = 0$  und  $0,2$  zugeordnet werden. Die Simulation ist im Sinne zu verstehen, dass viele (kleine) Studien mit geringer Teststärke ohne Effekt oder mit kleinem Effekt ( $\beta_1$  ist  $0$  und  $0,2$ ) vorhanden sind. Dies kann beispielsweise durch ein qualitativ "schlechtes" Studiendesign für die kleinen Studien hervorgerufen werden, sowie wenn eine (große) Studie mit hoher Teststärke und großen Effekt vorhanden ist ( $\beta_2 = 2$ ). Dies kann beispielsweise hervorgerufen werden durch ein qualitativ "hochwertiges" Studiendesign, welche Evidenz für die Alternativhypothese zeigt.

Die wichtigste Erkenntnis dieser Simulationsstudie ist, dass die Teststärke mit dem Ansatz der exakten Gewichte immer größer ist als die Teststärke mit der DSL-Methode. Dies spiegelt sich auch im exakten Gewicht wider, welches stets größer ist als das Gewicht für die große Studie beim DSL-Ansatz. Auffällig ist auch eine generell höhere Standardabweichung der Gewichte für das exakte RE-Modell im Bezug zum RE-Modell.

**Tabelle 4.2:** Teststärke für das zufällige Effekte Modell (p RE) und das zufällige Effekte Modell mit exakten Gewichten (p exRE) der Simulation II. Anzahl der kleinen Studien beträgt 3, 5 und 10. Die einzelne große Studie ist simuliert mit einer Effektgröße  $\beta_2 = 2$  und Varianz  $\sigma_2^2 = 0,2^2$ . Effektgrößen der kleinen Studien sind gewählt mit  $\beta_1 = 0$  sowie ein  $\beta_1 = 0,2$  mit Varianzen  $\sigma_1^2 = 0,3^2; 0,5^2; 0,6^2$ ; und  $0,8^2$ . Verteilung der  $\hat{\beta}_k$  ist aus einer Normalverteilung gezogen. Varianzen der Studieneffekte stammen aus einer Gleichverteilung mit  $U(0,1;0,4)$  für die kleinen Studien und  $U(0,001;0,2)$  für die große Studie. Mittlere Gewichte (Standardabweichung (SD) in Klammern) sind für die große Studie für das zufällige Effekte Modell (Gewicht RE) und das zufällige Effekte Modell mit exakten Gewichten (Gewicht exRE) angegeben.

$\beta_1$	$\sigma_1^2$	# Kleine Studien	Gewicht RE (SD)	Gewicht exRE (SD)	p RE	p exRE
0	$0,3^2$	3	26,91 (1,04)	37,59 (11,45)	0,0043	0,0132
		5	18,41 (0,79)	23,33 (5,55)	0,0059	0,0117
		10	10,61 (0,60)	11,92 (1,39)	0,0072	0,0101
	$0,5^2$	3	26,87 (1,20)	37,09 (11,17)	0,0405	0,0698
		5	18,31 (0,82)	23,19 (5,62)	0,0286	0,0477
		10	10,41 (0,54)	11,61 (1,29)	0,0324	0,0415
	$0,6^2$	3	26,91 (1,42)	37,36 (11,47)	0,0651	0,0995
		5	18,25 (0,87)	23,02 (5,63)	0,0473	0,0697
		10	10,30 (0,52)	11,41 (1,26)	0,0394	0,0472
$0,8^2$	3	26,87 (1,76)	37,02 (11,46)	0,1073	0,1410	
	5	18,14 (0,99)	22,77 (5,99)	0,0792	0,1050	
	10	10,11 (0,48)	11,09 (1,19)	0,0583	0,0664	
0,2	$0,3^2$	3	27,37 (1,40)	38,20 (11,87)	0,0636	0,1317
		5	18,82 (1,05)	24,48 (6,76)	0,1012	0,1609
		10	10,95 (0,81)	12,56 (1,89)	0,2694	0,3222
	$0,5^2$	3	27,34 (1,75)	38,05 (12,13)	0,1466	0,2150
		5	18,68 (1,13)	24,16 (6,60)	0,1553	0,2165
		10	10,66 (0,71)	12,11 (1,73)	0,2307	0,2620
	$0,6^2$	3	27,36 (2,03)	37,91 (11,83)	0,1616	0,2266
		5	18,60 (1,14)	23,97 (6,66)	0,1737	0,2261
		10	10,53 (0,68)	11,89 (1,73)	0,2289	0,2540
	$0,8^2$	3	27,34 (2,47)	37,65 (11,79)	0,1987	0,2522
		5	18,43 (1,45)	23,57 (6,78)	0,1880	0,2286
		10	10,26 (0,61)	11,41 (1,62)	0,2030	0,2226

## 4.4 Interpretation der Ergebnisse

Die Simulationsstudie zeigt auf, dass die Gewichte einen Effekt den empirischen  $\alpha$ -Fehler und die Teststärke haben, wenn Heterogenität zwischen den Studien vorhanden ist bzw. eine einzelne große Studie einen anderen, höheren wahren Effekt aufweist als die kleineren Studien. Dies ist beispielsweise der Fall, wenn eine Studie mit hoher Fallzahl und mit einem qualitativ hochwertigen Analyseplan durchgeführt wird, falls zuvor in kleineren älteren Studien mit qualitativ minderwertigem Studiendesign kein klarer Effekt nachgewiesen werden konnte. Der großen Studie wird mithilfe des exakten RE-Modells in allen Simulationsszenarien durchgehend ein höheres Gewicht zugeordnet. Dies hat dann auch Einfluss auf den Gesamteffektschätzer sowie den  $p$ -Wert. Während beide Modelle das  $\alpha$ -Niveau etwa im gleichen Ausmaß einhalten bzw. überschreiten, zeigt das exakte RE-Modell hinsichtlich der Teststärke eine klare Überlegenheit gegenüber dem RE-Modell. Somit ist dem neuen Modell in solchen Situationen der Vorzug zu geben, da insbesondere die Validität der Methode gewährleistet ist bzw. es in Standardsituationen (Kap. 3) zu keinen nennenswerten Abweichungen kommt.

# 5 Simulationsstudie III – Meta-Analysen zur Beurteilung von Abweichungen vom Hardy-Weinberg Gleichgewicht

## 5.1 Einführung

Diese Simulationsstudie befasst sich mit dem Vergleich von statistisch zusammengefassten (gepoolten) Maßzahlen, die eine Abweichung vom Hardy-Weinberg Gleichgewicht quantifizieren können. Anhand der Anwendung dieser Maßzahlen aus der Populationsgenetik wird ein Vergleich zwischen dem RE-Modell und dem exakten RE-Modell mithilfe einer Simulationsstudie durchgeführt.

Das Gesetz von Hardy und Weinberg beschreibt im Grundsatz den Zusammenhang zwischen Allel- und Genotypfrequenzen in einer idealen Population. Es besagt, dass sowohl die Allelfrequenzen als auch die Genotypfrequenzen über Generationen hinaus in einem festen Verhältnis zueinander stehen und stabil bleiben wenn die Population hinreichend groß ist und sich jedes Individuum mit jedem des anderen Geschlechtes mit gleicher Wahrscheinlichkeit paart [44].

Zur genaueren Betrachtung des sogenannten Hardy-Weinberg Gleichgewichts (HWG) wird ein diallelischer autosomaler Locus mit den Allelen  $A_1$  und  $A_2$  unter HWG angenommen. Die möglichen Genotypkombinationen auf diesem Locus sind dann entsprechend  $A_1A_1$ ,  $A_1A_2$  und  $A_2A_2$ . Die Genotyphäufigkeiten in einer Population entsprechen dann:

$$P(A_1A_1) = p_{11}, \quad P(A_1A_2) = p_{12}, \quad P(A_2A_2) = p_{22}$$

mit  $p_{11} + p_{12} + p_{22} = 1$ . Mit den Genotypfrequenzen lassen sich die Allelfrequenzen  $P(A_1) = p$  und  $P(A_2) = q$  berechnen mit:

$$P(A_1) = p = p_{11} + \frac{1}{2}p_{12}, \quad P(A_2) = q = p_{22} + \frac{1}{2}p_{12}$$

mit  $p + q = 1$ .

Untersucht werden der empirische  $\alpha$ -Fehler mit einem nominellem  $\alpha = 5\%$  und die Teststärke für die Maßzahlen: Disequilibrium-Koeffizient (Abschn. 5.2.2), Inzucht-Koeffizient (Abschn. 5.2.1) und relative Exzess-Heterozygotie (REH) (Abschn. 5.2.3).

Des Weiteren werden das RE-Modell und das exakte RE-Modell, basierend auf dem Resultat der Teststärke der jeweiligen Maßzahlen, durch Differenzbildung miteinander verglichen.

## 5.2 Maße für das Hardy-Weinberg Gleichgewicht

Für genetische Assoziationsstudien wird häufig empfohlen, die Genotypverteilung in den Kontrollgruppen auf HWG zu überprüfen. Die Untersuchung auf HWG dient als Surrogat für die Qualität der Daten einer Kontrollgruppe. Abweichungen vom HWG können auftreten bei Populationsstratifikation, Genotypisierungsfehlern oder durch eine Auswahlverzerrung [53]. Insbesondere ist HWG häufig eine Voraussetzung für, auf Allelfrequenzen basierende, statistische Tests. Eine Abweichung vom HWG in den Kontrollen führt i. Allg. bei diesen Tests zu einer Erhöhung von falsch positiven Ergebnissen [54, 55]. Dies kann letztendlich auch dazu führen, dass der Gesamteffektschätzer einer Meta-Analyse durch eine Abweichung vom HWG in den Kontrollen verzerrt ist und dies wiederum möglicherweise in einer Erhöhung von falsch positiven Ergebnissen resultiert.

Die Frage, ob Studien mit einer Abweichung vom HWG grundsätzlich von einer Meta-Analyse ausgeschlossen werden soll, ist nicht eindeutig geklärt. Minelli et al. [53] haben in ihrer Arbeit darauf hingewiesen, dass eine *do-nothing* Strategie (also die Kontrollstudien unangetastet zu lassen), zu den aussagekräftigsten Ergebnissen führt, wenn auch mit einem kleinen Verlust der Präzision. Sie empfehlen dennoch eine Untersuchung auf die Ursachen der Studien, die zu einer Abweichung vom HWG führen.

Sie kritisieren zudem, dass das eigentliche Problem nicht nur die häufig schlechte Teststärke bei gängigen Tests zum Testen auf Abweichung von HWG ist, sondern auch die fehlende Aussage über die Größenordnung der Abweichung. Dies wird

sehr deutlich bei der Betrachtung von statistischen Tests, die mit einem  $\alpha$ -Niveau die Alternativhypothese hinsichtlich einer Abweichung vom HWG überprüfen. Denn mit einem fixem  $\alpha$ -Niveau kann selbst bei einer leichten Abweichung vom HWG in einer (theoretischen) Kontrollgruppe mit stetiger Erhöhung der Fallzahl immer die Nullhypothese abgelehnt werden. Eine Aussage über die Größenordnung der Abweichung vom HWG oder ein Test auf HWG, entsprechend einem Äquivalenztest, würde dieses Defizit ausräumen.

Nachfolgend werden drei Maße (der Inzucht-Koeffizient  $f$ , der Disequilibrium-Koeffizient  $D$  und das REH (relative excess of heterozygosity;  $\omega$ )) beschrieben, mit denen man den Nachteil einer fehlenden Quantifizierung der Abweichung vom HWG ausräumen kann. Mit ihnen ist es möglich, den Grad der Abweichung vom HWG zu quantifizieren.

Zusätzlich ist es möglich eine Bewertung auf einer Abweichung vom HWG nicht nur auf einzelne Studien abzugeben, sondern auch beispielsweise über einen Gesamtschätzer, der mittels einer Meta-Analyse über alle beteiligten Studien erhoben wurde. Das REH  $\omega$  hebt sich hier von den anderen Maßen ab, da für dieses Maß bereits ein Test auf Äquivalenz etabliert wurde [56].

### 5.2.1 Der Inzucht-Koeffizient $f$

Der Inzucht-Koeffizient beschreibt die Wahrscheinlichkeit, dass beide Allele eines Individuums der Abstammung nach identisch sind (identical-by-descent; IBD). Eine Abschätzung für  $f$  ist definiert durch [44, 57, 58]:

$$\hat{f} = 1 - \frac{2n_{12}n}{(2n_{11})(n_{12} + 2n_{22})} = 1 - \frac{n_{12}}{2n\hat{p}\hat{q}}$$

und die Varianz von  $\hat{f}$  ist approximativ gegeben mit [44]:

$$\text{Var}(\hat{f}) = (1 - f)^2(1 - 2f) + \frac{f(1 - f)(2 - f)}{2np(1 - p)}$$

wobei  $n_{11}$ ,  $n_{12}$  und  $n_{22}$  die Anzahl der Genotypen  $A_1A_1$ ,  $A_1A_2$  und  $A_2A_2$  sind und  $n$  die Gesamtanzahl der Genotypen bzw. Individuen entspricht.

### 5.2.2 Der Disequilibrium-Koeffizient $D$

Der Disequilibrium-Koeffizient misst die Differenz der beobachteten und erwarteten Genotypfrequenzen bei Annahme eines HWG und ist definiert durch [59]:

$$D = p_{11} - p^2 = p_{22} - q^2 = -\frac{1}{2}p_{12} + pq \quad .$$

Erwartungswert und Varianz für den Schätzer  $\hat{D}$  sind gegeben mit [60]:

$$\begin{aligned} \mathbb{E}(\hat{D}) &= D - \frac{1}{2n}(pq + D) \quad \text{und} \\ \text{Var}(\hat{D}) &= \frac{1}{n}(p^2(1-p)^2 + (1-2p)^2D - D^2) \quad . \end{aligned}$$

### 5.2.3 Das REH $\omega$

Das REH (*relative excess of heterozygosity*) dient als Maß für den relativen Überschuss an Heterozygotie. Es baut maßgeblich auf die Vorarbeiten von Pereira & Rogatko [61] sowie Lindley [62] auf. Sie definieren  $\gamma$  als Maß für die Abweichung vom HWG mit [60]:

$$\gamma = \frac{4p_{11}p_{22}}{p_{12}^2}$$

wobei  $\gamma = 1$  dem HWG entspricht und  $\gamma < 1$  einen Überschuss von heterozygoten Genotypen bzw. ein Defizit von homozygoten Genotypen entspricht. Die Interpretation kehrt sich bei  $\gamma > 1$  entsprechend um und stellt somit einen Überschuss von homozygoten Genotypen bzw. ein Defizit von heterozygoten Genotypen dar.

Das REH ist eine einfache monotone Transformation von  $\gamma$  und ist definiert durch:

$$\omega = \gamma^{-1/2} = \frac{p_{12}}{2\sqrt{p_{11}p_{12}}}$$

Entsprechend ist für ein  $\omega > 1$  ein relativer Überschuss an Heterozygotie gegeben bzw. bei einem  $\omega < 1$  ein relativer Überschuss an Homozygotie.

Beide Maße beruhen auf einer zugehörigen linearen Exponentialfamilie mit bedingter Wahrscheinlichkeitsverteilung [63]. Dies ermöglicht u.a. die Verwendung

von Standardmethoden für statistisches Testen und Parameterschätzungen [60]. Des Weiteren lässt sich ein Anpassungstest mithilfe von Konfidenzintervalle (Gleichung (5.1)) konstruieren, um eine Abweichung vom HWG festzustellen. Für den Anpassungstest wird dafür das traditionelle zweiseitige z-Quantile mit  $z = z_{1-\frac{\alpha}{2}}$  verwendet.

Da  $\ln \omega$  einer asymptotischen Normalverteilung folgt und auch die Varianz mithilfe der beobachteten Genotyphäufigkeiten geschätzt werden kann, lässt sich das Konfidenzintervall mithilfe der entsprechenden z-Quantile aus der Normalverteilung herleiten [56]. Die Rücktransformation auf die Originalskala ergibt dann das Konfidenzintervall für das REH  $\omega$ :

$$\begin{aligned}\overline{KI}(\hat{\omega}) &= \hat{\omega} / \exp \left( z \sqrt{\frac{1}{n} \left( \frac{1 - \hat{p}_{12}}{4\hat{p}_{11}\hat{p}_{22}} + \frac{1}{\hat{p}_{12}} \right)} \right) \\ \underline{KI}(\hat{\omega}) &= \hat{\omega} \cdot \exp \left( z \sqrt{\frac{1}{n} \left( \frac{1 - \hat{p}_{12}}{4\hat{p}_{11}\hat{p}_{22}} + \frac{1}{\hat{p}_{12}} \right)} \right)\end{aligned}\tag{5.1}$$

Zudem kann mithilfe einer a priori festgelegten Region in Kombination mit den Anpassungstest ein Intervall-Inklusionstest auf Äquivalenz erstellt werden um eine Übereinstimmung mit HWG zu ermitteln. Eine Übereinstimmung ist dann gegeben wenn das Konfidenzintervall (Gleichung (5.1)) innerhalb der festgelegten Region liegt. Diese Prozedur ist asymptotisch valide, wenn ein z-Quantil mit  $z = z_{1-\alpha}$  verwendet wird [56]. Dies entspricht Konfidenzschranken  $\underline{KI}$  und  $\overline{KI}$ , die je einseitiges Konfidenzniveau von 95% aufweisen.

Wellek et al. und Ziegler et al. [56, 60] haben gezeigt, dass die Grenzen für eine Äquivalenzregion  $[\frac{5}{7}, \frac{7}{5}]$  hinreichend genügen, wenn folgende Bedingungen erfüllt sind:

- Über den gesamten Wertebereich der kleineren Allelfrequenz (MAF) berücksichtigt, darf die Wahrscheinlichkeit für eine Ablehnung des *Lack-of-Fit* unter einer Genotypverteilung, die genau mit dem Modell übereinstimmt, nicht kleiner als 90% sein. Dies kann mit einem klassischen  $\chi^2$ -Test für *Lack-of-Fit* [56] berechnet werden.
- Die Teststärke sinkt mit sinkender MAF um eine Assoziation in einer Fall-Kontroll Studie oder einer Kohortenstudie zu entdecken. Somit sollte der



Bereich zum Testen auf HWG zwischen  $0,1 < \text{MAF} < 0,5$  beschränkt werden.

- Die Stichprobengröße für eine genomweite Assoziationsstudie sollte 3000 Proben (Individuen) nicht übersteigen.

Die Nullhypothese für eine wesentliche Abweichung vom HWG kann dann entsprechend abgelehnt werden, wenn das Konfidenzintervall mit z-Quantil  $z = z_{1-\alpha}$  ein Teilintervall der Äquivalenzgrenzen  $[\frac{5}{7}, \frac{7}{5}]$  ist.

## 5.3 Simulationsaufbau

Als Grundlage dienen  $k$  Einzelstudien mit je  $n$  Individuen und den dazugehörigen simulierten Genotypen. Aus jeder Studie werden der Inzucht-Koeffizient  $f$ , der Disequilibrium-Koeffizienten  $D$ , und das REH  $\omega$  geschätzt. Mittels der Meta-Analyse werden die Schätzer der  $k$  Studien gepoolt, um den jeweiligen Gesamteffektschätzer zu erhalten.

Die Realisierungen  $n_{11}, n_{12}, n_{22}$  der Genotypen  $A_1A_1, A_1A_2$  und  $A_2A_2$  von  $n$  Individuen werden dazu mithilfe der Multinomialverteilung und den entsprechenden Wahrscheinlichkeiten  $p_{11}, p_{12}, p_{22}$  erzeugt mit:

$$\begin{aligned} p_{11} &= g^2 + f * g * (1 - g) \\ p_{12} &= (1 - g) * 2 * g * (1 - g) \\ p_{22} &= (1 - g)^2 + f * g * (1 - g) \end{aligned}$$

wobei gilt, dass  $p_{11} + p_{12} + p_{22} = 1$  und  $n_{11} + n_{12} + n_{22} = n$  ergibt. Die Variable  $g$  gibt die gewünschte Allelfrequenz vor und  $f$  stellt den Inzucht-Koeffizienten dar. Dementsprechend ergibt sich eine Abweichung von der erwarteten Genotypfrequenz unter HWG genau dann, wenn der Inzucht-Koeffizient  $f > 0$  gewählt wird.

Die Simulationsstudie baut auf mehreren Simulationsszenarien auf, die alle den folgenden gemeinsamen Grundaufbau besitzen. Zur Initialisierung werden Kombinationen aus  $k = 1, \dots, 5$  Einzelstudien und einer Allelfrequenz  $\tilde{g}$  (im

Bereich von  $[0, 1; 0, 4]$  in 0,1 Schritten) realisiert. Zudem wird eine Variation der Allelfrequenzen berücksichtigt. Die Variation wird mithilfe einer Gleichverteilung im Bereich  $[\tilde{g}; \tilde{g} + 0.1]$  realisiert und wird für jede Einzelstudie gezogen, woraus sich die Allelfrequenz  $g$  ergibt.

Die weiteren Variationen werden hinsichtlich der Studiengröße  $n_k$  sowie des Inzucht-Koeffizienten  $f$  (im Bereich von  $[0; 0, 1]$  in 0,01 Schritten) in folgende Szenarien unterschieden:

- dem ersten Szenario werden allen 5, 20 und 50 Einzelstudien ein  $n = 1000$  vorgegeben,
- das zweite Szenario erhält für die ersten 4 Einzelstudien jeweils ein  $n$  von 2500 und die letzten Einzelstudie ein  $n$  mit 10000 um eine überproportional große Einzelstudie vorzugeben um die Auswirkung bei Nichteinhaltung der Forderungen von [56, 60] zu analysieren,
- das dritte Szenario erhält für alle  $k = 1, \dots, 5$  Einzelstudien ein  $n_k$  mit 1000, jedoch wird der letzten Einzelstudie, ausgehend von den initialisierten Werten, dem Inzucht-Koeffizienten 0,2 hinzuaddiert.

Alle Simulationsszenarien wurden mit 10000 Iterationen durchgeführt. Zusätzlich wurde anhand der Teststärke der verschiedenen Simulationsszenarien und Maße ein Vergleich zwischen dem RE-Modell und dem exakten RE-Modell berücksichtigt.

## 5.4 Ergebnisse

### 5.4.1 $\alpha$ -Fehler und Teststärke

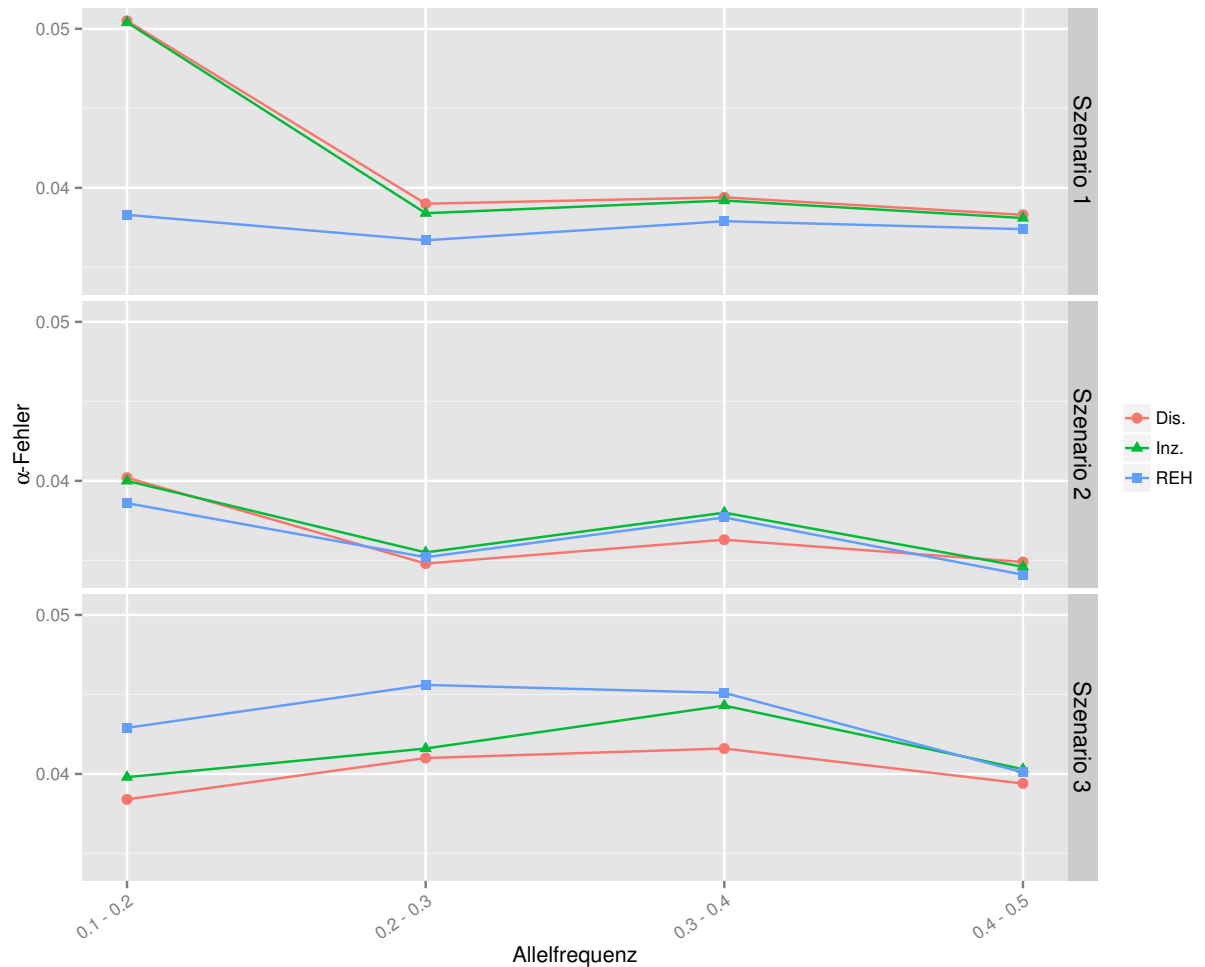
Die Simulationsergebnisse sind für die drei Szenarien, bei einem nominellem  $\alpha$ -Fehler von 5%, in Abbildung 5.1 und in Tabelle 5.1 dargestellt. Für Simulationsszenario 1 sind die Ergebnisse hierbei nur für fünf Studien dargestellt.

Für den empirischen  $\alpha$ -Fehler in Abbildung 5.1 zeigt sich für Szenario 1,

daß sowohl für den Disequilibrium-Koeffizienten als auch für den Inzucht-Koeffizienten das  $\alpha$ -Niveau für den Bereich der Allelfrequenz von  $g = [0, 1; 0, 2]$  knapp verfehlt wird, jedoch wird das Niveau für die restlichen Allelfrequenzbereiche eingehalten. Das REH hält das  $\alpha$ -Niveau über den gesamten Bereich zwar ein, schöpft das Niveau jedoch schlechter aus als die beiden anderen Maßzahlen.

Im zweiten Szenario wird das Niveau von allen Maßzahlen eingehalten, jedoch ist keine Überlegenheit einer bestimmten Maßzahl festzustellen, da die Ergebnisse über den gesamten Allelbereich heterogen verteilt sind.

Im dritten Szenario halten wiederum alle Maßzahlen das Niveau ein. Des Weiteren ist eine absolute Überlegenheit des REH gegenüber den anderen Maßzahlen festzustellen. Es schöpft das  $\alpha$ -Niveau in diesem Szenario gegenüber den anderen Maßzahlen am besten aus, ohne es jedoch zu überschreiten.



**Abbildung 5.1:** Die Grafik zeigt den empirischen  $\alpha$ -Fehler mit nominellem  $\alpha = 5\%$  (y-Achse) für die drei Simulationsszenarien (Abschn. 5.3) der drei gepoolten Hardy-Weinberg Gleichgewicht-Maße in Abhängigkeit von der gewählten Allelfrequenz (x-Achse). Die Hardy-Weinberg Gleichgewicht-Maße Disequilibrium-Koeffizient (Dis.), Inzucht-Koeffizient (Inz.) und das REH sind jeweils in rot, grün, und blau dargestellt. Die punktuellen Ergebnisse wurden mithilfe des zufälligen Effekte Modells erzielt. Sie sind ausschließlich zur Verdeutlichung mit Linien verbunden und stellen somit keinen kontinuierlichen Verlauf dar.

Tabelle 5.1 zeigt die Teststärke der jeweiligen Maßzahlen für jedes der drei Szenarien in Abhängigkeit vom Inzucht-Koeffizienten und der kleineren Allelfrequenz. Generell zeigt sich ein monotoner Anstieg der Teststärke mit steigenden simulierten Inzucht-Koeffizienten. Dieser Effekt gilt auch für ein Ansteigen der

kleineren Allelfrequenzbereiche von  $[0, 1; 0 - 2]$  hin zu  $[0, 4 - 0, 5]$ .

Für alle drei Simulationsszenarien zeigt das REH eine klare Überlegenheit hinsichtlich der Teststärke auf. Der Disequilibrium-Koeffizient liegt mit seinen Werten immer unterhalb und konvergiert später zu einer Teststärke von 100% als die beiden anderen Maße. Die Teststärke für den Inzucht-Koeffizient liegt damit immer zwischen den Werten vom Disequilibrium-Koeffizienten und vom REH.

**Tabelle 5.1:** Die Tabelle zeigt die Teststärke mit nominellem  $\alpha = 5\%$  für Simulationsszenario 1 bis 3 jeweils für den Inzucht-Koeffizienten (Inz.), den Disequilibrium-Koeffizienten (Dis.) und das REH für das zufällige Effekte Modell auf. Sortiert sind die Ergebnisse nach aufsteigender kleineren Allelfrequenz (MAF) zwischen 0,1 und 0,5 in jeweils 10% Abschnitten und mit initialen simulierten Inzucht-Koeffizienten (Sim. Inz.) für die Werte 0,1;0,3;0,5;0,7. Für Simulationsszenario 1 wurden nur das Unterszenario mit fünf Studien berücksichtigt.

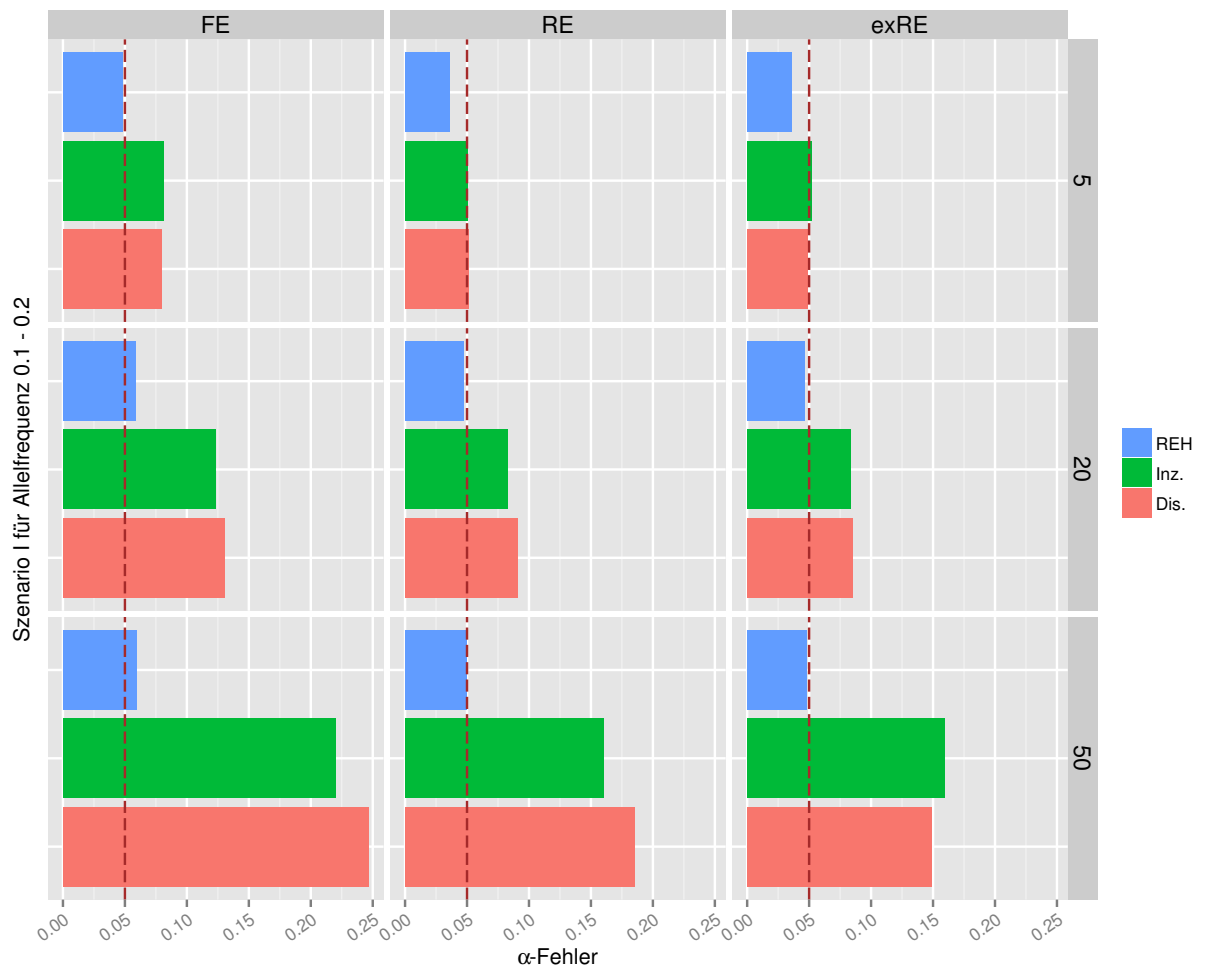
MAF	Szenario 1				Szenario 2			Szenario 3		
	Sim. Inz.	Dis.	Inz.	REH	Dis.	Inz.	REH	Dis.	Inz.	REH
0,1 – 0,2	0,01	0,053	0,060	0,093	0,184	0,194	0,239	0,076	0,088	0,136
	0,03	0,344	0,372	0,480	0,911	0,924	0,944	0,433	0,463	0,568
	0,05	0,780	0,802	0,873	0,998	0,999	1	0,833	0,856	0,912
	0,07	0,970	0,975	0,989	1	1	1	0,979	0,983	0,993
> 0,2 – 0,3	0,01	0,070	0,076	0,089	0,199	0,208	0,224	0,099	0,106	0,125
	0,03	0,422	0,440	0,477	0,939	0,942	0,949	0,513	0,532	0,566
	0,05	0,862	0,874	0,893	0,999	1	1	0,892	0,903	0,918
	0,07	0,984	0,986	0,990	1	1	1	0,989	0,992	0,994
> 0,3 – 0,4	0,01	0,070	0,075	0,079	0,215	0,221	0,226	0,109	0,114	0,120
	0,03	0,452	0,466	0,475	0,948	0,949	0,950	0,538	0,555	0,565
	0,05	0,878	0,886	0,891	0,999	1	1	0,914	0,920	0,924
	0,07	0,990	0,991	0,992	1	1	1	0,994	0,995	0,995
> 0,4 – 0,5	0,01	0,076	0,080	0,080	0,223	0,230	0,230	0,107	0,112	0,113
	0,03	0,455	0,471	0,471	0,947	0,949	0,949	0,550	0,563	0,564
	0,05	0,892	0,898	0,898	0,999	1	0,999	0,921	0,925	0,925
	0,07	0,993	0,993	0,993	1	1	1	0,994	0,994	0,994

Die Darstellung 5.2 zeigt das Simulationsszenario 1 mit einer Studienanzahl von 5, 20 und 50 Studien. Dabei entsprechen die Ergebnisse mit einer Studienanzahl von fünf genau den Ergebnissen aus 5.1 für den Allelfrequenzbereich von 0,1

bis 0,2 für das RE-Modell.

Grundsätzlich zeigt die Darstellung, daß bei steigender Anzahl von Studien das  $\alpha$ -Niveau für den Disequilibrium-Koeffizienten und den Inzucht-Koeffizienten nicht eingehalten wird. Bei dem Vergleich zwischen dem RE-Modell bzw. dem RE-Modell mit exakten Gewichten und dem FE-Modell fällt auf, daß die Zwischenstudienvarianz für den Disequilibrium-Koeffizienten und den Inzucht-Koeffizienten mit wachsender Anzahl von Studien proportional ansteigt und diese von den RE-Modellen, bis zu einem gewissen Grad, kompensiert werden kann. Dieser Effekt resultiert daraus, dass sowohl der Disequilibrium-Koeffizient als auch der Inzucht-Koeffizient nicht konstant über alle Studien bleibt wenn die Allelfrequenzen innerhalb dieser Studien variieren [64]. Dies verdeutlicht sich dann auch im Anteil der Überschreitung des  $\alpha$ -Niveaus, bzw. in der Länge der Balken, zwischen dem FE-Modell und den RE-Modellen. Der Anteil der berücksichtigten Zwischenstudienvarianz bei den RE-Modellen ist so indirekt über die Differenz der Länge der Balken zwischen den RE-Modellen und dem FE-Modell abzulesen.

Das REH hingegen überschreitet beim FE-Modell für die Studienanzahl 20 und 50 nur kaum das  $\alpha$ -Niveau und es hält das Niveau für alle Szenarien und Studienanzahlen immer ein, wenn eines der beiden RE-Modelle benutzt wurde, da die zuvor erwähnte fehlende Konstanz bei variierenden Allelfrequenzen nicht oder nur unbedeutend auftritt.



**Abbildung 5.2:** Die Grafik zeigt den empirischen  $\alpha$ -Fehler mit nominellem  $\alpha = 5\%$  (x-Achse) für die drei Simulationsszenarien (Abschn. 5.3) der drei gepoolten Hardy-Weinberg Gleichgewicht-Maße, in Abhängigkeit von der gewählten Allelfrequenz zwischen 0,1 und 0,2 und der Anzahl der Studien mit 5, 20 und 50 (y-Achse, rechte Seite). Die Hardy-Weinberg Gleichgewicht-Maße Disequilibrium-Koeffizient (Dis.) , Inzucht-Koeffizient (Inz.) und das REH sind jeweils in rot, grün, und blau dargestellt. Weiterhin sind die Ergebnisse unterteilt für das feste Effekte Modell (FE), das zufällige Effekte Modell (RE) und das Modell für zufällige Effekte mit exakten Gewichten (exRE) (x-Achse, oben).

### 5.4.2 Vergleich zwischen dem Modell mit zufälligen Effekten und dem Modell mit zufälligen Effekten mit exakten Gewichten

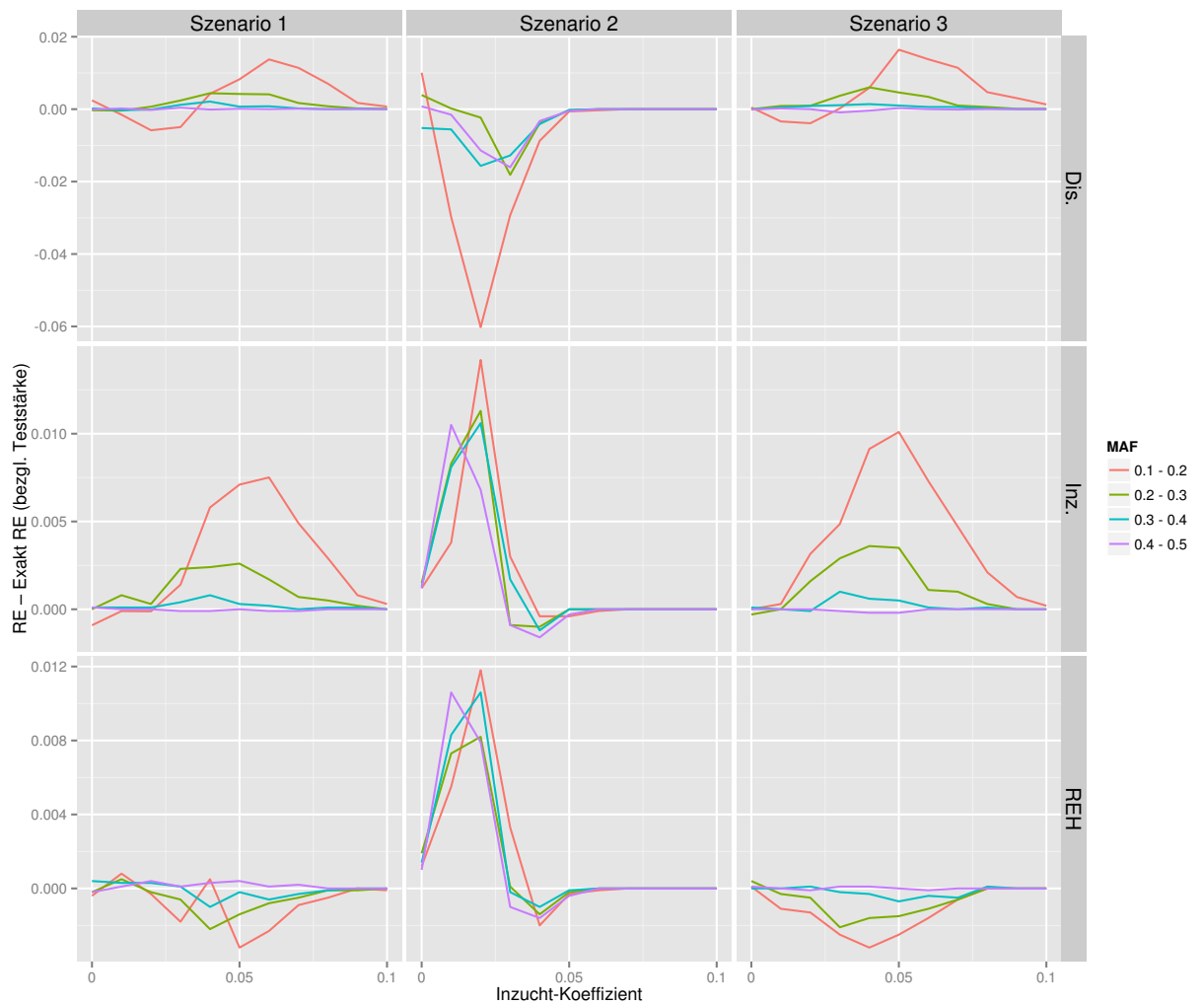
Der Vergleich zum Unterschied der Ergebnisse zwischen dem dem RE-Modell und dem exakten RE-Modell ist in Abbildung 5.3 dargestellt. Dabei wurden die Parameter zur Simulation für die Ergebnisse aus Tabelle 5.1 verwendet, um die Gesamteffektschätzer der jeweiligen Modelle zu ermitteln. Der Unterschied der Modelle ist mit einer Differenzbildung aus dem RE-Modell und dem exakten RE-Modell der beiden Gesamteffektschätzer bestimmt worden.

Positive Werte auf der y-Achse bedeuten eine höhere Teststärke für das RE-Modell und infolgedessen bedeuten negative Werte eine höhere Teststärke für das exakte RE-Modell.

Für den Disequilibrium-Koeffizienten ergeben sich im ersten und dritten Szenario nur bei einer Allelfrequenz zwischen  $[0,1 - 0,2]$  größere Abweichungen zugunsten des RE-Modells. Die größte Abweichung beträgt dabei ca. 1,6% bei einem simulierten Inzucht-Koeffizienten von ca. 0,06 für das erste Szenario und einem simulierten Inzucht-Koeffizienten von ca. 0,05 für das dritte Szenario. Das zweite Szenario zeigt für den Disequilibrium-Koeffizienten die höchste Abweichung zugunsten des exakten Modells hinsichtlich aller Szenarien und aller Maße. Während die Abweichung der Allelfrequenzbereiche ab 0,2 – 0,3 alle ein Maximum von ca. 1,8% bei einem simulierten Inzucht-Koeffizienten von 0,025 erreichen, ist der Wert für den Bereich 0,1 – 0,2 um ein Vielfaches höher und erreicht sein Maximum bei einem simulierten Inzucht-Koeffizienten von 0,02 mit einem Differenzwert von 6% zugunsten des exakten RE-Modells.

Sowohl der Inzucht-Koeffizient als auch das REH zeigen ein ähnliches Verhalten mit einer anfänglich stark steigenden Abweichung aller Allelfrequenzbereiche zugunsten des RE-Modells mit je einem Maximum von 10% – 12% bei einem simulierten Inzucht-Koeffizienten von ca. 0,02. Ab einem simulierten Inzucht-Koeffizienten von 0,05 senkt sich die Differenz der Modelle für alle Maße auf kaum nennenswerte Unterschiede ab, da hier die maximal erreichbare Teststärke für das RE-Modell erreicht wurde (Tab. 5.1) und das exakte RE-Modell nicht von einem monotonen Wachstum der Teststärke abweicht.





**Abbildung 5.3:** Die Grafik zeigt die Differenz des zufälligen Effekte Modells (RE) und des zufälligen Effekte Modells mit exakten Gewichten (Exakt RE) bezüglich der Teststärke (y-Achse; mit unterschiedlicher Skalierung) (Abb. 5.1) für die drei gepoolten Hardy-Weinberg Gleichgewicht-Maße Disequilibrium-Koeffizient (Dis.), Inzucht-Koeffizient (Inz.) und REH (Zeilen) und den drei Simulationsszenarien (Spalten) (Abschn. 5.3) in Abhängigkeit vom Inzucht-Koeffizienten (x-Achse). Die die Wahl der kleineren Allelfrequenz  $g$  (MAF) zwischen  $0,1 - 0,2; \dots; 0,4 - 0,5$  wurde farblich abgetragen (Legende und Abschn. 5.3).

## 5.5 Interpretation der Ergebnisse

Die HWG-Maße lassen sich alternativ zu den statistischen Tests auf HWG (z.B. basierend auf einer  $\chi^2$ -Verteilung) verwenden und bieten zusätzlich den Vorteil einer Quantifizierung der Abweichung vom HWG. Damit kann beispielsweise das Problem ausgeräumt werden, dass eine beliebige minimale Abweichung vom HWG mit hinreichend großer Fallzahl immer nachgewiesen werden kann, obwohl dies aus biologischer Sicht keine Rolle spielt. So lässt sich z.B. eine Grenze für die HWG-Maße aufstellen, die den Grad einer gerade noch hinnehmbaren Abweichung vom HWG festlegt.

Die Maße eignen sich nicht nur für einzelne genomweite Assoziationsstudien sondern lassen sich auch, wie die Studieneffekte aus genomweiten Assoziationsstudien, poolen. Dann kann auch über die gepoolte Gesamtkohorte auf eine Abweichungen vom HWG geprüft werden.

Die dafür angelegten Simulationsergebnisse zeigen auf, dass das REH besser abschneidet als die beiden anderen betrachteten Maße. Es hält als einziges Maß das  $\alpha$ -Niveau für Simulationsszenario 1 ein und zeigt auch eine klare Überlegenheit in Simulationsszenario 3.

Der Hauptgrund für das bessere Abschneiden des REH-Maß ist die fehlende Konstanz des Disequilibrium-Koeffizienten und des Inzucht-Koeffizienten bei variierenden Allelfrequenzen innerhalb von Einzelstudien [64]. Dies führt zu einer Zwischenstudienvarianz, die sich mit steigender Anzahl von Studien verstärken kann und somit für das FE-Modell zu einer Nichteinhaltung des  $\alpha$ -Niveaus führt. Beide RE-Modelle können diese Eigenschaft ebenfalls nur unzureichend kompensieren.

Der Vergleich der Maße hinsichtlich der RE-Modelle führt nur für den Inzucht-Koeffizienten zu einer eindeutigen Überlegenheit im Bezug zum RE-Modell. Hier ist ein Unterschied von bis zu 1,5% in der Teststärke zu verzeichnen. Die anderen Maße schneiden unterschiedlich bezüglich der Methode und den Simulationsszenarien ab, so dass hier keine eindeutigen Empfehlungen für eines der beiden meta-analytischen Modelle gegeben werden kann.

## 6 Illustration anhand von Realdaten

In diesem Abschnitt werden anhand von vier Realdaten die eintretenden möglichen Unterschiede zwischen dem Standard RE-Modell und dem RE-Modell mit exakten Gewichten aufgezeigt. Es handelt sich dabei um folgende Datensätze:

1. Diuretika und Präeklampsie [50]: Bei diesem Beispiel wurden 9 Studien in einer Meta-Analyse berücksichtigt um die Prävention von Präeklampsie mithilfe der Gabe von Diuretika zu bewerten. Dieser Datensatz diente schon des öfteren in verschiedenen methodischen Artikeln als Beispieldatensatz [1, 45, 65–67].
2. Krankenhaus zu Hause bei früher Entlassung (orig. „Hospital at home early discharge“) [68, Analysis 2.2]. In dieser Studie wurde eine frühe Entlassung aus dem Krankenhaus nach Hause mit der stationären Krankenhausbehandlung verglichen.
3. Inkubatoren für die Reduzierung der Wärmeverluste bei Säuglingen mit sehr geringem Geburtsgewicht (orig. „Incubators for reducing heat loss in very low birth weight (VLBW) infants“) [69, Analysis 1.5]. In dieser Meta-Analyse wurde bei Säuglingen mit sehr geringem Geburtsgewicht verglichen, ob ein einwandiger Inkubator zu einer höheren abdominalen Temperatur als ein doppelwandiger Inkubator führt.
4. Der *CASP8* -652 6N del Promoter Polymorphismus und Brustkrebsrisiko: Eine multizentrische Studie [70]. In dieser Studie wurde die sechs-Nukleotid-Deletion im Gen *CASP8* analysiert. Biochemische Analysen legen nahe, dass diese Deletion das Risiko für verschiedene Typen von Krebs, insbesondere Brustkrebs, erhöht.

Dabei werden mit dem ersten Datensatz Unterschiede hinsichtlich der Gewichtung und der Heterogenitätsmaße aufgezeigt (Abschn. 6.1). Mithilfe der Datensätze zwei und drei werden Unterschiede hinsichtlich der Signifikanz des Gesamteffektschätzers aufgezeigt. Mit dem letzten Datensatz werden für die Kontrollen die HWG-Maße (Abschn. 5.2), die im Anschluss den beiden RE-Modellen gepoolt werden.

## 6.1 Diuretika und Präeklampsie Meta-Analyse

Der Diuretika und Präeklampsie Datensatz aus Tabelle 6.1 besteht aus einer Übersichtsarbeit [50] von randomisierten Präventionsstudien mit Präeklampsiepatientinnen hinsichtlich der Gabe von Diuretika. Die Tabelle listet die Einzelstudien mit ihren Odds-Ratio-Schätzern und 95%-Konfidenzintervalle sowie die Varianzen auf. Des Weiteren wurden die Gewichte des FE-Modells sowie des Standard RE-Modells und des exakten RE-Modells ergänzt. Bei den Gewichten zeigt sich zunächst der große Unterschied zwischen dem FE-Modell und den RE-Modellen. Dies ist nicht überraschend, da eine große klinische Heterogenität (Abschn. 1.3) vorliegt, die sich mit der großen Variabilität der Krankheitsausprägung begründen lässt [50]. Vergleicht man die RE-Modelle untereinander zeigt sich, dass das exakte Modell, mit Ausnahme der Campell Studie, alle Gewichte zwischen dem RE-Modell und dem asymptotischen Modell liegen. Bei der Campell Studie ist das exakte Gewicht das Größte im Vergleich zu allen anderen Studien. Dies stellt somit eine Ausnahme von der Regel dar, da (zumindest approximativ) normalerweise alle exakten RE-Gewichte zwischen den Gewichten vom FE-Modell und den Gewichten vom RE-Modell liegen [45].

**Tabelle 6.1:** Diuretika und Präeklampsie Datensatz. Angezeigt sind auf der linken Seite die geschätzten Odds-Ratios pro Studie  $i$  ( $\widehat{OR}_i$ ) sowie 95%-Konfidenzintervall (KI), die log Odds-Ratios mit  $Y_i$  ( $Y_i = \ln(OR_i)$ ) und die geschätzten Varianzen  $\hat{\sigma}_i^2$ . Auf der rechten Seite finden sich die relativen Gewichte in Prozent nach Berechnung der drei Modelle, das feste Effekte Modell (FE Gewicht), das zufällige Effekte Modell (RE Gewicht) und das zufällige Effekte Modell mit exakten Gewichten (ex RE Gewicht). Die Studiendaten stammen aus [50], S. 1128; siehe auch [45], Tabelle I.

Studie	95%-KI	$\widehat{OR}_i$	$\hat{\sigma}_i^2$	FE Gewicht	RE Gewicht	exRE Gewicht
Weseley	1,04	0,160	[0,48;2,28]	5,06	10,66	8,77
Flowers	0,40	0,118	[0,20;0,78]	6,86	11,94	10,46
Menzies	0,33	0,178	[0,14;0,74]	4,55	10,19	8,22
Fallis	0,23	0,299	[0,08;0,67]	2,71	7,86	5,87
Cuadros	0,25	0,114	[0,13;0,48]	7,10	12,08	10,66
Landesman	0,74	0,015	[0,59;0,94]	53,96	16,95	28,54
Krans	0,77	0,121	[0,39;1,52]	6,69	11,84	10,31
Tervila	2,97	0,686	[0,59;15,07]	1,18	4,54	3,16
Campbell	1,14	0,068	[0,69;1,91]	11,90	13,94	13,99

In Tabelle 6.2 werden die Ergebnisse für den Gesamteffektschätzer für die drei Modelle angezeigt. Zusätzlich sind das  $I^2$  und das trunkierte  $I^2$  (Spalte  $I^2$ , Zeile exakt RE) mit der exakten Verteilung von Cochran's Q aufgelistet [1] (Abschn. 2.1). Die Unterschiede zwischen dem RE-Modell und dem exakten RE-Modell sind in diesem Beispiel relativ gering. Der Gesamteffektschätzer des exakten RE-Modells liegt zwischen den beiden Gesamteffektschätzern des FE-Modells und des RE-Modells. Die Ähnlichkeit der Konfidenzintervalle vom RE-Modell und dem exakten RE-Modell stehen im Kontrast zu den teils erheblichen Unterschieden bei den Gewichten der beiden Modelle in Tabelle 6.1.

Der Schätzer für das  $I^2$  ist für das exakte Modell niedriger als beim klassischen RE-Modell. Zudem ist das Konfidenzintervall vom exakten  $I^2$  breiter. Es umschließt zudem die 0, was beim RE-Modell nicht der Fall ist.

**Tabelle 6.2:** Ergebnisse der Meta-Analyse für den Diuretika und Präeklampsie Datensatz. Die log Odds-Ratios (log OR) sind zusammen mit ihren 95%-Konfidenzintervallen (95%-KI für log OR) und ihrem  $p$ -Wert angegeben. Zusätzlich sind die Ergebnisse für das  $I^2$ , basierend auf dem zufälligen Effekte Modell (Zeile RE, Spalte  $I^2$ ) und den trunkierten zufälligen Effekte Modell mit exakten Gewichten (Zeile exRE, Spalte  $I^2$ ) zusammen mit den 95%-Konfidenzintervallen (95%-KI für  $I^2$ ) angegeben.

Modell	log OR	p-Wert	95%-KI für log OR	$I^2$	95%-KI für $I^2$
FE	0,67	$8,8 \times 10^{-6}$	[0,56;0,80]	—	—
RE	0,60	0,01	[0,40;0,89]	70,64	[40,27;85,57]
exRE	0,62	0,02	[0,41;0,94]	59,85	[ 0,00;88,47]

Tabelle 6.3 listet den  $\tau^2$  Schätzer und das Konfidenzintervall für den Diuretika und Präeklampsie Datensatz 6.1 von mehreren Verfahren auf, die bereits in der Literatur vorgeschlagen wurden [1, 45].

Des Weiteren wurde die Schätzung des Konfidenzintervalls für das  $\tau^2$  auf Grundlage der exakten Cochran's  $Q$ -Verteilung ergänzt (Tab. 6.3). Dieser Ansatz bietet den Vorteil einer gewissen Kohärenz, da sowohl das  $\tau_{DSL}^2$  und die Berechnung der Konfidenzintervalle auf der Cochran's  $Q$ -Verteilung basieren (Gleichung 2.5). Die anderen Ansätze können zu Inkonsistenzen führen, da generell der klassische DSL  $\tau_{DSL}^2$  Schätzer für RE-Meta-Analysen genutzt wird, jedoch andere Schätzer für das  $\tau^2$  entstehen (beispielsweise erhalten durch das Maximum-Likelihood-Verfahren) und benutzt werden um eine Schätzung für das Konfidenzintervall von  $\tau_{DSL}^2$  zu erhalten.

Bei der ersten Methode in Tabelle 6.3 handelt es sich um den klassischen DSL-Schätzer in Verbindung mit einer Satterthwaite Approximation [71] der Verteilung von  $Q$  unter der Bedingung, dass jedes  $\sigma_i^2 \ll \tau^2$  [45] ist. Biggerstaff und Tweedie kritisieren, dass diese Limitierung für die Praxis nicht akzeptabel ist [45]. Daraufhin entwickelten sie alternative Likelihood-Verfahren zur Schätzung von  $\tau^2$  und dem Konfidenzintervall, der asymptotische Verteilung vom Likelihood-Ratio von  $\tau^2$  (ALR) und die asymptotische Verteilung vom Maximum-Likelihood von  $\tau^2$  (AML) auf der Grundlage einer  $\chi^2$ -Verteilung sowie entsprechend einer Normalverteilung. Die letzte Methode (Exakt) basiert auf der Grundlage der exakten Schätzung der Cochran's  $Q$ -Verteilung (Abschn. 2.1).

**Tabelle 6.3:** Diuretika und Präeklampsie Datensatz. Geschätztes  $\tau^2$  und sein 95%-Konfidenzintervall (95%-KI) mit der DSL Momenten-Methode (MM) (Satterthwaite-Approximation bei Berechnung des Konfidenzintervalls), das asymptotische Likelihood-Ratio-Methode (ALR), das asymptotische Maximum-Likelihood-Methode (AML) und der exakten Methode (Exakt).

Methode	$\hat{\tau}^2$	95%-KI
MM	0,23	[0,04; 2,35]
ALR	0,24	[0,03; 1,13]
AML	0,24	[ 0; 0,57]
Exakt	0,23	[0,05; 1,43]

## 6.2 Unterschiede bei der Signifikanz bezüglich des zufällige Effekte Modells und des exakten zufällige Effekte Modells anhand von realen Datensätzen

Mithilfe von zwei weiteren Datensätzen wird gezeigt, wie sich die Gesamtschlussfolgerung einer Meta-Analyse, je nach Wahl der Modelle, ändern kann. Beide Beispiele, der Hospital-Datensatz sowie der Inkubator-Datensatz (Einleitung, Kap. 6), haben eine kleine Studienanzahl gemeinsam.

Der Fokus der wiederholten Analyse liegt auf die Veränderung der Gewichte sowie auf die Varianz des Gesamteffektschätzers im Bezug auf das exakte RE-Modell bei einem  $\tau^2 > 0$ . Die Ergebnisse der Einzelstudien finden sich auf der linken Seite der Tabelle 6.4 wieder und die Gewichte der drei Modelle finden sich auf der rechten Seite wieder.

**Tabelle 6.4:** Effekte der Einzelstudien (geschätztes log Hazard-Ratio  $\log HR$ ) für die Hospital Meta-Analyse und (gewichtete Mittelwertdifferenz) für die Inkubator Meta-Analyse (Kap. 6) und ihre geschätzten Standardfehler ( $S.E.$ ) sowie Gewichte für das FE-Modell (FE Gewicht), das RE-Modell (RE Gewicht) und das exakte RE-Modell (exRE Gewicht).

Datensatz	Studie	Effekt	S.E.	FE Gewicht	RE Gewicht	exRE Gewicht
Hospital	Cunliffe 2004	0,226	0,213	71,09	57,63	67,15
	Harris 2005	0,996	0,677	7,04	12,27	8,80
	Shepperd 1998	0,867	0,384	21,87	30,11	24,05
Inkubator	Bell 1983	-0,038	0,087	15,99	28,10	19,50
	Marks 1981	0,180	0,043	67,46	43,34	60,46
	Yeh 1980	0,210	0,086	16,55	28,55	20,04

In Tabelle 6.5 führen beide Beispielen mit der Berechnung des FE-Modells zu einer Ablehnung der Null-Hypothese: keinen Unterschied zwischen den Behandlungsgruppen. Aber es gibt wesentliche Unterschiede zwischen dem RE-Modell und dem exakten RE-Modell. Für das „Krankenhaus zu Hause bei früher Entlassung“ (Hospital) Beispiel erhält man mit dem RE-Modell einen  $p$ -Wert von 0,0430. Dieser ist somit kleiner als das nominelle  $\alpha = 0,05$ . Das exakte RE-Modell überschreitet jedoch klar das nominelle  $\alpha$  mit einem  $p$ -Wert von 0,0830. Die Ursache dieser Diskrepanz liegt vor allem bei der sich unterscheidenden Gewichtung der Cunliffe Studie. Diese Studie weist einen kleinen Studieneffekt auf, besitzt jedoch eine hohe Präzision mit  $S.E. = 0,213$  (Tab. 6.4). Das Gewicht fällt entsprechend hoch aus und beträgt für das exakte RE-Modell 67,15 jedoch im direkten Vergleich zu einem Gewicht von 57,63 für das RE-Modell.

Mit dem zweiten Beispiel „Inkubatoren für die Reduzierung der Wärmeverluste bei Säuglingen mit sehr geringem Geburtsgewicht“ (Inkubator) kehrt sich der Effekt entsprechend um. Das RE-Modell führt zu einem  $p = 0,0616 > 0,05$ , während für das exakte RE-Modell ein  $p = 0,0465 < 0,05$  berechnet wird.

Wiederum liegt die Ursache für den Unterschied bei der Gewichtung der Studien, vor allem bei der Bell Studie. Diese Studie hat einen Behandlungseffektschätzer der in die entgegengesetzte Richtung geht als es bei den beiden anderen Studien der Fall ist. Die Studie besitzt ein wesentlich höheres Gewicht beim RE-Modell mit 28,10 als beim exakten RE-Modell mit 19,50.



Solche Unterschiede können nur beobachtet werden, wenn eine große Heterogenität vorhanden ist. Und in der Tat beträgt das  $I^2$  für das erste Beispiel 31,24% (95%-KI [0,0%;92,8%]) und für das zweite Beispiel 64,3% (95%-KI [0,0%;89,8%]). Die Schätzungen für das  $\tau^2$  betragen 0,07 (95%-KI [0,00;9,98]) und 0,009 (95%-KI [0,00;0,57]), je für das erste und zweite Beispiel.

**Tabelle 6.5:** Ergebnisse der Meta-Analyse für das Hospital-Beispiel und das Inkubator-Beispiel. Das FE-Modell (FE), das RE-Modell (RE) und das exakte RE-Modell (Exakt RE) wurden für die Schätzung verwendet. Aufgelistet sind die gepoolten log Hazard-Ratios für den Hospital-Datensatz und die gepoolten gewichteten Mittelwertdifferenzen für den Inkubator Datensatz. Die zugehörigen 95%-Konfidenzintervalle (95%-KI) und die  $p$ -Werte für den Test der Nullhypothese: „keinen Unterschied zwischen den Behandlungsgruppen“.

Datensatz	Modell	Effekt	95%-KI	$p$ -Wert
Hospital	FE	0,420	[ 0,068;0,772]	0,0192
	RE	0,513	[ 0,016;1,011]	0,0430
	Exakt RE	0,448	[−0,059;0,954]	0,0830
Inkubator	FE	0,150	[ 0,082;0,219]	$1,7 \cdot 10^{-5}$
	RE	0,127	[−0,006;0,261]	0,0616
	Exakt RE	0,144	[ 0,002;0,285]	0,0465

### 6.3 Multizentrische Studie für den *CASP8* -652 6N del Promoter Polymorphismus

Mithilfe der Kontrollen aus der multizentrischen Studie werden die HWG-Maße berechnet und im Anschluss mithilfe der beiden meta-analytischen Modelle gepoolt. Die Auswertung demonstriert eine Erweiterung der Analysen von Ziegler et al. [60], bei dem ausschließlich das REH Maß für HWG sowie das FE-Modell und das RE-Modell benutzt wurde.

**Tabelle 6.6:** Ergebnisse für den CASP8 -652 6N del Promoter Polymorphismus Datensatz. Im oberen Teil der Tabelle sind die vier Studien mit den dazugehörigen Genotypen sortiert nach homozygoter Insertion (ins/ins), heterozygoter Insertion und Deletion (ins/del) und homozygoter Deletionen (del/del) und die studienweise berechneten Hardy-Weinberg Gleichgewicht-Maße Disequilibrium-Koeffizient (Dis.), Inzucht-Koeffizient (Inz.) und das REH sowie die dazugehörigen 95%-Konfidenzintervalle aufgelistet. Im unteren Teil der Tabelle sind die drei gepoolten Werte für die jeweiligen Maße mit den drei meta-analytischen Modellen, das feste Effekte Modell (FE-Modell), das zufällige Effekte Modell (RE-Modell) und das zufällige Effekte Modell mit exakten Gewichten (EX-Modell) sowie die dazugehörigen Konfidenzintervalle für die gepoolten Schätzer aufgelistet.

Studie	Genotyp			HWG-Maße			95%-KI HWG-Maße		
	ins/ins	ins/del	del/del	Dis.	Inz.	REH	Dis.	Inz.	REH
GFBCS	270	506	263	0,006	0,026	0,950	[-0,009;0,022]	[-0,035;0,088]	[0,841;1,072]
SBCS	245	608	321	-0,010	-0,040	1,084	[-0,024;0,004]	[-0,097;0,017]	[0,966;1,216]
GENICA	285	492	229	0,005	0,019	0,963	[-0,011;0,020]	[-0,043;0,081]	[0,851;1,090]
SEARCH	1149	2263	1062	-0,003	-0,012	1,024	[-0,010;0,004]	[-0,041;0,017]	[0,966;1,086]
FE - Modell				-0,0019	-0,0072	1,014	[-0,007;0,004]	[-0,030;0,015]	[0,970;1,060]
RE - Modell				-0,0018	-0,0067	1,013	[-0,008;0,004]	[-0,030;0,017]	[0,966;1,060]
EX - Modell				-0,0017	-0,0067	1,013	[-0,008;0,004]	[-0,030;0,017]	[0,966;1,060]

In Tabelle 6.6 kennzeichnen die drei HWG-Maße die Richtung für den Überschuss der homozygoten Genotypen im Vergleich zu HWG anhand der positiven oder negativen Abweichung vom Erwartungswert 0 bzw. für das REH eine Abweichung vom Erwartungswert 1. So ist bei GFBCS Studie ein Überschuss bei der Genotypkombination ins/ins im Vergleich zu HWG zu verzeichnen. Dies ist jeweils durch den positiven Disequilibrium-Koeffizienten und den positiven Inzucht-Koeffizienten bzw. einem REH  $\omega < 1$  charakterisiert.

Bei der studienweisen Auswertung der Kontrollgruppen zeigt sich bei allen HWG-Maße nur eine leichte Abweichung vom jeweiligen Erwartungswert. Zudem beinhalten alle zweiseitigen 95%-Konfidenzintervalle den jeweiligen Erwartungswert der Maße unter HWG, was einer perfekten Übereinstimmung mit HWG entspricht. Das bedeutet, dass keine Abweichung von HWG bei einem Signifikanzniveau von 5% festgestellt werden konnte.

Auch die Ergebnisse für die gepoolten Werte im unteren Teil der Tabelle 6.6 weisen für alle meta-analytischen Modelle eine perfekte Übereinstimmung mit HWG auf. Des Weiteren weisen die Modelle auch untereinander nur kaum eine Abweichung bezüglich der Gesamteffektschätzer und der zugehörigen Konfidenzintervalle auf.

## 7 Diskussion

In dieser Arbeit wurde die Gleichung zur Abschätzung der exakten Gewichte des RE-Modells vereinfacht [1]. Mit der Vereinfachung entfällt die Berechnung der Dichte von Cochran's  $Q$ , so dass nur noch die Berechnung der exakten Verteilungsfunktion von Cochran's  $Q$  anfällt. Dadurch reduziert sich die Rechenzeit und erhöht zudem die numerische Stabilität und reduziert damit auch den Approximationsfehler.

Des Weiteren ermöglicht die exakte Verteilung von Cochran's  $Q$  eine Abschätzung der Variabilität von  $\tau^2$ . Mit ihr besteht die Möglichkeit Konfidenzintervalle für  $\hat{\tau}_{DSL}^2$  vorgenommen. Dieser kohärente Ansatz ist im Vorteil gegenüber anderen Ansätzen, da dieser Ansatz zur Berechnung der exakten Gewichte für das RE-Modell ebenfalls die Verteilung von  $\hat{\tau}_{DSL}^2$ , basierend auf der exakten Verteilung von Cochran's  $Q$ , beinhaltet. D.h. auch der Punktschätzer  $\hat{\tau}_{DSL}^2$  bleibt gleich, während bei den meisten anderen Ansätzen dies nicht gewährleistet ist.

Die Vereinfachung des RE-Modells mit exakten Gewichten wurde in einer Monte-Carlo Simulation intensiv auf ihre Validität überprüft. In den Simulationsstudien I und II wurde der neue Ansatz mit dem klassischen DSL RE-Modell hinsichtlich  $\alpha$ -Fehler und Teststärke verglichen. Ein weiterer Schwerpunkt dieser Simulationsstudien war es die Eigenschaften der Verfahren in Gegenwart einer einzelnen großen Studie mit großem Effekt und vielen kleinen Studien mit kleinerem Effekt zu untersuchen. Dabei war der  $\alpha$ -Fehler erstaunlicherweise stets kleiner für den exakten RE-Ansatz als für den klassischen DSL-Ansatz und zugleich war die Teststärke für den exakten RE-Ansatz meistens größer als für den klassischen DSL-Ansatz. Für dieses Szenario zeigt das exakte RE-Modell eine klare Überlegenheit gegenüber dem klassischen DSL-Ansatz auf.

Dieser Befund erklärt sich durch die Verteilung von  $\hat{\tau}^2$ , die ihrerseits auf der Cochran's  $Q$ -Verteilung basiert. Die Variabilität von  $\hat{\tau}^2$  kann sich speziell für Meta-Analysen mit einer niedrigen Anzahl von Studien stark erhöhen, da ihre Verteilung von den vorhandenen Studien geschätzt werden muss und die Unsicherheit dementsprechend für eine kleine Studienanzahl groß ist. Der neue Ansatz berücksichtigt diese Variabilität. Sie fließt somit in die Berechnung des neuen Effektschätzers mit ein. Des Weiteren sind die Gewichte bei Anwendung des exakten Ansatzes größer als für den klassischen DSL-Ansatz, wenn einzelne überproportional große Studien in die Meta-Analyse einfließen. Dieser Unter-

schied macht sich in Gegenwart von wenigen kleinen Studien dann noch stärker bemerkbar.

Weiterhin zeigen die Simulationsstudien I und II, dass für eine hinreichende Anzahl von Studien, die beiden RE-Ansätze zu ähnlichen Ergebnissen führen. Dies dürfte eine Folge aus der dadurch verbesserten Approximation der Normalverteilung sein. Für zukünftige Arbeiten wäre es dennoch von großem Interesse, den Approximationsfehler für die Verteilung von Cochran's  $Q$  zu untersuchen.

In der Simulationsstudie III wurden der DSL-Ansatz und der exakte RE-Ansatz für Meta-Analysen im Kontext zu HWG betrachtet. Dafür wurden drei Maße aus der Populationsgenetik herangezogen, die eine Abweichung vom Hardy-Weinberg Gleichgewicht nicht nur feststellen, sondern auch quantifizieren können. Hinsichtlich der Simulationsstudie für den  $\alpha$ -Fehler zeichnet sich eine Überlegenheit für das gepoolte REH gegenüber den anderen Maßen ab, da es als einziges Maß das Niveau für alle Allelfrequenzen in allen betrachteten Szenarien einhält.

Bezüglich der Teststärke zeigen alle Maßzahlen ähnliche Ergebnisse. Jedoch ist eine Besonderheit für das REH zu verzeichnen, welches bezüglich der Allelfrequenzen und innerhalb der betrachteten Simulationsszenarien zu nahezu gleichen Ergebnissen führt. Dies ist ein Indiz für eine größere Unabhängigkeit gegenüber den Allelfrequenzen im Vergleich zu den anderen Maßen.

Für den Vergleich bezüglich der Teststärke zwischen dem RE-Modell und dem RE-Modell mit exakten Gewichten zeichnet sich ein heterogenes Bild ab. Ein klares Muster ist nicht zu erkennen. In zukünftigen Arbeiten sollte versucht werden, Merkmale für die jeweilige Richtung der Teststärke für die verschiedenen Ansätze und Maße zu identifizieren um beispielsweise situationsbedingt eine passende Entscheidung für den jeweiligen Ansatz bzw. Maß zu finden.

Da die Maße u.a. eine Abweichung vom HWG quantifizieren können, lassen sich Äquivalenztests auf Basis von Intervall-Inklusionstests konstruieren. Für das REH stehen bereits Methoden zur Verfügung [56]. Für zukünftige Anwendungen bietet dieser Sachverhalt die interessante Möglichkeit an, den tatsächlichen Nachweis der Gleichheit für das HWG innerhalb einer Stichprobe oder in gepoolten Stichproben durch Meta-Analysen zu erbringen.

In der praktischen Anwendung auf Realdaten wurde gezeigt, dass der klassische DSL-Ansatz und der exakte RE-Ansatz zu unterschiedlichen Ergebnissen führen

kann. Diese Diskrepanz ist in der Praxis aber nur bei einer niedrigen Anzahl von Studien und einer großen Heterogenität zwischen den Studien zu erwarten.

Im Zuge dieser Arbeit entstand eine Implementierung der beschriebenen Ansätze. Dies umfasst u.a. das exakte RE-Modell für Meta-Analysen, das klassische  $I^2$ , das exakte trunkierte  $I^2$  sowie der DSL  $\tau^2$  Schätzer mit exaktem Konfidenzintervall. Alle Methoden wurden dem R-Paket `metaxa` (Meta-Analyse mit exakten Gewichten)) mit objektorientierter S4 Programmierung in der Programmiersprache R [3] implementiert und stehen frei zur Verfügung [4].

Die Simulationsstudien haben gezeigt, dass das RE-Modell mit exakten Gewichten im Vergleich zum klassischem DSL RE-Modell im Wesentlichen zu äquivalenten Ergebnissen gelangt. Darüber hinaus kann mit dem exakten RE-Modell in bestimmten Situationen, insbesondere bei hoher Heterogenität zwischen den Studien und einer kleinen Studienanzahl, eine höhere Teststärke erreicht werden. Der kohärente Ansatz zur Berechnung von Konfidenzintervalle für  $\hat{\tau}_{DSL}^2$  bereichert zusätzlich das exakte RE-Modell. Demzufolge sollte, insbesondere für die erwähnten Situationen, dem exakten RE-Modell gegenüber dem klassischen DSL RE-Modell in zukünftigen Anwendungen der Vorzug gegeben werden.

## 8 Zusammenfassung

Das Modell mit zufälligen Effekten ist das Standardmodell bei Meta-Analysen im Fall von Heterogenität zwischen den Studien. Die Heterogenität zwischen den Studien wird unter Verwendung von Cochran's  $Q$  Statistik geschätzt und mit der inversen Varianz-Methode zu den studienspezifischen Varianzen addiert. Auf Basis der Arbeit von Biggerstaff und Jackson [1], die die exakte Verteilung von Cochran's  $Q$  entwickelten, ist es möglich, die Verteilungseigenschaften der Heterogenität zu untersuchen. Somit lässt sich die Variabilität der Heterogenität  $\tau^2$  abschätzen und die Unsicherheit der Schätzung von  $\tau^2$  im Meta-Analyse Modell mit zufälligen Effekten berücksichtigen.

Auf Grundlage dieser Vorarbeiten wird in dieser Arbeit erstmalig die Variabilität von  $\tau^2$  mithilfe der exakten Verteilung von Cochran's  $Q$  im Modell für zufällige Effekte nach DerSimonian und Laird (DSL RE-Modell) für Meta-Analysen berücksichtigt, aus dem sich das Meta-Analyse Modell für zufällige Effekte mit exakten Gewichten (das exakte RE-Modell) ergibt. Zudem wird der ursprüngliche Ausdruck für die Berechnung der asymptotischen Gewichte für das exakte RE-Modell angepasst, so dass eine numerische Berechnung der Dichtefunktion von Cochran's  $Q$  entfällt. Die Vereinfachung führt zu einer Reduzierung der Berechnungszeit, erhöht die numerische Stabilität und reduziert den Approximierungsfehler bei Meta-Analysen. Die exakte Verteilung von Cochran's  $Q$  ermöglicht zudem die Konstruktion von Konfidenzintervalle für das DSL  $\hat{\tau}^2$ . Dies stellt einen kohärenten Ansatz dar, da auch der neue Ansatz die Verteilung von DSL  $\hat{\tau}^2$  mit einbezieht. Die im Zuge dieser Dissertationsschrift näher betrachteten Verfahren sind in dem R-Paket `metaxa` (Meta-Analyse mit exakten Gewichten) mit objektorientierter S4 Programmierung in der Programmiersprache R [3] neu implementiert worden und stehen frei zur Verfügung.

Die Validität und das Leistungsverhalten des neuen Ansatzes werden mit einer umfangreichen Simulationsstudie sowie mit Anwendungen auf Realdaten bestätigt. Der neue Ansatz hält den nominellen Typ I Fehler ein. Hingegen hält die übliche DSL Methode ohne exakte Gewichte in verschiedenen Situationen das nominelle Niveau nicht ein. Der DSL Ansatz mit exakten Gewichten weist im Fall von vielen kleinen Studien und einer großen Studien darüber hinaus eine größere Teststärke auf als die übliche DSL Methode ohne exakte Gewichte. In

einer weiteren Simulationsstudie wird der DSL Ansatz mit exakten Gewichten für Meta-Analysen im Kontext des Hardy-Weinberg Gleichgewichts betrachtet. Dafür werden drei Maße aus der Populationsgenetik herangezogen, die die Stärke der Abweichung vom Hardy-Weinberg Gleichgewicht quantifizieren. Hier zeigt sich, dass die sogenannte relative Exzessheterozygotie (REH) den anderen beiden Maßen gegenüber überlegen ist, da es als einziges Maße das nominelle Signifikanzniveau in allen betrachteten Szenarien einhält.

Die Simulationsstudien zeigen, dass das DSL Modell mit exakten Gewichten in vielen Fällen äquivalente Ergebnisse mit dem klassischen DSL Modell ohne exakte Gewichte liefert, diesem aber in einigen Fällen deutlich überlegen ist. Diese Fälle besitzen insbesondere in der Genetischen Epidemiologie eine große Bedeutung. Ein weiterer Vorteil des DSL Ansatzes mit exakten Gewichten ist, dass er im Vergleich mit dem klassischen DSL Ansatz ohne exakte Gewichte kohärent eine Schätzung von Konfidenzintervallen gestattet. Demzufolge sollte dem DSL Modell mit exakten Gewichten gegenüber dem klassischen DSL Modell ohne exakte Gewichten in zukünftigen Anwendungen verwendet werden.



# Literaturverzeichnis

- [1] Biggerstaff BJ, Jackson D. The exact distribution of Cochran's heterogeneity statistic in one-way random effects meta-analysis. *Stat Med* 2008; 27:6093–6110. DOI: 10.1002/sim.3428.
- [2] DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7:177–188. DOI: 10.1016/0197-2456(86)90046-2.
- [3] Team RC. R: a language and environment for statistical computing, 2013. Vienna. <http://www.R-project.org/>
- [4] Preuß M. R-Paket metaxa, 2013. Lübeck. [www.imbs-luebeck.de/imbs/de/software](http://www.imbs-luebeck.de/imbs/de/software)
- [5] Chalmers I, Hedges LV, Cooper H. A brief history of research synthesis. *Eval Health Prof* 2002; 25:12–37. DOI: 10.1177/0163278702025001003.
- [6] Higgins J, Green S. *Cochrane handbook for systematic reviews of interventions*, Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011.
- [7] Ziegler A, Lange S, Bender R. Systematische Übersichten und Meta-Analysen. *Dtsch Med Wochenschr* 2005; 129:e1. DOI: 10.1055/s-2004-925563.
- [8] Glass GV. Primary, secondary, and meta-analysis of research. *Educational Res* 1976; 5:3–8. DOI: 10.2307/1174772.
- [9] Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T, Friedenreich C. Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int J Epidemiol* 1999; 28:1–9. DOI: 10.1093/ije/28.1.1.
- [10] Hedges LV, Olkin I. *Statistical Methods for Meta-Analysis*, 1985. Academic Press: Orlando.

- [11] Hunter JE, Schmidt FL. *Methods of meta-analysis: correcting error and bias in research findings*, 1990. Sage: Thousand Oaks.
- [12] Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for meta-analysis in medical research*, 2000. Wiley: Chichester.
- [13] Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *PLoS Med* 2009; 6:e1000097. DOI: 10.1371/journal.pmed.1000097.
- [14] von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Preventive Medicine* 2007; 45:247–251. DOI: 10.1016/j.ypmed.2007.08.012.
- [15] The PLoS Medicine Editors. Best practice in systematic reviews: the importance of protocols and registration. *PLoS Med* 2011; 8:e1001009. DOI: 10.1371/journal.pmed.1001009.
- [16] Birge R. The calculation of errors by the method of least squares. *Phys Rev* 1932; 40:207–227. DOI: 10.1103/PhysRev.40.207.
- [17] Cochran WG. Problems arising in the analysis of a series of similar experiments. *J R Stat Soc* 1937; 4:102–118. DOI: 10.2307/2984123.
- [18] Rao PSRS, Kaplan J, Cochran WG. Estimators for the one-way random effects model with unequal error variances. *J Am Stat Assoc* 1981; 76:89–97. DOI: 10.2307/2287050.
- [19] Harville DA. Maximum likelihood approaches to variance component estimation and to related problems. *J Am Stat Assoc* 1977; 72:320–338. DOI: 10.2307/2286796.

- [20] Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiol Rev* 1987; 9:1–30.
- [21] Berkey CS, Hoaglin DC, Mosteller F, Colditz GA. A random-effects regression model for meta-analysis. *Stat Med* 1995; 14:395–411. DOI: 10.1002/sim.4780140406.
- [22] Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? *Stat Med* 2002; 21:1559–1573. DOI: 10.1002/sim.1187.
- [23] Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. An introduction to bayesian methods in health technology assessment. *BMJ* 1999; 508–512. DOI: 10.1136/bmj.319.7208.508.
- [24] Gilks WR, Richardson S, Spiegelhalter DJ. *Markov chain monte carlo in practice*, 1996. Chapman & Hall: London.
- [25] Machin D, Cheung YB, Parmar MKB. *Survival analysis: a practical approach*, 2006. John Wiley and Sons: Chichester.
- [26] Glass GV, McGaw B, Smith ML. *Meta-analysis in social research: individual and neighbourhood reactions*, 1981. Sage Publications: London.
- [27] Lipsey MW, Wilson DB. *Practical meta-analysis*, 2001. Sage Publications: London.
- [28] Cooper H, Hedges LV. *The handbook of research synthesis and meta-analysis*, 1994. Russell Sage Foundation: New York.
- [29] Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *New Engl J Med* 1992; 248–254. DOI: 10.1056/NEJM199207233270406.

- [30] Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med* 1998; 17:2815–2834. DOI: 10.1002/sim.1605.
- [31] Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working Group. *Stat Med* 1995; 14:2057–2079. DOI: 10.1002/sim.4780141902.
- [32] Gueyffier F, Bouitrie F, Boissel JP, Coope J, Cutler J, Ekbom T, et al. IN-DANA: a meta-analysis on individual patient data in hypertension. Protocol and preliminary results. *Therapie* 1995; 50:353–362.
- [33] IQWiG, GMDS, IBS-DR. Stellenwert von Ergebnissen aus indirekten Vergleichen, 2013. [www.gmds.de/pdf/publikationen/stellungnahmen/120202\\_IQWIG\\_GMDS\\_IBS\\_DR.pdf](http://www.gmds.de/pdf/publikationen/stellungnahmen/120202_IQWIG_GMDS_IBS_DR.pdf)
- [34] Terrin N, Schmid CH, Lau J, Olkin I. Adjusting for publication bias in the presence of heterogeneity. *Stat Med* 2003; 22:2113–2126. DOI: 10.1002/sim.1461.
- [35] Higgins J, Thompson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. *J Health Serv Res Policy* 2002; 7:51–61. DOI: 10.1258/1355819021927674.
- [36] Cochran WG. The Combination of Estimates from Different Experiments. *Biometrics* 1954; 10:101–129. DOI: 10.2307/3001666.
- [37] Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. *Stat Med* 1998; 17:841–856. DOI: 10.1002/(SICI)1097-0258(19980430)17:8<841::AID-SIM781>3.0.CO;2-D pp.841 - 856.
- [38] Paul SR, Donner A. Small sample performance of tests of homogeneity of odds ratios in  $K \times 2$  tables. *Stat Med* 1992; 11:159–165. DOI: 10.1002/sim.4780110203.

- [39] Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; 21:1539–1558. DOI: 10.1002/sim.1186.
- [40] Rücker G, Schwarzer G, Carpenter J, Schumacher M. Undue reliance on  $I^2$  in assessing heterogeneity may mislead. *BMC Med Res Methodol* 2008; 8:79. DOI: 10.1186/1471-2288-8-79.
- [41] Higgins JPT. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327:557–560. DOI: 10.1136/bmj.327.7414.557.
- [42] Mittlböck M, Heinzl H. A simulation study comparing properties of heterogeneity measures in meta-analyses. *Stat Med* 2006; 25:4321–4333. DOI: 10.1002/sim.2692.
- [43] Higgins JPT. Commentary: heterogeneity in meta-analysis should be expected and appropriately quantified. *Int J Epidemiol* 2008; 37:1158–1160. DOI: 10.1093/ije/dyn204.
- [44] Ziegler A, König IR, Pahlke F. *A Statistical approach to genetic epidemiology*, 2010. Wiley-VCH: Weinheim.
- [45] Biggerstaff BJ, Tweedie RL. Incorporating variability in estimates of heterogeneity in the random effects model in meta-analysis. *Stat Med* 1997; 16:753–768. DOI: 10.1002/(SICI)1097-0258(19970415)16:7<753::AID-SIM494>3.0.CO;2-G.
- [46] Farebrother RW. Algorithm AS 204: the distribution of a positive linear combination of  $\chi^2$  random variables. *Appl Stat-J Roy St C* 1984; 33:332–339. DOI: 10.2307/2347721.
- [47] Imhof JP. Computing the distribution of quadratic forms in normal variables. *Biometrika* 1961; 48:419–426. DOI: 10.1093/biomet/48.3-4.419.
- [48] Duchesne P, Lafaye De Micheaux P. Computing the distribution of quadratic forms: Further comparisons between the Liu–Tang–Zhang approxima-

- tion and exact methods. *Comput Stat Data An* 2010; 54:858–862. DOI: 10.1016/j.csda.2009.11.025.
- [49] Muenchen RA. The popularity of data analysis software, 2013. <http://r4stats.com/articles/popularity/>
- [50] Collins R, Yusuf S, Peto R. Overview of randomised trials of diuretics in pregnancy. *BMJ* 1985; 290:17–23. DOI: 10.1136/bmj.290.6470.788-d.
- [51] Thorlund K, Wetterslev J, Awad T, Thabane L, Gluud C. Comparison of statistical inferences from the DerSimonian–Laird and alternative random-effects model meta-analyses – an empirical assessment of 920 Cochrane primary outcome meta-analyses. *Res Syn* 2011; 2:238–253. DOI: 10.1002/jrsm.53.
- [52] Viechtbauer W. Confidence intervals for the amount of heterogeneity in meta-analysis. *Stat Med* 2007; 26:37–52. DOI: 10.1002/sim.2514.
- [53] Minelli C, Thompson JR, Abrams KR, Thakkinstian A, Attia J. How should we use information about HWE in the meta-analyses of genetic association studies? *Int J Epidemiol* 2008; 37:136–146. DOI: 10.1093/ije/dym234.
- [54] Schaid DJ, Jacobsen SJ. Biased tests of association: comparisons of allele frequencies when departing from Hardy-Weinberg proportions. *Am J Hum Genet* 1999; 149:706–711.
- [55] Schaid DJ, Batzler AJ, Jenkins GD, Hildebrandt MAT. Exact tests of Hardy-Weinberg equilibrium and homogeneity of disequilibrium across strata. *Am J Hum Genet* 2006; 79:1071–1080. DOI: 10.1086/510257.
- [56] Wellek S, Goddard KAB, Ziegler A. A confidence-limit-based approach to the assessment of Hardy-Weinberg equilibrium. *BMJ* 2010; 52:253–270. DOI: 10.1002/bimj.200900249.

- [57] Malécot G, Yermanos DM. The mathematics of heredity, 1969. Freeman: San Francisco.
- [58] Hedrick PW. Genetics of populations, 2011. Jones & Bartlett Publishers: Sudbury.
- [59] Jiang R, Dong J, Wang D, Sun FZ. Fine-scale mapping using Hardy–Weinberg disequilibrium. *Ann Hum Genet* 2001; 65:207–219. DOI: 10.1046/j.1469-1809.2001.6520207.x.
- [60] Ziegler A, Van Steen K, Wellek S. Investigating Hardy–Weinberg equilibrium in case–control or cohort studies or meta-analysis. *Breast Cancer Res Tr* 2010; 128:197–201. DOI: 10.1007/s10549-010-1295-z.
- [61] Pereira C, Rogatko A. The Hardy–Weinberg equilibrium under a Bayesian perspective. *Brazil J Genetics* 1984; 7:689–707.
- [62] Lindley DV. Statistical inference concerning Hardy-Weinberg equilibrium, 1987. Institute of Statistics and Decision Sciences, Duke University.
- [63] Elston RC, Forthofer R. Testing for Hardy-Weinberg equilibrium in small samples. *Biometrics* 1977; 33:536–542. DOI: 10.2307/2529370.
- [64] Olson JM. Testing the Hardy-Weinberg law across strata. *Ann Hum Genet* 1993; 291–295. DOI: 10.1111/j.1469-1809.1993.tb00903.x.
- [65] Thompson SG, Pocock SJ. Can meta-analyses be trusted? *Lancet* 1991; 338:1127–1130. DOI: 10.1016/0140-6736(91)91975-Z.
- [66] Guolo A, Varin C. The R package metaLik for likelihood inference in meta-analysis. *J Stat Softw* 2012; 21:979–980. DOI: 10.1080/10618600.2012.754272.
- [67] Sørensen H. Small sample distribution of the likelihood ratio test in the random effects model. *J Stat Plan Infer* 2008; 138:1605–1614. DOI: 10.1016/j.jspi.2007.05.047.

- [68] Shepperd S, Doll H, Broad J, Gladman J, Iliffe S, Langhorne P, et al. Hospital at home early discharge. *Cochrane DB Syst Rev* 2009; 1:CD000356. DOI: 10.1002/14651858.CD000356.pub3.
- [69] Laroia N, Phelps DL, Roy J. Double wall versus single wall incubator for reducing heat loss in very low birth weight infants in incubators. *Cochrane DB Syst Rev* 2007; 2:CD004215. DOI: 10.1002/14651858.CD004215.pub2.
- [70] Frank B, Rigas SH, Bermejo JL, Wiestler M, Wagner K, Hemminki K, et al. The CASP8 -652 6N del promoter polymorphism and breast cancer risk: a multicenter study. *Breast Cancer Res Treat* 2008; 139–144. DOI: 10.1007/s10549-007-9752-z.
- [71] Satterthwaite FE. An approximate distribution of estimates of variance components. *Biometrics Bull* 1946; 2:110–114. DOI: 10.2307/3002019.



# Danksagung

An erster Stelle richtet sich mein Dank an Herrn Univ.-Prof. Dr. Andreas Ziegler für die umfassende wissenschaftliche Betreuung und die Möglichkeit, neben der Dissertation an verschiedenen spannenden Forschungsprojekten mitzuarbeiten.

Frau Dr. Claudia Hemmelmann und Herrn Dr. Arne Schillert möchte ich für die vielen hilfreichen und kreativen Diskussionen sowie zahlreichen fachlichen Hinweise danken.

Ein großes Dankeschön geht an alle Mitarbeiterinnen und Mitarbeiter des Instituts für Medizinische Biometrie und Statistik für die hervorragende Zusammenarbeit und Hilfestellung bei allen möglichen Problemen. Besonders möchte ich mich bei Dipl. Math. oec. Andrea Senft für die Namensgebung meines R-Pakets bedanken.

Mein Dank gebührt außerdem meinen Eltern Birgit und Werner Preuß sowie meiner Schwester Esther Preuß für die moralische Unterstützung.

# Lebenslauf

## Personalien

Name	Michael Preuß
Geburtsdatum	4. Juli 1979
Geburtsort	Würselen
Familienstand	ledig
Nationalität	deutsch



## Berufliche Tätigkeiten

seit 11/2007	Wissenschaftlicher Mitarbeiter im Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck
07/2007 – 10/2007	Wissenschaftliche Hilfskraft am Institut für Tropenhygiene und öffentliches Gesundheitswesen, Universität Heidelberg

## Hochschulausbildung

seit 11/2007	Promotionsstudium an der Universität zu Lübeck
06/2007	Diplom der Medizinischen Informatik, Universität Heidelberg & Hochschule Heilbronn
04/2004	Vordiplom
10/2000	Studium der Medizinischen Informatik

## Schulbildung

1999	Abitur
1997 – 1999	Mercator Berufskolleg, Moers

# Publikationsverzeichnis

## Zeitschriftenartikel

- [1] **Preuß M**, Ziegler A. A simplification and implementation of random-effects meta-analyses based on the exact distribution of Cochran's  $Q$ . *Method Inform Med* 2013; (im Druck).
- [2] **Preuss M**, König IR, Thompson JR, Erdmann J, Absher D, Assimes TL, Blankenberg S, Boerwinkle E, Chen L, Cupples LA, Hall AS, Halperin E, Hengstenberg C, Holm H, Laaksonen R, Li M, März W, McPherson R, Musunuru K, Nelson CP, Burnett MS, Epstein SE, O'Donnell CJ, Quertermous T, Rader DJ, Roberts R, Schillert A, Stefansson K, Stewart AFR, Thorleifsson G, Voight BF, Wells GA, Ziegler A, Kathiresan S, Reilly MP, Samani NJ, Schunkert H, CARDIoGRAM Consortium. Design of the Coronary ARtery DIsease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) study: a genome-wide association meta-analysis involving more than 22 000 cases and 60 000 controls. *Circ Cardiovasc Genet* 2010; 3:475–483. DOI: 10.1161/CIRCGENETICS.109.899443.
- [3] Keller A, Westenberger A, Sobrido MJ, García-Murias M, Domingo A, Sears RL, Lemos RR, Ordoñez-Ugalde A, Nicolas G, da Cunha JEG, Rushing EJ, Hugelshofer M, Wurnig MC, Kaech A, Reimann R, Lohmann K, Dobričić V, Carracedo A, Petrović I, Miyasaki JM, Abakumova I, Mäe MA, Raschperger E, Zatz M, Zschiedrich K, Klepper J, Spiteri E, Prieto JM, Navas I, **Preuss M**, Dering C, Janković M, Paucar M, Svenningsson P, Saliminejad K, Khorshid HRK, Novaković I, Aguzzi A, Boss A, Le Ber I, Defer G, Hannequin D, Kostić VS, Champion D, Geschwind DH, Coppola G, Betsholtz C, Klein C, Oliveira JRM. Mutations in the gene encoding PDGF-B cause brain calcifications in humans and mice. *Nat Genet* 2013; 45:1077–1082. DOI: 10.1038/ng.2723.
- [4] Berndt SI, Gustafsson S, Mägi R, Ganna A, Wheeler E, Feitosa MF, Justice AE, Monda KL, Croteau-Chonka DC, Day FR, Esko T, Fall T, Ferreira T, Gentilini D, Jackson AU, Luan J, Randall JC, Vedantam S, Willer CJ, Winkler TW, Wood AR, Workalemahu T, Hu YJ, Lee SH, Liang L, Lin DY, Min JL, Neale BM, Thorleifsson G, Yang J, Albrecht E, Amin N, Bragg-Gresham JL, Cad-

by G, den Heijer M, Eklund N, Fischer K, Goel A, Hottenga JJ, Huffman JE, Jarick I, Johansson A, Johnson T, Kanoni S, Kleber ME, König IR, Kristiansson K, Kutalik Z, Lamina C, Lecoeur C, Li G, Mangino M, McArdle WL, Medina-Gomez C, Müller-Nurasyid M, Ngwa JS, Nolte IM, Paternoster L, Pechlivanis S, Perola M, Peters MJ, **Preuss M**, Rose LM, Shi J, Shungin D, Smith AV, Strawbridge RJ, Surakka I, Teumer A, Trip MD, Tyrer J, Van Vliet-Ostaptchouk JV, Vandenput L, Waite LL, Zhao JH, Absher D, Asselbergs FW, Atalay M, Attwood AP, Balmforth AJ, Basart H, Beilby J, Bonnycastle LL, Brambilla P, Bruinenberg M, Campbell H, Chasman DI, Chines PS, Collins FS, Connell JM, Cookson WO, de Faire U, de Vegt F, Dei M, Dimitriou M, Edkins S, Estrada K, Evans DM, Farrall M, Ferrario MM, Ferrières J, Franke L, Frau F, Gejman PV, Grallert H, Grönberg H, Gudnason V, Hall AS, Hall P, Hartikainen AL, Hayward C, Heard-Costa NL, Heath AC, Hebebrand J, Homuth G, Hu FB, Hunt SE, Hyppönen E, Iribarren C, Jacobs KB, Jansson JO, Jula A, Kähönen M, Kathiresan S, Kee F, Khaw KT, Kivimäki M, Koenig W, Kraja AT, Kumari M, Kuulasmaa K, Kuusisto J, Laitinen JH, Lakka TA, Langenberg C, Launer LJ, Lind L, Lindström J, Liu J, Liuzzi A, Lokki ML, Lorentzon M, Madden PA, Magnusson PK, Manunta P, Marek D, März W, Mateo Leach I, McKnight B, Medland SE, Mihailov E, Milani L, Montgomery GW, Mooser V, Mühleisen TW, Munroe PB, Musk AW, Narisu N, Navis G, Nicholson G, Nohr EA, Ong KK, Oostra BA, Palmer CNA, Palotie A, Peden JF, Pedersen N, Peters A, Polasek O, Pouta A, Pramstaller PP, Prokopenko I, Pütter C, Radhakrishnan A, Raitakari O, Rendon A, Rivadeneira F, Rudan I, Saaristo TE, Sambrook JG, Sanders AR, Sanna S, Saramies J, Schipf S, Schreiber S, Schunkert H, Shin SY, Signorini S, Sinisalo J, Skrobek B, Soranzo N, Stančáková A, Stark K, Stephens JC, Stirrups K, Stolk RP, Stumvoll M, Swift AJ, Theodoraki EV, Thorand B, Tregouet DA, Tremoli E, Van der Klauw MM, van Meurs JBJ, Vermeulen SH, Viikari J, Virtamo J, Vitart V, Waeber G, Wang Z, Widén E, Wild SH, Willemsen G, Winkelmann BR, Witteman JCM, Wolffenbuttel BHR, Wong A, Wright AF, Zillikens MC, Amouyel P, Boehm BO, Boerwinkle E, Boomsma DI, Caulfield MJ, Chanock SJ, Cupples LA, Cusi D, Dedoussis GV, Erdmann J, Eriksson JG, Franks PW, Froguel P, Gieger C, Gyllensten U, Hamsten A, Harris TB, Hengstenberg C, Hicks AA, Hingorani A, Hinney A. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture.

Nat Genet 2013; 45:501–512. DOI: 10.1038/ng.2606.

- [5] Yang J, Loos RJJ, Powell JE, Medland SE, Speliotes EK, Chasman DI, Rose LM, Thorleifsson G, Steinthorsdottir V, Mägi R, Waite L, Smith AV, Yerges-Armstrong LM, Monda KL, Hadley D, Mahajan A, Li G, Kapur K, Vitart V, Huffman JE, Wang SR, Palmer C, Esko T, Fischer K, Zhao JH, Demirkan A, Isaacs A, Feitosa MF, Luan J, Heard-Costa NL, White C, Jackson AU, **Preuss M**, Ziegler A, Eriksson J, Kutalik Z, Frau F, Nolte IM, Van Vliet-Ostaptchouk JV, Hottenga JJ, Jacobs KB, Verweij N, Goel A, Medina-Gomez C, Estrada K, Bragg-Gresham JL, Sanna S, Sidore C, Tyrer J, Teumer A, Prokopenko I, Mangino M, Lindgren CM, Assimes TL, Shuldiner AR, Hui J, Beilby JP, McArdle WL, Hall P, Haritunians T, Zgaga L, Kolcic I, Polasek O, Zemunik T, Oostra BA, Junttila MJ, Grönberg H, Schreiber S, Peters A, Hicks AA, Stephens J, Foad NS, Laitinen J, Pouta A, Kaakinen M, Willemsen G, Vink JM, Wild SH, Navis G, Asselbergs FW, Homuth G, John U, Iribarren C, Harris T, Launer L, Gudnason V, O'Connell JR, Boerwinkle E, Cadby G, Palmer LJ, James AL, Musk AW, Ingelsson E, Psaty BM, Beckmann JS, Waeber G, Vollenweider P, Hayward C, Wright AF, Rudan I, Groop LC, Metspalu A, Khaw KT, van Duijn CM, Borecki IB, Province MA, Wareham NJ, Tardif JC, Huikuri HV, Cupples LA, Atwood LD, Fox CS, Boehnke M, Collins FS, Mohlke KL, Erdmann J, Schunkert H, Hengstenberg C, Stark K, Lorentzon M, Ohlsson C, Cusi D, Staessen JA, Van der Klauw MM, Pramstaller PP, Kathiresan S, Jolley JD, Ripatti S, Jarvelin MR, de Geus EJC, Boomsma DI, Penninx B, Wilson JF, Campbell H, Chanock SJ, van der Harst P, Hamsten A, Watkins H, Hofman A, Witteman JC, Zillikens MC, Uitterlinden AG, Rivadeneira F, Zillikens MC, Kiemeneij LA, Vermeulen SH, Abecasis GR, Schlessinger D, Schipf S, Stumvoll M, Tönjes A, Spector TD, North KE, Lettre G, McCarthy MI, Berndt SI, Heath AC, Madden PAF, Nyholt DR, Montgomery GW, Martin NG, McKnight B, Strachan DP, Hill WG, Snieder H, Ridker PM, Thorsteinsdottir U, Stefansson K, Frayling TM, Hirschhorn JN, Goddard ME, Visscher PM. FTO genotype is associated with phenotypic variability of body mass index. *Nature* 2012; 490:267–272. DOI: 10.1038/nature11401.

- [6] Altmann A, Weber P, Bader D, **Preuss M**, Binder EB, Müller-Myhsok B. A beginners guide to SNP calling from high-throughput DNA-sequencing data. *Hum Genet* 2012; 131:1541–1554. DOI: 10.1007/s00439-012-1213-z.
- [7] Demirkan A, van Duijn CM, Ugocsai P, Isaacs A, Pramstaller PP, Liebisch G, Wilson JF, Johansson A, Rudan I, Aulchenko YS, Kirichenko AV, Janssens ACJW, Jansen RC, Gnewuch C, Domingues FS, Pattaro C, Wild SH, Jonasson I, Polasek O, Zorkoltseva IV, Hofman A, Karssen LC, Struchalin M, Floyd J, Igl W, Biloglav Z, Broer L, Pfeufer A, Pichler I, Campbell S, Zaboli G, Kolcic I, Rivadeneira F, Huffman J, Hastie ND, Uitterlinden A, Franke L, Franklin CS, Vitart V, DIAGRAM Consortium, Nelson CP, **Preuss M**, CARDIoGRAM Consortium, Bis JC, O'Donnell CJ, Franceschini N, CHARGE Consortium, Witteman JCM, Axenovich T, Oostra BA, Meitinger T, Hicks AA, Hayward C, Wright AF, Gyllenstein U, Campbell H, Schmitz G, EUROSPAN consortium. Genome-wide association study identifies novel loci associated with circulating phospho- and sphingolipid concentrations. *PLoS Genet* 2012; 8:e1002490. DOI: 10.1371/journal.pgen.1002490.
- [8] Herder C, Peeters W, Illig T, Baumert J, de Kleijn DPV, Moll FL, Poschen U, Klopp N, Müller-Nurasyid M, Roden M, **Preuss M**, CARDIoGRAM Consortium, Karakas M, Meisinger C, Thorand B, Pasterkamp G, Koenig W, Assimes TL, Deloukas P, Erdmann J, Holm H, Kathiresan S, König IR, McPherson R, Reilly MP, Roberts R, Samani NJ, Schunkert H, Stewart AFR. RANTES/CCL5 and risk for coronary events: results from the MONICA/KORA Augsburg case-cohort, Athero-Express and CARDIoGRAM studies. *Plos One* 2011; 6:e25734. DOI: 10.1371/journal.pone.0025734.
- [9] Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T, Johnson AD, Bochud M, Rice KM, Henneman P, Smith AV, Ehret GB, Amin N, Larson MG, Mooser V, Hadley D, Dörr M, Bis JC, Aspelund T, Esko T, Janssens ACJW, Zhao JH, Heath S, Laan M, Fu J, Pistis G, Luan J, Arora P, Lucas G, Pirastu N, Pichler I, Jackson AU, Webster RJ, Zhang F, Peden JF, Schmidt H, Tanaka T, Campbell H, Igl W, Milaneschi Y, Hottenga JJ, Vitart V, Chasman DI, Trompet S, Bragg-Gresham JL, Alizadeh BZ, Chambers JC, Guo X, Lehtimäki T, Kühnel B, Lopez LM, Polašek O, Boban M, Nelson CP, Morrison AC, Pihur V, Ganesh

SK, Hofman A, Kundu S, Mattace-Raso FUS, Rivadeneira F, Sijbrands EJG, Uitterlinden AG, Hwang SJ, Vasani RS, Wang TJ, Bergmann S, Vollenweider P, Waeber G, Laitinen J, Pouta A, Zitting P, McArdle WL, Kroemer HK, Völker U, Völzke H, Glazer NL, Taylor KD, Harris TB, Alavere H, Haller T, Keis A, Tammesoo ML, Aulchenko Y, Barroso I, Khaw KT, Galan P, Hercberg S, Lathrop M, Eyheramendy S, Org E, Söber S, Lu X, Nolte IM, Penninx BW, Corre T, Masciullo C, Sala C, Groop L, Voight BF, Melander O, O'Donnell CJ, Salomaa V, d'Adamo AP, Fabretto A, Faletra F, Ulivi S, Del Greco FM, Facheris M, Collins FS, Bergman RN, Beilby JP, Hung J, Musk AW, Mangino M, Shin SY, Soranzo N, Watkins H, Goel A, Hamsten A, Gider P, Loitfelder M, Zeginigg M, Hernandez D, Najjar SS, Navarro P, Wild SH, Corsi AM, Singleton A, de Geus EJC, Willemsen G, Parker AN, Rose LM, Buckley B, Stott D, Orru M, Uda M, LifeLines Cohort Study, van der Klauw MM, Zhang W, Li X, Scott J, Chen YDI, Burke GL, Kähönen M, Viikari J, Döring A, Meitinger T, Davies G, Starr JM, Emilsson V, Plump A, Lindeman JH, Hoen PACT, König IR, EchoGen consortium, Felix JF, Clarke R, Hopewell JC, Ongen H, Breteler M, Debette S, Destefano AL, Fornage M, AortaGen Consortium, Mitchell GF, CHARGE Consortium Heart Failure Working Group, Smith NL, KidneyGen consortium, Holm H, Stefansson K, Thorleifsson G, Thorsteinsdottir U, CKDGen consortium, Cardiogenics consortium, CardioGram, Samani NJ, **Preuss M**, Rudan I, Hayward C, Deary IJ, Wichmann HE, Raitakari OT, Palmas W, Kooner JS, Stolk RP, Jukema JW, Wright AF, Boomsma DI, Bandinelli S, Gyllenstein UB, Wilson JF, Ferrucci L, Schmidt R, Farrall M, Spector TD, Palmer LJ, Tuomilehto J, Pfeufer A, Gasparini P, Siscovick D, Altshuler D, Loos RJJ, Toniolo D, Snieder H, Gieger C, Meneton P, Wareham NJ, Oostra BA, Metspalu A, Launer L, Rettig R, Strachan DP, Beckmann JS, Witteman JCM, Erdmann J, van Dijk KW, Boerwinkle E, Boehnke M, Ridker PM, Jarvelin MR, Chakravarti A, Abecasis GR, Gudnason V, Newton-Cheh C, Levy D, Munroe PB, Psaty BM, Caulfield MJ, Rao DC, Tobin MD, Elliott P, van Duijn CM. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nat Genet* 2011; 43:1005–1011. DOI: 10.1038/ng.922.

- [10] Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, **Preuss M**, Stewart AFR, Barbalic M, Gieger C, Absher D, Aherrahrou Z,

Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I, Cardiogenics, Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NEE, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJP, Khaw KT, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Mühleisen TW, Muhlestein JB, Münzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nöthen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schäfer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz SM, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoop JD, Soranzo N, Spertus JA, Stark K, Stirrups K, Stoll M, Tang WHW, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rij AM, Voight BF, Wareham NJ, Wells GA, Wichmann HE, Wild PS, Willenborg C, Witteman JCM, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, März W, Hengstenberg C, Blankenberg S, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann J, CARDIoGRAM Consortium, Samani NJ. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* 2011; 43:333–338. DOI: 10.1038/ng.784.

- [11] Erdmann J, Willenborg C, Nahrstaedt J, **Preuss M**, König IR, Baumert J, Linsel-Nitschke P, Gieger C, Tennstedt S, Belcredi P, Aherrahrou Z, Klopp N, Loley C, Stark K, Hengstenberg C, Bruse P, Freyer J, Wagner AK, Medack A, Lieb W, Grosshennig A, Sager HB, Reinhardt A, Schäfer A, Schreiber S, El Mokhtari NE, Raaz-Schrauder D, Illig T, Garlachs CD, Ekici AB, Reis A, Schrezenmeir J, Rubin D, Ziegler A, Wichmann HE, Doering A, Meisinger C,



- Meitinger T, Peters A, Schunkert H. Genome-wide association study identifies a new locus for coronary artery disease on chromosome 10p11.23. *Eur Heart J* 2011; 32:158–168. DOI: 10.1093/eurheartj/ehq405.
- [12] Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Lango Allen H, Lindgren CM, Luan J, Mägi R, Randall JC, Vedantam S, Winkler TW, Qi L, Workalemahu T, Heid IM, Steinthorsdottir V, Stringham HM, Weedon MN, Wheeler E, Wood AR, Ferreira T, Weyant RJ, Segrè AV, Estrada K, Liang L, Nemesh J, Park JH, Gustafsson S, Kilpeläinen TO, Yang J, Bouatia-Naji N, Esko T, Feitosa MF, Kutalik Z, Mangino M, Raychaudhuri S, Scherag A, Smith AV, Welch R, Zhao JH, Aben KK, Absher DM, Amin N, Dixon AL, Fisher E, Glazer NL, Goddard ME, Heard-Costa NL, Hoesel V, Hottenga JJ, Johansson A, Johnson T, Ketkar S, Lamina C, Li S, Moffatt MF, Myers RH, Narisu N, Perry JRB, Peters MJ, **Preuss M**, Ripatti S, Rivadeneira F, Sandholt C, Scott LJ, Timpson NJ, Tyrer JP, van Wingerden S, Watanabe RM, White CC, Wiklund F, Barlassina C, Chasman DI, Cooper MN, Jansson JO, Lawrence RW, Pellikka N, Prokopenko I, Shi J, Thiering E, Alavere H, Alibrandi MTS, Almgren P, Arnold AM, Aspelund T, Atwood LD, Balkau B, Balmforth AJ, Bennett AJ, Ben-Shlomo Y, Bergman RN, Bergmann S, Biebermann H, Blakemore AIF, Boes T, Bonnycastle LL, Bornstein SR, Brown MJ, Buchanan TA, Busonero F, Campbell H, Cappuccio FP, Cavalcanti-Proença C, Chen YDI, Chen CM, Chines PS, Clarke R, Coin L, Connell J, Day INM, den Heijer M, Duan J, Ebrahim S, Elliott P, Elosua R, Eiriksdottir G, Erdos MR, Eriksson JG, Facheris MF, Felix SB, Fischer-Posovszky P, Folsom AR, Friedrich N, Freimer NB, Fu M, Gaget S, Gejman PV, Geus EJC, Gieger C, Gjesing AP, Goel A, Goyette P, Grallert H, Grässler J, Greenawalt DM, Groves CJ, Gudnason V, Guiducci C, Hartikainen AL, Hassanali N, Hall AS, Havulinna AS, Hayward C, Heath AC, Hengstenberg C, Hicks AA, Hineney A, Hofman A, Homuth G, Hui J, Igl W, Iribarren C, Isomaa B, Jacobs KB, Jarick I, Jewell E, John U, Jørgensen T, Jousilahti P, Jula A, Kaakinen M, Kajantie E, Kaplan LM, Kathiresan S, Kettunen J, Kinnunen L, Knowles JW, Kolcic I, König IR, Koskinen S, Kovacs P, Kuusisto J, Kraft P, Kvaløy K, Laitinen J, Lantieri O, Lanzani C, Launer LJ, Lecoeur C, Lehtimäki T, Lettre G, Liu J, Lokki ML, Lorentzon M, Luben RN, Ludwig B, MAGIC, Manunta P, Marek D, Marre M, Martin NG, McArdle WL, McCarthy A, McK-

- night B, Meitinger T, Melander O, Meyre D, Midthjell K, Montgomery GW, Morken MA, Morris AP, Mulic R, Ngwa JS, Nelis M, Neville MJ, Nyholt DR, O'Donnell CJ, O'Rahilly S, Ong KK, Oostra B, Paré G, Parker AN, Perola M, Pichler I, Pietiläinen KH, Platou CGP, Polasek O, Pouta A, Rafelt S, Raitakari O, Rayner NW, Ridderstråle M, Rief W, Ruukonen A, Robertson NR, Rzehak P, Salomaa V, Sanders AR, Sandhu MS, Sanna S, Saramies J, Savolainen MJ, Scherag S, Schipf S, Schreiber S, Schunkert H, Silander K, Sinisalo J, Siscovick D. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 2010; 42:937–948. DOI: 10.1038/ng.686.
- [13] Assimes TL, Hólm H, Kathiresan S, Reilly MP, Thorleifsson G, Voight BF, Erdmann J, Willenborg C, Vaidya D, Xie C, Patterson CC, Morgan TM, Burnett MS, Li M, Hlatky MA, Knowles JW, Thompson JR, Absher D, Iribarren C, Go A, Fortmann SP, Sidney S, Risch N, Tang H, Myers RM, Berger K, Stoll M, Shah SH, Thorgeirsson G, Andersen K, Havulinna AS, Herrera JE, Faraday N, Kim Y, Kral BG, Mathias RA, Ruczinski I, Suktitipat B, Wilson AF, Yanek LR, Becker LC, Linsel-Nitschke P, Lieb W, König IR, Hengstenberg C, Fischer M, Stark K, Reinhard W, Winogradow J, Grassl M, Grosshennig A, **Preuss M**, Schreiber S, Wichmann HE, Meisinger C, Yee J, Friedlander Y, Do R, Meigs JB, Williams G, Nathan DM, MacRae CA, Qu L, Wilensky RL, Matthai J William H, Qasim AN, Hakonarson H, Pichard AD, Kent KM, Satler L, Lindsay JM, Waksman R, Knouff CW, Waterworth DM, Walker MC, Mooser VE, Marrugat J, Lucas G, Subirana I, Sala J, Ramos R, Martinelli N, Olivieri O, Trabetti E, Malerba G, Pignatti PF, Guiducci C, Mirel D, Parkin M, Hirschhorn JN, Asselta R, Duga S, Musunuru K, Daly MJ, Purcell S, Eifert S, Braund PS, Wright BJ, Balmforth AJ, Ball SG, Myocardial Infarction Genetics Consortium, Wellcome Trust Case Control Consortium, Cardiogenics, Ouwehand WH, Deloukas P, Scholz M, Cambien F, Hogue A, Scheffold T, Salomaa V, Girelli D, Granger CB, Peltonen L, McKeown PP, Altshuler D, Melander O, Devaney JM, Epstein SE, Rader DJ, Elosua R, Engert JC, Anand SS, Hall AS, Ziegler A, O'Donnell CJ, Spertus JA, Siscovick D, Schwartz SM, Becker D, Thorsteinsdottir U, Stefansson K, Schunkert H, Samani NJ, Quertemous T. Lack of association between the Trp719Arg polymorphism in kinesin-like protein-6 and coronary artery disease in 19 case-control studies. *J Am Coll Cardiol* 2010; 56:1552–1563. DOI: 10.1016/j.jacc.2010.06.022.

- [14] Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S, Ferreira T, Wood AR, Weyant RJ, Segrè AV, Speliotes EK, Wheeler E, Soranzo N, Park JH, Yang J, Gudbjartsson D, Heard-Costa NL, Randall JC, Qi L, Vernon Smith A, Mägi R, Pastinen T, Liang L, Heid IM, Luan J, Thorleifsson G, Winkler TW, Goddard ME, Sin Lo K, Palmer C, Workalemahu T, Aulchenko YS, Johansson A, Zillikens MC, Feitosa MF, Esko T, Johnson T, Ketkar S, Kraft P, Mangino M, Prokopenko I, Absher D, Albrecht E, Ernst F, Glazer NL, Hayward C, Hottenga JJ, Jacobs KB, Knowles JW, Kutalik Z, Monda KL, Polasek O, **Preuss M**, Rayner NW, Robertson NR, Steinthorsdottir V, Tyrer JP, Voight BF, Wiklund F, Xu J, Zhao JH, Nyholt DR, Pellikka N, Perola M, Perry JRB, Surakka I, Tammesoo ML, Altmaier EL, Amin N, Aspelund T, Bhangale T, Boucher G, Chasman DI, Chen C, Coin L, Cooper MN, Dixon AL, Gibson Q, Grundberg E, Hao K, Juhani Juntila M, Kaplan LM, Kettunen J, König IR, Kwan T, Lawrence RW, Levinson DF, Lorentzon M, McKnight B, Morris AP, Müller M, Suh Ngwa J, Purcell S, Rafelt S, Salem RM, Salvi E, Sanna S, Shi J, Sovio U, Thompson JR, Turchin MC, Vandenput L, Verlaan DJ, Vitart V, White CC, Ziegler A, Almgren P, Balmforth AJ, Campbell H, Citterio L, De Grandi A, Dominiczak A, Duan J, Elliott P, Elosua R, Eriksson JG, Freimer NB, Geus EJC, Glorioso N, Haiqing S, Hartikainen AL, Havulinna AS, Hicks AA, Hui J, Igl W, Illig T, Jula A, Kajantie E, Kilpeläinen TO, Koiranen M, Kolcic I, Koskinen S, Kovacs P, Laitinen J, Liu J, Lokki ML, Marusic A, Maschio A, Meitinger T, Mulas A, Paré G, Parker AN, Peden JF, Petersmann A, Pichler I, Pietiläinen KH, Pouta A, Ridderstråle M, Rotter JI, Sambrook JG, Sanders AR, Schmidt CO, Sinisalo J, Smit JH, Stringham HM, Bragi Walters G, Widen E, Wild SH, Willemsen G, Zagato L, Zgaga L, Zitting P, Alavere H, Farrall M, McArdle WL, Nelis M, Peters MJ, Ripatti S, van Meurs JBJ, Aben KK, Ardlie KG, Beckmann JS, Beilby JP, Bergman RN, Bergmann S, Collins FS, Cusi D, den Heijer M, Eiriksdottir G, Gejman PV, Hall AS, Hamsten A, Huikuri HV, Iribarren C, Kähönen M, Kaprio J, Kathiresan S, Kiemenev L, Kocher T, Launer LJ, Lehtimäki T, Melander O, Mosley J Tom H, Musk AW, Nieminen MS, O'Donnell CJ, Ohlsson C, Oostra B, Palmer LJ, Raitakari O, Ridker PM, Rioux JD, Rissanen A, Rivolta C, Schunkert H, Shuldiner AR, Siscovick DS, Stumvoll M, Tönjes A, Tuomilehto J, van Ommen GJ, Viikari J, Heath AC, Martin NG, Montgomery GW, Province MA, Kayser M, Arnold

- AM, Atwood LD, Boerwinkle E, Chanock SJ, Deloukas P, Gieger C, Grönberg H, Hall P, Hattersley AT, Hengstenberg C, Hoffman W, Lathrop GM, Salomaa V, Schreiber S, Uda M, Waterworth D, Wright AF, Assimes TL, Barroso I, Hofman A, Moh. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010; 467:832–838. DOI: 10.1038/nature09410.
- [15] Kynast-Wolf G, Preuß M, Sié A, Kouyaté B, Becher H. Seasonal patterns of cardiovascular disease mortality of adults in Burkina Faso, West Africa. *Trop Med Int Health* 2010; 15:1082–1089. DOI: 10.1111/j.1365-3156.2010.02586.x.
- [16] Trégouët DA, König IR, Erdmann J, Munteanu A, Braund PS, Hall AS, Grosshennig A, Linsel-Nitschke P, Perret C, DeSuremain M, Meitinger T, Wright BJ, **Preuss M**, Balmforth AJ, Ball SG, Meisinger C, Germain C, Evans A, Arveiler D, Luc G, Ruidavets JB, Morrison C, van der Harst P, Schreiber S, Neureuther K, Schäfer A, Bugert P, El Mokhtari NE, Schrezenmeir J, Stark K, Rubin D, Wichmann HE, Hengstenberg C, Ouwehand W, Wellcome Trust Case Control Consortium, Cardiogenics Consortium, Ziegler A, Tiret L, Thompson JR, Cambien F, Schunkert H, Samani NJ. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nat Genet* 2009; 41:283–285. DOI: 10.1038/ng.314.
- [17] Myocardial Infarction Genetics Consortium, Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, Anand S, Engert JC, Samani NJ, Schunkert H, Erdmann J, Reilly MP, Rader DJ, Morgan T, Spertus JA, Stoll M, Girelli D, McKeown PP, Patterson CC, Siscovick DS, O'Donnell CJ, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Melander O, Altshuler D, Ardissino D, Merlini PA, Berzuini C, Bernardinelli L, Peyvandi F, Tubaro M, Celli P, Ferrario M, Fetiveau R, Marziliano N, Casari G, Galli M, Ribichini F, Rossi M, Bernardi F, Zonzin P, Piazza A, Mannucci PM, Schwartz SM, Siscovick DS, Yee J, Friedlander Y, Elosua R, Marrugat J, Lucas G, Subirana I, Sala J, Ramos R, Kathiresan S, Meigs JB, Williams G, Nathan DM, MacRae CA, O'Donnell CJ, Salomaa V, Havulinna AS, Peltonen L, Melander O, Berglund G, Voight BF, Kathiresan S, Hirschhorn JN, Asselta R, Duga S, Spreatico M,

Musunuru K, Daly MJ, Purcell S, Voight BF, Purcell S, Nemesh J, Korn JM, McCarroll SA, Schwartz SM, Yee J, Kathiresan S, Lucas G, Subirana I, Elosua R, Surti A, Guiducci C, Gianniny L, Mirel D, Parkin M, Burt N, Gabriel SB, Samani NJ, Thompson JR, Braund PS, Wright BJ, Balmforth AJ, Ball SG, Hall AS, Wellcome Trust Case Control Consortium, Schunkert H, Erdmann J, Linsel-Nitschke P, Lieb W, Ziegler A, König I, Hengstenberg C, Fischer M, Stark K, Grosshennig A, Preuss M, Wichmann HE, Schreiber S, Schunkert H, Samani NJ, Erdmann J, Ouwehand W, Hengstenberg C, Deloukas P, Scholz M, Cambien F, Reilly MP, Li M, Chen Z, Wilensky R, Matthai W, Qasim A, Hakonarson HH, Devaney J, Burnett MS, Pichard AD, Kent KM, Satler L, Lindsay JM, Waksman R, Knouff CW, Waterworth DM, Walker MC, Mooser V, Epstein SE, Rader DJ, Scheffold T, Berger K, Stoll M, Hugel A, Girelli D, Martinelli N, Olivieri O, Corrocher R, Morgan T, Spertus JA, McKeown P, Patterson CC, Schunkert H, Erdmann E, Linsel-Nitschke P, Lieb W, Ziegler A, König IR, Hengstenberg C, Fischer M, Stark K, Grosshennig A, **Preuss M**, Wichmann HE, Schreiber S, Hólm H, Thorleifsson G, Thorsteinsdottir U, Stefansson K, Engert JC, Do R, Xie C, Anand S, Kathiresan S, Ardissino D, Mannucci PM, Siscovick D, O'Donnell CJ, Samani NJ, Melander O, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Altshuler D. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet* 2009; 41:334–341. DOI: 10.1038/ng.327.

## Zeitschriftenartikel als Mitglied eines Konsortiums

- [1] Dastani Z, Johnson T, Kronenberg F, Nelson CP, Assimes TL, März W, **CARDIoGRAM Consortium**, ADIPOGen Consortium, Richards JB. The shared allelic architecture of adiponectin levels and coronary artery disease. *Atherosclerosis* 2013; 229:145–148. DOI: 10.1016/j.atherosclerosis.2013.03.034.
- [2] Reiner AP, Hartiala J, Zeller T, Bis JC, Dupuis J, Fornage M, Baumert J, Kleber ME, Wild PS, Baldus S, Bielinski SJ, Fontes JD, Illig T, Keating BJ, Lange LA, Ojeda F, Müller-Nurasyid M, Munzel TF, Psaty BM, Rice K, Rotter JJ, Schnabel RB, Tang WHW, Thorand B, Erdmann J, **CARDIoGRAM Consortium**, Jacobs J David R, Wilson JG, Koenig W, Tracy RP, Blankenberg S, März W, Gross MD, Benjamin EJ, Hazen SL, Allayee H. Genome-wide and gene-centric analyses of circulating myeloperoxidase levels in the charge and care consortia. *Hum Mol Genet* 2013; 22:3381–3393. DOI: 10.1093/hmg/ddt189.
- [3] Lieb W, Jansen H, Loley C, Pencina MJ, Nelson CP, Newton-Cheh C, Kathiresan S, Reilly MP, Assimes TL, Boerwinkle E, Hall AS, Hengstenberg C, Laaksonen R, McPherson R, Thorsteinsdottir U, Ziegler A, Peters A, Thompson JR, König IR, Erdmann J, Samani NJ, Vasan RS, Schunkert H, **CARDIoGRAM**, Assimes TL, Deloukas P, Erdmann J, Holm H, Kathiresan S, König IR, McPherson R, Reilly MP, Roberts R, Samani NJ, Schunkert H, Stewart AFR. Genetic predisposition to higher blood pressure increases coronary artery disease risk. *Hypertension* 2013; 61:995–1001. DOI: 10.1161/HYPERTENSIONAHA.111.00275.
- [4] Vimalaswaran KS, Berry DJ, Lu C, Tikkanen E, Pilz S, Hiraki LT, Cooper JD, Dastani Z, Li R, Houston DK, Wood AR, Michaëlsson K, Vandenput L, Zgaga L, Yerges-Armstrong LM, McCarthy MI, Dupuis J, Kaakinen M, Kleber ME, Jameson K, Arden N, Raitakari O, Viikari J, Lohman KK, Ferrucci L, Melhus H, Ingelsson E, Byberg L, Lind L, Lorentzon M, Salomaa V, Campbell H, Dunlop M, Mitchell BD, Herzig KH, Pouta A, Hartikainen AL, **Genetic Investigation of Anthropometric Traits-GIANT Consortium**, Streeten EA, Theodoratou E, Jula A, Wareham NJ, Ohlsson C, Frayling TM, Kritchevsky

- SB, Spector TD, Richards JB, Lehtimäki T, Ouwehand WH, Kraft P, Cooper C, März W, Power C, Loos RJF, Wang TJ, Järvelin MR, Whittaker JC, Hingorani AD, Hyppönen E. Causal relationship between obesity and vitamin D status: bi-directional Mendelian randomization analysis of multiple cohorts. *PLoS Med* 2013; 10:e1001383. DOI: 10.1371/journal.pmed.1001383.
- [5] Lu X, Wang L, Chen S, He L, Yang X, Shi Y, Cheng J, Zhang L, Gu CC, Huang J, Wu T, Ma Y, Li J, Cao J, Chen J, Ge D, Fan Z, Li Y, Zhao L, Li H, Zhou X, Chen L, Liu D, Chen J, Duan X, Hao Y, Wang L, Lu F, Liu Z, Yao C, Shen C, Pu X, Yu L, Fang X, Xu L, Mu J, Wu X, Zheng R, Wu N, Zhao Q, Li Y, Liu X, Wang M, Yu D, Hu D, Ji X, Guo D, Sun D, Wang Q, Yang Y, Liu F, Mao Q, Liang X, Ji J, Chen P, Mo X, Li D, Chai G, Tang Y, Li X, Du Z, Liu X, Dou C, Yang Z, Meng Q, Wang D, Wang R, Yang J, Schunkert H, Samani NJ, Kathiresan S, Reilly MP, Erdmann J, Coronary ARtery DIsease Genome-Wide Replication And Meta-Analysis **CARDIoGRAM Consortium**, Peng X, Wu X, Liu D, Yang Y, Chen R, Qiang B, Gu D. Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease. *Nat Genet* 2012; 44:890–894. DOI: 10.1038/ng.2337.
- [6] IL6R Genetics Consortium Emerging Risk Factors Collaboration, Sarwar N, Butterworth AS, Freitag DF, Gregson J, Willeit P, Gorman DN, Gao P, Saleheen D, Rendon A, Nelson CP, Braund PS, Hall AS, Chasman DI, Tybjærg-Hansen A, Chambers JC, Benjamin EJ, Franks PW, Clarke R, Wilde AAM, Trip MD, Steri M, Wittteman JCM, Qi L, van der Schoot CE, de Faire U, Erdmann J, Stringham HM, Koenig W, Rader DJ, Melzer D, Reich D, Psaty BM, Kleber ME, Panagiotakos DB, Willeit J, Wennberg P, Woodward M, Adamovic S, Rimm EB, Meade TW, Gillum RF, Shaffer JA, Hofman A, Onat A, Sundström J, Wassertheil-Smoller S, Mellström D, Gallacher J, Cushman M, Tracy RP, Kauhanen J, Karlsson M, Salonen JT, Wilhelmsen L, Amouyel P, Cantin B, Best LG, Ben-Shlomo Y, Manson JE, Davey-Smith G, de Bakker PIW, O'Donnell CJ, Wilson JF, Wilson AG, Assimes TL, Jansson JO, Ohlsson C, Tivesten A, Ljunggren O, Reilly MP, Hamsten A, Ingelsson E, Cambien F, Hung J, Thomas GN, Boehnke M, Schunkert H, Asselbergs FW, Kastelein JJP, Gudnason V, Salomaa V, Harris TB, Kooner JS, Allin KH, Nordestgaard BG, Hopewell JC, Goodall AH, Ridker PM, Hólm H, Watkins H, Ouwehand

WH, Samani NJ, Kaptoge S, Di Angelantonio E, Harari O, Danesh J. **Collaborators:** Assimes TL, Quertermous T, Go AS, Hlatky MA, Knowles JW, Gudnason V, Smith AV, Panagiotakos DB, Chrysohoou C, Pitsavos C, Stefanadis C, Nelson CP, Braund PS, Samani NJ, Hall AS, Balmforth AJ, Thompson JR, Rendon A, Wilde AA, Trip MD, van der Schoot CE, Kastelein JJ, Ouwehand WH, Sivapalaratnam S, Maiwald S, Basart H, Motazacker M, de Jong JS, Dekker LR, Tanck M, Bezzina CR, Whincup PH, Morris RW, Wannamethee SG, Willeit J, Kiechl S, Gallacher J, Yarnell JW, Lowe G, Rumley A, Samani NJ, Goodall AH, Cambien F, Cushman M, Mukamal KJ, Nordestgaard BG, Tybjaerg-Hansen A, Allin KH, Salomaa V, Havulinna AS, Lokki ML, Nieminen MS, Ripatti S, Sinisalo J, Hung J, McQuillan BM, Beilby JP, Thompson PL, Hólm H, Thorleifsson G, Thorgeirsson G, Thorsteinsdóttir U, Stefansson K, Salomaa V, Jula A, Männistö S, Perola M, Tikkanen E, Asselbergs FW, Boer JM, Onland-Moret NC, van der Schouw YT, Verschuren WM, de Bakker PI, Wennberg P, Jansson JH, Salomaa V, Woodward M, Benjamin EJ, Dupuis J, Fontes JD, Yin X, O'Donnell CJ, Boehnke M, Stringham HM, Tuomilehto J, Schunkert H, Erdmann J, Koenig IR, Nahrstaedt J, Loley C, Stark K, Wilkenborg C, Hengstenberg C, Schreiber S, **Preuss M**, Franks PW, Barroso I, Hallmans G, Shungin D, Thomas GN, Cheng KK, Lam TH, Jiang CC, Reich D, Harris TB, Rimm EB, Pai J, Hopewell JC, Collins R, Parish S, Armitage J, Stringham HM, Jackson A, Hveem K, Psaty BM, Wiggins KL, Heckbert SR, Smith NL, Bis JC, Melzer D, Ferrucci L, Guralnik JM, Bandinelli S, Singleton AB, Kauhanen J, Salonen JT, Tuomainen TP, Kurl S, Meade TW, Chambers JC, Kooner JS, Zhang W, Kooner AS, Das D, Kleber ME, März W, Scharnagl H, Böhm BO, Winkelmann BR, Cushman M, Tracy RP, Folsom AR, Psaty BM, Shea SJ, Stringham HM, Laakso M, Kuusisto J, Koenig W, Baumert J, Thorand B, Illig T, Meisinger C, Wilhelmsen L, Rosengren A, Jansson JO, Adamovic S, Karlsson MK, Ljunggren Ö, Mellström D, Ohlsson C, Tivesten Å, Gillum RF, Rimm EB, Manson JE, Qi L, Hu FB, Hankinson SE, Shaffer JA, Davidson KW, Wilson JF, Fraser R, Wild S, Campbell H, Rader DJ, Reilly MP, Qasim A, Qu L, Li M, Ingelsson E, Lind L, Sundström J, Syvänen AC, Amouyel P, Arveiler D, Clarke R, Watkins H, Farrall M, Hopewell JC, Peden JF, Saleheen D, Deloukas P, Sheikh N, Rasheed A, Danesh J, Cantin B, Dagenais GR, Witteman JC, Hofman A, Dehghan A, van Duijn CM, Uitterlinden AG, Steri M, Abecasis GR, Cucca F, Sanna S, Uda M, Schlessinger D, Ham-



- sten A, Sabater-Lleal M, Silveira A, de Faire U, Gigante B, Best LG, Howard BV, Davey-Smith G, Ben-Shlomo Y, Onat A, Ingelsson E, Sundström J, Lind L, Basu S, Syvänen AC, Ridker PM, Chasman DI, Rose LM, Wassertheil-Smoller S, Ridker PM, Buring J. Interleukin-6 receptor pathways in coronary heart disease: a collaborative meta-analysis of 82 studies. *Lancet* 2012; 379:1205–1213. DOI: 10.1016/S0140-6736(11)61931-4.
- [7] O'Donnell CJ, Kavousi M, Smith AV, Kardia SLR, Feitosa MF, Hwang SJ, Sun YV, Province MA, Aspelund T, Dehghan A, Hoffmann U, Bielak LF, Zhang Q, Eiriksdottir G, van Duijn CM, Fox CS, de Andrade M, Kraja AT, Sigurdsson S, Elias-Smale SE, Murabito JM, Launer LJ, van der Lugt A, Kathiresan S, **CARDIoGRAM Consortium**, Krestin GP, Herrington DM, Howard TD, Liu Y, Post W, Mitchell BD, O'Connell JR, Shen H, Shuldiner AR, Altshuler D, Elosua R, Salomaa V, Schwartz SM, Siscovick DS, Voight BF, Bis JC, Glazer NL, Psaty BM, Boerwinkle E, Heiss G, Blankenberg S, Zeller T, Wild PS, Schnabel RB, Schillert A, Ziegler A, Münzel TF, White CC, Rotter JJ, Nalls M, Oudkerk M, Johnson AD, Newman AB, Uitterlinden AG, Massaro JM, Cunningham J, Harris TB, Hofman A, Peyser PA, Borecki IB, Cupples LA, Gudnason V, Witteman JCM. Genome-wide association study for coronary artery calcification with follow-up in myocardial infarction. *Circulation* 2011; 124:2855–2864. DOI: 10.1161/CIRCULATIONAHA.110.974899.
- [8] Lucas G, Lluís-Ganella C, Subirana I, Sentí M, Willenborg C, Musameh MD, Schwartz SM, O'Donnell CJ, Melander O, Salomaa V, Elosua R, **CARDIoGRAM Consortium**. Post-genomic update on a classical candidate gene for coronary artery disease: ESR1. *Circ Cardiovasc Genet* 2011; 4:647–654. DOI: 10.1161/CIRCGENETICS.111.960583.
- [9] **IBC 50K CAD Consortium**. Large-scale gene-centric analysis identifies novel variants for coronary artery disease. *PLoS Genet* 2011; 7:e1002260. DOI: 10.1371/journal.pgen.1002260.
- [10] International Consortium for Blood Pressure Genome-Wide Association Studies, Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, Smith AV, Tobin MD, Verwoert GC, Hwang SJ, Pihur V, Vollenweider P,

O'Reilly PF, Amin N, Bragg-Gresham JL, Teumer A, Glazer NL, Launer L, Zhao JH, Aulchenko Y, Heath S, Söber S, Parsa A, Luan J, Arora P, Dehghan A, Zhang F, Lucas G, Hicks AA, Jackson AU, Peden JF, Tanaka T, Wild SH, Rudan I, Igl W, Milaneschi Y, Parker AN, Fava C, Chambers JC, Fox ER, Kumari M, Go MJ, van der Harst P, Kao WHL, Sjögren M, Vinay DG, Alexander M, Tabara Y, Shaw-Hawkins S, Whincup PH, Liu Y, Shi G, Kuusisto J, Tayo B, Seielstad M, Sim X, Nguyen KDH, Lehtimäki T, Matullo G, Wu Y, Gaunt TR, Onland-Moret NC, Cooper MN, Platou CGP, Org E, Hardy R, Dahgam S, Palmen J, Vitart V, Braund PS, Kuznetsova T, Uiterwaal CSPM, Adeyemo A, Palmas W, Campbell H, Ludwig B, Tomaszewski M, Tzoulaki I, Palmer ND, **CARDIoGRAM Consortium**, CKDGen Consortium, KidneyGen Consortium, EchoGen Consortium, CHARGE-HF Consortium, Aspelund T, Garcia M, Chang YPC, O'Connell JR, Steinle NI, Grobbee DE, Arking DE, Kardia SL, Morrison AC, Hernandez D, Najjar S, McArdle WL, Hadley D, Brown MJ, Connell JM, Hingorani AD, Day INM, Lawlor DA, Beilby JP, Lawrence RW, Clarke R, Hopewell JC, Ongen H, Dreisbach AW, Li Y, Young JH, Bis JC, Kähönen M, Viikari J, Adair LS, Lee NR, Chen MH, Olden M, Pattaro C, Bolton JAH, Köttgen A, Bergmann S, Mooser V, Chaturvedi N, Frayling TM, Islam M, Jafar TH, Erdmann J, Kulkarni SR, Bornstein SR, Grässler J, Groop L, Voight BF, Kettunen J, Howard P, Taylor A, Guarrera S, Ricceri F, Emilsson V, Plump A, Barroso I, Khaw KT, Weder AB, Hunt SC, Sun YV, Bergman RN, Collins FS, Bonnycastle LL, Scott LJ, Stringham HM, Peltonen L, Perola M, Vartiainen E, Brand SM, Staessen JA, Wang TJ, Burton PR, Soler Artigas M, Dong Y, Snieder H, Wang X, Zhu H, Lohman KK, Rudock ME, Heckbert SR, Smith NL, Wiggins KL, Doumatey A, Shriner D, Veldre G, Viigimaa M, Kinra S, Prabhakaran D, Tripathy V, Langefeld CD, Rosengren A, Thelle DS, Corsi AM, Singleton A, Forrester T, Hilton G, McKenzie CA, Salako T, Iwai N, Kita Y, Ogiwara T, Ohkubo T, Okamura T, Ueshima H, Umemura S, Eyheramendy S, Meitinger T, Wichmann HE, Cho YS, Kim HL, Lee JY, Scott J, Sehmi JS, Zhang W, Hedblad B, Nilsson P, Smith GD, Wong A, Narisu N, Stančáková A, Raffel LJ, Yao J, Kathiresan S, O'Donnell CJ, Schwartz SM, Ikram MA, Longstreth J W T, Mosley TH, Seshadri S, Shrine NRG, Wain LV, Morken MA, Swift AJ, Laitinen J, Prokopenko I, Zitting P, Cooper JA, Humphries SE, Danesh J, Rasheed A, Goel A, Hamsten A, Watkins H, Bakker SJL, van Gilst WH, Janipalli CS, Mani KR, Yajnik CS, Hofman A, Mattace-Raso FUS,

- Oostra BA, Demirkan A, Isaacs A, Rivadeneira F, Lakatta. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* 2011; 478:103–109. DOI: 10.1038/nature10405.
- [11] Suhre K, Shin SY, Petersen AK, Mohny RP, Meredith D, Wägele B, Altmaier E, **CARDIoGRAM**, Deloukas P, Erdmann J, Grundberg E, Hammond CJ, de Angelis MH, Kastenmüller G, Köttgen A, Kronenberg F, Mangino M, Meisinger C, Meitinger T, Mewes HW, Milburn MV, Prehn C, Raffler J, Ried JS, Römisch-Margl W, Samani NJ, Small KS, Wichmann HE, Zhai G, Illig T, Spector TD, Adamski J, Soranzo N, Gieger C. Human metabolic individuality in biomedical and pharmaceutical research. *Nature* 2011; 477:54–60. DOI: 10.1038/nature10354.
- [12] Strawbridge RJ, Dupuis J, Prokopenko I, Barker A, Ahlqvist E, Rybin D, Petrie JR, Travers ME, Bouatia-Naji N, Dimas AS, Nica A, Wheeler E, Chen H, Voight BF, Taneera J, Kanoni S, Peden JF, Turrini F, Gustafsson S, Zabena C, Almgren P, Barker DJP, Barnes D, Dennison EM, Eriksson JG, Eriksson P, Eury E, Folkersen L, Fox CS, Frayling TM, Goel A, Gu HF, Horikoshi M, Iso-maa B, Jackson AU, Jameson KA, Kajantie E, Kerr-Conte J, Kuulasmaa T, Kusisto J, Loos RJF, Luan J, Makrilakis K, Manning AK, Martínez-Larrad MT, Narisu N, Nastase Mannila M, Ohrvik J, Osmond C, Pascoe L, Payne F, Sayer AA, Sennblad B, Silveira A, Stancáková A, Stirrups K, Swift AJ, Syvänen AC, Tuomi T, van 't Hooft FM, Walker M, Weedon MN, Xie W, Zethelius B, **DIA-GRAM Consortium**, **GIANT Consortium**, **MuTHER Consortium**, **CARDIo-GRAM Consortium**, **C4D Consortium**, Ongen H, Mälarstig A, Hopewell JC, Saleheen D, Chambers J, Parish S, Danesh J, Kooner J, Ostenson CG, Lind L, Cooper CC, Serrano-Ríos M, Ferrannini E, Forsen TJ, Clarke R, Franzosi MG, Seedorf U, Watkins H, Froguel P, Johnson P, Deloukas P, Collins FS, Laakso M, Dermitzakis ET, Boehnke M, McCarthy MI, Wareham NJ, Groop L, Pat-tou F, Gloyn AL, Dedoussis GV, Lyssenko V, Meigs JB, Barroso I, Watanabe RM, Ingelsson E, Langenberg C, Hamsten A, Florez JC. Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes. *Diabetes* 2011; 60:2624–2634. DOI: 10.2337/db11-0415.

- [13] Speliotes EK, Yerges-Armstrong LM, Wu J, Hernaez R, Kim LJ, Palmer CD, Gudnason V, Eiriksdottir G, Garcia ME, Launer LJ, Nalls MA, Clark JM, Mitchell BD, Shuldiner AR, Butler JL, Tomas M, Hoffmann U, Hwang SJ, Massaro JM, O'Donnell CJ, Sahani DV, Salomaa V, Schadt EE, Schwartz SM, Siscovick DS, NASH CRN, **GIANT Consortium**, MAGIC Investigators, Voight BF, Carr JJ, Feitosa MF, Harris TB, Fox CS, Smith AV, Kao WHL, Hirschhorn JN, Borecki IB, GOLD Consortium. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet* 2011; 7:e1001324. DOI: 10.1371/journal.pgen.1001324.
- [14] **Coronary Artery Disease Consortium**, Samani NJ, Deloukas P, Erdmann J, Hengstenberg C, Kuulasmaa K, McGinnis R, Schunkert H, Soranzo N, Thompson J, Tiret L, Ziegler A. Large scale association analysis of novel genetic loci for coronary artery disease. *Arterioscler Thromb Vasc Biol* 2009; 29:774–780. DOI: 10.1161/ATVBAHA.108.181388.
- [15] Hu YJ, Berndt S, Gustafsson S, Ganna A, **Genetic Investigation of Anthropometric Traits (GIANT) Consortium**, Hirschhorn J, North KE, Ingelsson E, Lin DY. Meta-analysis of gene-level associations for rare variants based on single-variant statistics. *Am J Hum Genet* 2013; 93:236–248. DOI: 10.1016/j.ajhg.2013.06.011.

## Vorträge

- [1] **Preuß M.** Methodik genomweiter Assoziationsstudien, 2012. Deutsche Gesellschaft für Kinder- und Jugendmedizin / Gesellschaft für Neonatologie und pädiatrische Intensivmedizin (DGKJ/GNPI), Hamburg.
  
- [2] **Preuß M, Ziegler A.** Ein R-Paket zur Berechnung der exakten Verteilung von  $Q$  und  $I^2$  für Meta-Analysen, 2011. Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS), Mainz.
  
- [3] **Preuß M, König IR** on behalf of CARDIoGRAM. Design of the Coronary ARtery DIsease Genome-wide Replication And Meta-analysis (CARDIoGRAM) Study, 2011. Greifswalder Seminarberichte; Biometrische Aspekte der Genomanalyse IV, Greifswald.
  
- [4] **Preuß M, König IR** on behalf of CARDIoGRAM. Design of the Coronary ARtery DIsease Genome-wide Replication And Meta-analysis (CARDIoGRAM) Study, 2010. Deutsche Arbeitsgemeinschaft Statistik (DAGStat), Dortmund.