

Bayesianische Meta-Analyse mit PROC MCMC

Kristina Weber
 Medizinische Hochschule Hannover
 Institut für Biometrie
 Carl-Neuberg-Str. 1
 30625 Hannover
 Weber.Kristina@mh-hannover.de

Martin Jäger
 Medizinische Hochschule Hannover
 Institut für Biometrie
 Carl-Neuberg-Str. 1
 30625 Hannover
 Jaeger.Martin@mh-hannover.de

Zusammenfassung

Meta-Analysen werden heutzutage häufig verwendet, um Informationen verschiedener Studien zu den gleichen Behandlungsgruppen zusammenzufassen und zu interpretieren. Das Konzept der bayesianischen Meta-Analyse sieht die Möglichkeit vor, unterschiedliche a-priori bekannte Informationen zusätzlich zu den Studieninformationen in die Auswertung miteinfließen zu lassen. Außerdem können die unterschiedlichen Studien als Teil eines Random-Effects-Modells, bei dem die einzelnen Studienergebnisse um den tatsächlichen Behandlungseffekt streuen oder eines Fixed-Effects-Modells, das diese Streuung nicht berücksichtigt, gesehen werden. Seit SAS/STAT 9.2 ist es in PROC MCMC möglich bayesianische Meta-Analysen in beiden Modellen mittels Markov-Chain-Monte-Carlo-Algorithmen (MCMC) durchzuführen. Anhand eines Beispiels werden der Programmcode und ein Teil der zugehörigen Outputs erläutert.

Schlüsselwörter: Meta-Analyse, Bayesianische Statistik, PROC MCMC

1 Einleitung

Häufig wird mehr als eine klinische Studie zu einer bestimmten Fragestellung oder auch Behandlung durchgeführt. Dies gilt vor allem für Studien, die die Wirkung einer Medikation oder anderweitigen medizinischen Behandlung untersuchen. Um die Ergebnisse dieser Studien besser zueinander in eine Beziehung zu setzen, kann man diese mittels einer statistischen Meta-Analyse zusammenfassen. Es handelt sich bei einer solchen Meta-Analyse nicht nur um eine narrative Zusammenfassung der Studien und Ergebnisse, sondern vor allem um eine systematische Zusammenfassung. Es soll sichergestellt werden, dass die Einschätzung des Behandlungseffekts nicht nur auf den Ergebnissen einer einzelnen Studie beruht, sondern auf allen vorhandenen Studien. Man erhält im Prinzip über die Meta-Analyse einen Behandlungseffektschätzer, der sich gewichtet aus den Schätzern der einzelnen Studien zusammensetzt.

1.1 Motivierendes Beispiel

Da publizierte Meta-Analysen im Median meist nur drei Studien umfassen [1], bezieht sich das im Folgenden verwendete Beispiel auf eine Meta-Analyse aus zwei Studien. Ein weiterer Grund ist, dass bei seltenen Erkrankungen meist maximal zwei Studien zu

einer Fragestellung vorliegen und somit auch nur eine Meta-Analyse mit zwei Studien durchgeführt werden kann. In den Studien Bond und Opera [2] wurde der Effekt eines Protonenpumpeninhibitors auf funktionelle Dyspepsie untersucht. Patienten mit einer klinisch diagnostizierten funktionellen Dyspepsie wurden eingeschlossen und anschließend zu den Behandlungsgruppen Omeprazol 20mg, 10mg und Placebo randomisiert. Primärer Endpunkt ist in beiden Studien die vollständige Symptomlinderung. Die Studien umfassen jeweils etwa 400 Patienten und wurden doppelblind durchgeführt. Die genauen Stichprobenumfänge und Anteile an Patienten, die in den Vergleichsgruppen Omeprazol 20mg und Placebo den primären Endpunkt erreicht haben, beinhaltet Tabelle 1. Wenn man die beiden Studien voneinander getrennt betrachtet, so zeigt sich bei Bond unter Omeprazol eine Verbesserung der Symptome von 42.5% gegenüber einer Symptomlinderung von 26.0% unter Placebo. In der Opera-Studie ist dieser starke Effekt jedoch nicht mehr deutlich sichtbar. Der Anteil der Patienten, die eine Linderung der Symptome erfahren haben, ist unter Omeprazol nur um 3.5 Prozentpunkte höher als unter Placebo. Der Effekt von Omeprazol auf die Symptomlinderung ist in dieser Studie also nur sehr geringfügig vorhanden.

Tabelle 1: Bond- und Opera-Studie

Studie	Omeprazol 20 mg Linderung/behandelt	Placebo Linderung/behandelt
Bond	93/219 (42.5%)	57/219 (26.0%)
Opera	68/202 (33.7%)	62/203 (30.5%)

Um auf Basis dieser beiden Studien eine Aussage über den Behandlungseffekt treffen zu können, muss eine statistische Meta-Analyse durchgeführt werden, was entweder frequentistisch oder mithilfe bayesianischer Methoden geschehen kann. Die theoretischen Grundlagen und Methoden hierzu werden in Kapitel 2 erläutert. Die Umsetzung in SAS mittels PROC MCMC wird anhand von Code-Beispielen in Kapitel 3 gezeigt. Ergebnisse und Zusammenfassung befinden sich in den Kapiteln 4 und 5.

2 Theoretische Grundlagen und Methoden

2.1 Fixed-Effects- und Random-Effects-Meta-Analyse

In einer Meta-Analyse werden Ergebnisse mehrerer Studien zu einer Fragestellung betrachtet und es wird ein Gesamteffekt, in den die einzelnen Studien einfließen, geschätzt. Zunächst werden daher die einzelnen Studieneffekte, bezogen auf das gleiche Effektmaß, und anschließend der Gesamteffekt basierend auf diesen Studieneffekten geschätzt. Da es sich in Kapitel 1.1 um Studien mit dichotomen Endpunkten handelt, kann pro Studie eine Vierfeldertafel aufgestellt und darauf basierend die Risikodifferenz, das Relative Risiko oder das Odds-Ratio berechnet werden. Hierbei ist zu beachten, dass im Folgenden alle Ausführungen bezogen auf das Odds-Ratio bzw. das Log-

Odds-Ratio erläutert werden. Es ist jedoch relativ einfach den theoretischen Hintergrund und auch die Programme auf die beiden anderen Effektmaße zu erweitern.

Um den Gesamteffekt zu berechnen, werden meist zwei statistische Modelle verwendet: das Fixed-Effects-Modell und das Random-Effects-Modell [3].

Im Fixed-Effects-Modell wird angenommen, dass den beobachteten Studieneffekten ein wahrer Behandlungseffekt zugrunde liegt. Beobachtete Unterschiede der Studieneffekte zum wahren Effekt entstehen also in der gleichen Population und bei der gleichen Fragestellung nur aufgrund des Stichprobenfehlers. Es gibt somit nur eine mögliche Fehlerquelle und wenn man diese ausschalten könnte sollten alle Studieneffekte dem wahren Effekt entsprechen.

Dies ist im Random-Effects-Modell anders. Hier unterliegt jeder Studieneffekt zusätzlich zum Stichprobenfehler einer weiteren Variationsquelle. Es wird also wieder angenommen, dass es nur einen wahren Effekt gibt, da alle Studien die gleiche Fragestellung in der gleichen Population untersuchen, aber zusätzlich ist eine Variation um diesen wahren Effekt möglich. Daher müsste man sowohl den Stichprobenfehler als auch die zusätzliche Heterogenität für alle Studien ausschalten um den wahren Effekt in den einzelnen Studien zu sehen.

Da die Ausschaltung dieser Fehlerquellen nicht möglich ist, kann man die beiden Modelle verwenden, um die Fehler möglichst gut abzubilden. Die Verteilungsannahmen in den beiden Modellen sind im Folgenden dargestellt. Für jede Studie i wird der Studieneffekt δ_i zunächst unter Berücksichtigung der Events r_i , der Eventraten p_i und der Fallzahlen n_i in den jeweiligen Gruppen Kontrolle (C) und Behandlung (T) als Log-Odds-Ratio geschätzt:

$$\begin{aligned} r_i^C &\sim \text{Bin}(p_i^C, n_i^C), \\ r_i^T &\sim \text{Bin}(p_i^T, n_i^T), \\ \text{logit}(p_i^C) &= \mu_i, \\ \text{logit}(p_i^T) &= \mu_i + \delta_i. \end{aligned}$$

Der Behandlungseffekt in der Kontrollgruppe wird dabei mit μ_i bezeichnet. Es wird im Modell davon ausgegangen, dass die Events binomialverteilt sind.

Im Fixed-Effects-Modell nimmt man an, dass die einzelnen Studieneffekte dem Gesamteffekt d entsprechen,

$$\delta_i = d,$$

und der Gesamteffekt um den wahren Behandlungseffekt θ normalverteilt ist mit einer Varianz σ^2 :

$$d \sim N(\theta, \sigma^2).$$

Über diese Normalverteilungsannahme werden die beobachteten Abweichungen der Studieneffekte vom wahren Effekt abgebildet.

Im Random-Effects-Modell wird für die Studieneffekte ebenfalls eine Normalverteilungsannahme angenommen:

$$\delta_i \sim N(d, \tau^2).$$

Die Studieneffekte sind also mit der Varianz τ^2 um den geschätzten Behandlungseffekt d normalverteilt. Dieser ist wiederum wie im Fixed-Effects-Modell mit der Varianz σ^2 um den wahren Behandlungseffekt θ normalverteilt.

Somit werden hier zwei Variationsquellen berücksichtigt. Die zusätzliche Varianz τ^2 um den wahren Effekt wird auch als Heterogenitätsparameter bezeichnet, da dadurch eine Möglichkeit eingeräumt wird heterogene Ergebnisse in den einzelnen Studien abzubilden.

Bayesianische und Frequentistische Statistik

In der frequentistischen Statistik, wie sie klassischerweise benutzt wird, werden die einzelnen Studieneffekte, die Varianzen und der Gesamteffekt für das Fixed- oder das Random-Effects-Modell auf Basis erhobener Daten und ihrer Likelihood geschätzt.

In der bayesianischen Statistik ist es möglich, zusätzlich Informationen außerhalb der Studie mit einfließen zu lassen. Neben den Daten und ihrer Likelihood gehen daher auch so genannte a-priori-Informationen mit in die Entscheidungsfindung ein. Mit diesen beiden Informationsquellen bestimmt man in der bayesianischen Analyse die a-posteriori-Verteilung, aus der Schätzer gebildet und Inferenzen abgeleitet werden können. Inferenzen basieren unter anderem auf dem Credible-Intervall, welches analog zum Konfidenzintervall einen Bereich um den Schätzer angibt. Das Credible-Intervall kann für ein Signifikanzniveau von 5% so interpretiert werden, dass unter Berücksichtigung der aktuellen Daten der wahre Behandlungseffekt mit 95-prozentiger Wahrscheinlichkeit in dem Credible-Bereich liegt. Ein Konfidenzintervall zum gleichen Signifikanzniveau hingegen gibt an, dass mit 95-prozentiger Wahrscheinlichkeit ein Konfidenzintervall für Daten dieser Art berechnet werden kann, das den wahren Behandlungseffekt enthält [4]. Während bei einem Konfidenzintervall also der Parameter als fix und die Konfidenzgrenzen als Zufallsvariablen angesehen werden, werden bei einem Credible-Intervall der Parameter als Zufallsvariable und die Intervallgrenzen als fix interpretiert.

A-priori-Informationen in der bayesianischen Statistik

Zusatzinformationen können aus Expertenwissen und Annahmen aus vorhergehenden oder ähnlichen Studien gewonnen werden. Es ist also möglich, für jede der oben genannten unbekannt Variablen a-priori-Informationen in Form einer Verteilung anzunehmen.

Im Prinzip gibt es zwei Arten der a-priori-Verteilungen. Bei der Verwendung von nicht-informativen a-priori-Verteilungen handelt es sich um eine objektive bayesianische Analyse, da die Ergebnisse nur auf den beobachteten Daten und dem angenommen Modell basieren. Die a-priori-Verteilung geht in die Analyse mit ein, es wird aber angenommen, dass diese keine zusätzlichen Informationen in sich trägt und die Auswertung somit in dieser Hinsicht nicht verzerrt [5]. Im Regelfall handelt es sich bei nicht-informativen a-priori-Verteilungen um vage Annahmen auf einem möglichst weiten Parameterbereich. Aber wie bestimmt man einen nicht-informativen Prior und ist dies überhaupt möglich? Nach Lambert et al. [6] existieren keine tatsächlich nicht-informativen a-priori-Verteilungen, da jede zusätzliche Verteilungsannahme auch in irgendeiner

Form Informationen enthält, wie den Wertebereich, die Form und die Verteilungsfamilie. Daher kann es auch bei nicht-informativen Verteilungen zu unrealistischen Schätzern und demzufolge zu Verzerrungen kommen.

Werden bewusst Informationen außerhalb der Studie als a-priori-Annahmen in der Analyse betrachtet, so handelt es sich um eine subjektive bayesianische Analyse. Man erhält nur dann gleiche Ergebnisse in der Auswertung, wenn sowohl Datensatz, und dadurch die Likelihood, als auch die a-priori-Verteilung in den Auswertungen gleich sind. Wenn sich nur ein Bestandteil ändert, so ergeben sich in jedem Fall unterschiedliche Ergebnisse, die zu unterschiedlichen Inferenzen führen können [7].

Im Regelfall wird eine Studie gestartet, da angenommen wird, dass die neue Therapie einen positiven Effekt gegenüber der Kontrolltherapie oder einem Placebo hat. Hieraus folgt logischerweise, dass in der bayesianischen Statistik häufig positiv gefärbte oder zumindest nicht-informative Zusatzinformationen in die Analyse mit einfließen.

Es stellt sich daher in der bayesianischen Statistik die Frage, inwiefern neu erhobene Daten die getroffenen Annahmen ändern können bzw. welchen Einfluss die a-priori-Informationen auf die Analyse und Inferenz haben. Sutton und Abrams [8] geben an, dass vor allem in kleinen Studien die Schätzung des Heterogenitätsparameters stark von den a-priori-Informationen abhängen und somit auch einen großen Einfluss auf die Analysen haben kann. Lambert et al. [6] verschärfen diese Aussage sogar noch, indem sie angeben, dass in jedem Fall verschiedene a-priori-Annahmen zu verschiedenen Analyseergebnissen führen und eventuell sogar zu unterschiedlichen Inferenzen.

3 Umsetzung in SAS mittels PROC MCMC

Die Umsetzung der bayesianischen Meta-Analyse in SAS ist mittels der Prozedur MCMC möglich. Der Name der Prozedur stammt daher, dass mithilfe von Markov-Chain-Monte-Carlo-Algorithmen (MCMC) über Simulationen und nicht über analytische oder numerische Verfahren Schätzer gebildet und Inferenzen abgeleitet werden können. Dabei wird aus der gemeinsamen a-posteriori-Verteilung aller zu schätzenden Parameter simuliert, indem die vollständigen bedingten Verteilungen der unbekannt Parameter genutzt werden. Die Eigenschaften der Markov-Ketten führen dann zu einer asymptotischen Konvergenz auf die korrekte a-posteriori-Verteilung. Aus den simulierten Werten kann die a-posteriori-Verteilung über Histogramme und empirische Zusammenfassungen charakterisiert werden. Weitere Informationen zu dem methodischen Hintergrund können im SAS/STAT 9.3 User's Guide [9] oder bei Spiegelhalter et al. [3] nachgelesen werden. In den folgenden Unterkapiteln wird SAS-Code und ein Teil des resultierenden Outputs anhand der Beispielstudien in Kapitel 1.1 erläutert. Die hier vorgestellten Programme geben nur einen kurzen Einblick. Eine ausführliche Präsentation der MCMC Prozedur gibt Chen [10]. Die Ergebnisse der Meta-Analysen werden anschließend in Kapitel 4 erläutert.

Als erstes werden in Code 1 die Daten der Beispielstudien über einen Data-Set eingelesen um diese mit der Prozedur MCMC zu verarbeiten. Wichtig ist, dass neben den Stu-

diennamen (`studyid`) die Anzahl an Events (`rt` und `rc`) und die Fallzahlen (`nt` und `nc`) in den einzelnen Studien und Behandlungsgruppen angegeben werden.

```
data bond_opera;
  length studyid $7;
  input studyid $ rt nt rc nc;
  index = _n_;
  datalines;
Bond 93 219 57 219
Opera 68 202 62 203
;
```

Code 1: Einlesen der Daten

3.1 Nicht-informative bayesianische Fixed-Effects-Meta-Analyse

Der zunächst einfachste Fall einer bayesianischen Meta-Analyse ist die Verwendung von nicht-informativen a-priori-Informationen und einem Fixed-Effects-Modell. Code 2 gibt ein Beispiel für eine mögliche Umsetzung in SAS. In der ersten Zeile werden nach dem Prozedur-Aufruf der Datensatz (`data=`), die Anzahl an Simulationen (`nmc=`), der Thinning-Faktor (`thin=`), der Seed (`seed=`) und die zu schätzenden Variablen (`monitor=`) übergeben. Dabei gibt der Thinning-Faktor an, dass in diesem Fall nur jede fünfte Simulation für die Charakterisierung der a-posteriori-Verteilung und somit für die Schätzungen verwendet werden soll. Diese Angabe ist optional und reduziert die Stärke der Autokorrelationen zwischen den einzelnen Simulationen. Da es sich um eine Meta-Analyse mit zwei Studien handelt, wird für die Basiseffekte in den Studien ein zwei-stufiges Array (`array mu[2]`) angegeben. Bei größeren Meta-Analysen muss diese Zahl angepasst werden.

```
proc mcmc data=bond_opera nmc=50000 thin=5 seed=5677 monitor=(mu d);
array mu[2];
/*Initialwerte*/
parm mu: 0;
parm d 0;
/*a-priori-Verteilungen*/
prior d ~ n(mean=0, var=10); /*Behandlungseffekt*/
prior mu: ~ n(mean=0, var=4); /*Basiseffekt in Studie i*/
/*Modelldefinition*/
p = logistic(d + mu[index]); /*Rate der Behandlungsgruppe*/
q = logistic(mu[index]); /*Rate der Kontrollgruppe*/
model rt ~ binomial(nt, p); /*Modellierung der Behandlungsevents*/
model rc ~ binomial(nc, q); /*Modellierung der Kontrollevents*/
run;
```

Code 2: Nichtinformative Fixed-Effects-Meta-Analyse

Anschließend werden für alle unbekanntenen Variablen, die nicht direkt aus den Daten berechnet werden können, Startwerte für die Durchführung der Simulation festgelegt. In diesem Fall sind das die Basiseffekte in den Studien und der Gesamtschätzer. Die a-

priori-Verteilungen werden mit dem Wort `prior` definiert. Da zunächst ein Fixed-Effects-Modell betrachtet wird, sind hier ebenfalls nur für die Basiseffekte und den Gesamteffekt nicht-informative a-priori-Verteilungen gegeben, welche wie folgt aussehen:

$$d \sim N(0,10),$$

$$\mu_i \sim N(0,4).$$

Die Wahl der a-priori-Verteilungen basiert auf Smith et al. [11], Studieneffekte werden nicht angegeben und sind nicht direkt von Interesse, da bei einem Fixed-Effects-Modell angenommen wird, dass die einzelnen Studieneffekte dem Gesamteffekt bis auf einen Stichprobenfehler entsprechen. Nachdem Funktionen für die Berechnung der Raten in den einzelnen Gruppen der Studien hinterlegt wurden, kann über die Signalworte `model` das den Studien zugrunde liegende Binomialmodell für die Events angegeben werden. Die Prozedur liefert, angewandt auf das einführende Beispiel in 1.1, im Output automatisch mehrere Grafiken und Tabellen. Die Schätzer und 95%-Credible-Intervalle der zu monitorierenden Variablen werden in den Tabellen „Posterior Summaries“ und „Posterior Intervals“ ausgegeben. Diese sind in Abbildung 1 aufgeführt. Da es sich um ein Fixed-Effects-Modell handelt, werden nur Schätzer für die Basiseffekte in den einzelnen Studien und den Behandlungseffekt berechnet. Es wird angegeben, dass die Schätzung von Mittelwert und Standardabweichung der jeweiligen Variablen auf 10000 Simulationen beruhen. Das zweiseitige 95%-Credible-Intervall, basierend auf den 2.5%- und 97.5%-Quantilen, wird unter der Bezeichnung „Equal-Tail Interval“ angegeben. Das „HPD Interval“ steht für High Posterior Density und gibt ein 95%-Intervall an, das nicht auf Basis der Quantile berechnet wird und somit bei schiefen Verteilungen geeigneter ist.

Posterior Summaries						
Parameter	N	Mean	Standard Deviation	Percentiles		
				25%	50%	75%
mu1	10000	-0.8864	0.1263	-0.9702	-0.8857	-0.7989
mu2	10000	-0.9862	0.1338	-1.0734	-0.9838	-0.8967
d	10000	0.4532	0.1461	0.3529	0.4522	0.5506

Posterior Intervals					
Parameter	Alpha	Equal-Tail Interval		HPD Interval	
mu1	0.050	-1.1368	-0.6480	-1.1354	-0.6476
mu2	0.050	-1.2565	-0.7289	-1.2518	-0.7263
d	0.050	0.1712	0.7441	0.1816	0.7513

Abbildung 1: Schätzer und Credible-Intervalle der nicht-informativen Fixed-Effects-Meta-Analyse

Zusätzlich zu diesen Maßzahlen werden Plots zur Konvergenzdiagnostik für jeden Parameter von Interesse ausgegeben. Für den Behandlungseffekt d sind diese in **Abbildung 2** gezeigt. Anhand dieser Grafiken kann überprüft werden, ob die Markov-Kette konvergiert und somit die richtige a-posteriori-Verteilung simuliert wird. Die obere Grafik ist ein Trace-Plot der einzelnen simulierten Werte über alle Ziehungen. Es ist kein Trend erkennbar und es scheint, als ob die Markov-Kette auf eine stationäre a-posteriori-Verteilung konvergiert, da Mittelwert und Varianz über alle Simulationen hinweg relativ konstant sind. Der Autokorrelationsplot links unten zeigt eine niedrige Autokorrelation an, welche für eine gute Durchmischung der simulierten Werte notwendig ist. Die a-posteriori-Dichte rechts unten in der Abbildung ist unimodal und normalverteilt mit einem Maximum um den geschätzten Wert von 0.4532 (s. Abbildung 1, „Posterior Summaries“). Die Plots zur Konvergenzdiagnostik und a-posteriori-Dichte sehen für die Basiseffekte ähnlich aus und werden deshalb nicht berichtet. Der angegebene Code produziert diese jedoch automatisch.

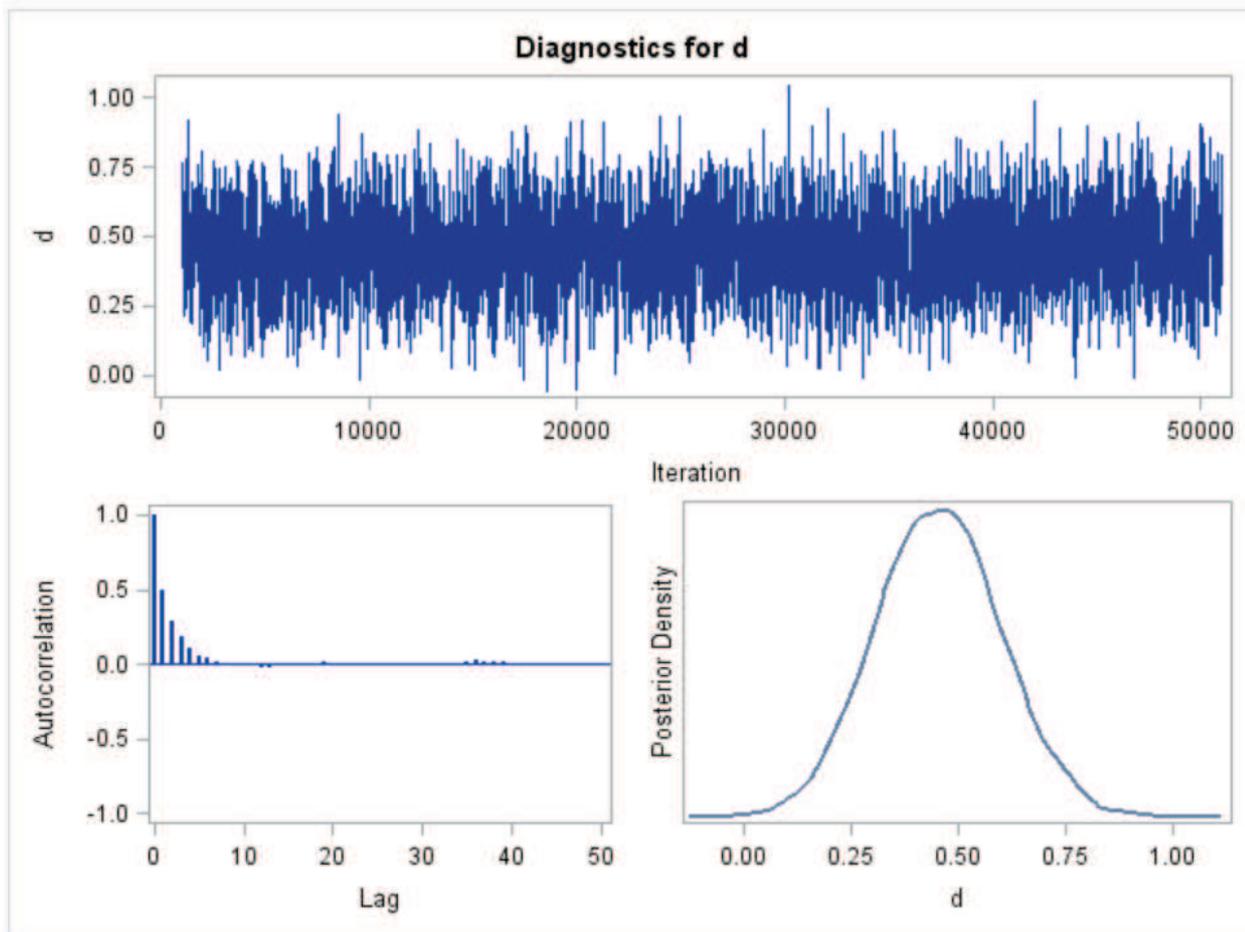


Abbildung 2: Konvergenzdiagnostik und a-posteriori-Dichte für den Behandlungseffekt d der nicht-informativen Fixed-Effects-Meta-Analyse

3.2 Nicht-informative bayesianische Random-Effects-Meta-Analyse

Um nun nicht nur einen gemeinsamen Behandlungseffekt für alle Studien zu schätzen, sondern den Studien eine Heterogenität zu ermöglichen, benutzt man das Random-Effects-Modell. Die Umsetzung mittels PROC MCMC für nicht-informative a-priori-

Verteilungen ist in Code 3 dargestellt. Im Prinzip ist dieser baugleich zu Code 2. Da jedoch im Random-Effects-Modell zusätzliche Parameter für die Studieneffekte und die Heterogenitätsvarianz enthalten sind, müssen für diese Initialwerte und auch a-priori-Verteilungen festgelegt werden. In der Modelldefinition wird der Studieneffekt δ_i statt dem Gesamt-Behandlungseffekt d in der Berechnung der Rate genutzt. Das Ziel ist es jedoch, wieder den Behandlungseffekt basierend auf allen in die Meta-Analyse eingehenden Studien zu schätzen. Die a-priori-Verteilungen werden wie folgt festgelegt:

$$\begin{aligned}d &\sim N(0, 10), \\ \mu_i &\sim N(0, 4), \\ \theta_i &\sim N\left(d, \frac{1}{\tau^2}\right), \\ \tau^2 &\sim \text{Gamma}(3, 1).\end{aligned}$$

Auf den Behandlungseffekt d und den Basiseffekt in den einzelnen Studien werden allgemeine nicht-informative a-priori-Verteilungen gelegt, die in diesem Fall denen im Fixed-Effects-Modell entsprechen. Die nicht-informative a-priori-Verteilung der Studieneffekte θ_i ist eine Normalverteilung mit dem Behandlungseffekt d als Erwartungswert und der Inversen des Heterogenitätsparameters τ^2 als Varianz. Für diesen Parameter ist die nicht-informative a-priori-Verteilung eine Gammaverteilung mit Erwartungswert 3 und Varianz 3. Die Wahl dieser a-priori-Verteilungen orientiert sich ebenfalls an Smith et al. [11].

```
proc mcmc data=bond_opera nmc=50000 thin=5 seed=62511
monitor=(mu theta d tau2);

array mu[2];
array theta[2];

/*Initialwerte*/
parm theta: 0;
parm mu: 0;
parm tau2 1;
parm d 0;

/*a-priori-Verteilungen*/
prior d ~ n(mean=0, var=10); /*Behandlungseffekt*/
prior mu: ~ n(mean=0, var=4); /*Basiseffekt in Studie i*/
prior theta: ~ n(mean=d, var=1/tau2); /*Behandlungseffekt Studie i*/
prior tau2 ~ gamma(shape=3, iscale=1); /*Heterogenitätsparameter*/

/*Modelldefinition*/
p = logistic(theta[index]+mu[index]); /*Rate der Behandlungsgruppe*/
q = logistic(mu[index]); /*Rate der Kontrollgruppe*/
model rt ~ binomial(nt, p); /*Modellierung der Behandlungsevents*/
model rc ~ binomial(nc, q); /*Modellierung der Kontrollevents*/

run;
```

Code 3: Nicht-informative Random-Effects-Meta-Analyse

Wie auch für die vorhergehende Analyse stellt SAS die Ergebnisse automatisch in Form von mehreren Tabellen und Abbildungen dar. Abbildung 4 zeigt die Schätzer und 95%-

Credible-Intervalle der einzelnen Parameter von Interesse. Der Aufbau entspricht weitestgehend der Abbildung 1, da jedoch im Random-Effects-Modell zusätzliche Parameter untersucht werden, sind diese mit aufgeführt. Für jeden dieser Parameter werden ebenfalls Plots zur Konvergenzdiagnostik ausgegeben. Da diese sich nicht nennenswert von Abbildung 2 unterscheiden, wird auf eine Darstellung verzichtet.

Posterior Summaries						
Parameter	N	Mean	Standard Deviation	Percentiles		
				25%	50%	75%
mu1	10000	-1.0229	0.1495	-1.1254	-1.0215	-0.9204
mu2	10000	-0.8381	0.1454	-0.9336	-0.8357	-0.7386
theta1	10000	0.7045	0.1969	0.5704	0.7101	0.8381
theta2	10000	0.1738	0.1994	0.0393	0.1729	0.3090
d	10000	0.4314	0.4765	0.1431	0.4343	0.7257
tau2	10000	3.2504	1.7770	1.9695	2.9196	4.2279

Posterior Intervals					
Parameter	Alpha	Equal-Tail Interval		HPD Interval	
mu1	0.050	-1.3150	-0.7283	-1.3214	-0.7400
mu2	0.050	-1.1347	-0.5537	-1.1407	-0.5627
theta1	0.050	0.3162	1.0885	0.3229	1.0931
theta2	0.050	-0.2247	0.5645	-0.2154	0.5682
d	0.050	-0.5276	1.3689	-0.5240	1.3711
tau2	0.050	0.7693	7.5687	0.5042	6.7178

Abbildung 4: Schätzer und Credible-Intervalle der nicht-informative Random-Effects-Meta-Analyse

3.3 Informative bayesianische Fixed- und Random-Effects-Meta-Analyse

Um eine informative Meta-Analyse für beide Modelle durchzuführen, müssen in Code 2 und Code 3 nur die a-priori-Verteilungen anders definiert werden. Der restliche Code bleibt unverändert und auch die Outputs unterscheiden sich bezüglich Aufbau und Interpretation nicht von den bereits vorgestellten.

Eine Möglichkeit informativer a-priori-Verteilungen ist die Annahme eines positiven Effekts für den Behandlungseffekt mit verhältnismäßig kleiner Varianz. Es ist natürlich auch möglich, größere oder auch kleinere und sogar negative Effekte als a-priori-Annahme zu wählen. In Kapitel 4 werden neben den Ergebnissen für die angegeben nicht-

informativen a-priori-Verteilungen auch die Ergebnisse für die folgenden informativen a-priori-Annahmen auf den Behandlungseffekt dargestellt:

$$d \sim N(2, 1)$$

und

$$d \sim N(4, 1).$$

Auf die Basiseffekte der einzelnen Studien, die Studienbehandlungseffekte und den Heterogenitätsparameter wird keine zusätzliche Vorinformation gelegt, so dass es sich immer noch um die gleichen nicht-informativen a-priori-Verteilungen handelt wie in den Kapiteln 3.1 und 3.2. Die Umsetzung dieser Annahmen für das Random-Effects-Modell in PROC MCMC ist in Code 4 dargestellt.

```
prior d ~ n(mean=2, var=1); *oder prior d ~ n(mean=4, var=1);
prior mu: ~ n(mean=0, var=4);
prior theta: ~ n(mean=d, var=1/tau2);
prior tau2 ~ gamma(shape=3, iscale=1);
```

Code 4: Informative a-priori-Verteilungen für das Random-Effects-Modell

4 Ergebnisse

In diesem Kapitel werden die Behandlungseffektschätzer \hat{d} der vorgestellten bayesianischen Meta-Analysen basierend auf den beiden Beispielstudien Opera und Bond erläutert. Zusätzlich wird der Schätzer aus einer frequentistischen Meta-Analyse angegeben. Zu beachten ist, dass alle Angaben auf der Log-Odds-Ratio-Skala basieren.

In Tabelle 2 werden zunächst nur Schätzer und zugehörige 95%-Konfidenz- bzw. Credible-Intervalle für Fixed-Effects-Meta-Analysen angegeben. In der frequentistischen Meta-Analyse werden keine Vorinformationen angenommen und als Schätzer für den Behandlungseffekt ergibt sich ein Log-Odds-Ratio von 0.455. Anhand des Konfidenzintervalls kann davon ausgegangen werden, dass dieser positive Effekt für Omeprazol signifikant zu einem Signifikanzniveau von 5% ist. Werden in der bayesianischen Fixed-Effects-Meta-Analyse ebenfalls keine Vorinformationen angenommen, ändern sich Schätzer und Intervall nur leicht und auch hier würde man auf einen vorhandenen Effekt von Omeprazol gegenüber Placebo schließen. Die Hinzunahme der a-priori-Annahme eines positiven Behandlungseffekts hat im Fixed-Effects-Modell bei gegebener Datenlage nur einen leichten Einfluss auf den Schätzer und das Credible-Intervall. Bei beiden a-priori-Verteilungen ändern sich die Schätzer nur leicht und man kann immer noch von einem tatsächlich vorhandenen Effekt ausgehen.

Tabelle 2: Ergebnisse der Fixed-Effects-Meta-Analysen

Ansatz	a-priori-Verteilung	Behandlungs- effekt \hat{d}	95%-Konfidenz-/ Credible-Intervall
Frequentistisch	-	0.455	[0.167, 0.745]
Bayesianisch	nicht-informativ	0.456	[0.176, 0.743]
	informativ, $d \sim N(2,1)$	0.489	[0.388, 0.589]
	informativ, $d \sim N(4,1)$	0.529	[0.245, 0.816]

Die Ergebnisse bezüglich des Behandlungseffektschätzers \hat{d} und das zugehörige Intervall unter Annahme eines Random-Effects-Modells sind in Tabelle 3 dargestellt. In der frequentistischen Random-Effects-Meta-Analyse zeigt sich ein ähnlicher Effekt wie in der Fixed-Effects-Meta-Analyse. Betrachtet man jedoch das 95%-Konfidenzintervall, so fällt auf, dass es sich unter diesem Modell nicht um einen signifikanten Effekt zwischen Omeprazol und Placebo handelt. Auch hier sind Schätzer und 95%-Credible-Intervall in der nicht-informativen bayesianischen Meta-Analyse den frequentistischen Ergebnissen sehr ähnlich. Je mehr von den Studien unabhängige Informationen jedoch über die a-priori-Annahmen einfließen, desto stärker verschiebt sich das Credible-Intervall. Der Schätzer an sich ändert sich nur leicht. Bei einer zusätzlichen Annahme eines relativ großen Effekts von 4 auf der Log-Odds-Ratio-Skala würde man auf Basis der bayesianischen informativen Random-Effects-Meta-Analyse von einem tatsächlich vorhandenen Behandlungseffekt von Omeprazol gegenüber Placebo sprechen. In diesem Fall schließt das 95%-Credible-Intervall die Null nicht mehr mit ein.

Tabelle 3: Ergebnisse der Random-Effects-Meta-Analysen

Ansatz	a-priori-Verteilung	Behandlungs- effekt \hat{d}	95%-Konfidenz-/ Credible-Intervall
Frequentistisch	-	0.434	[-0.515, 1.387]
Bayesianisch	nicht-informativ	0.456	[-0.537, 1.394]
	informativ, $d \sim N(2,1)$	0.489	[-0.101, 1.696]
	informativ, $d \sim N(4,1)$	0.529	[0.280, 2.994]

5 Zusammenfassung

Bayesianische Analysen spielen neben klassischen frequentistischen Analysen eine immer größer werdende Rolle in statistischen Auswertungen. Das Konzept der bayesianischen Analyse lässt sich auch auf Meta-Analysen übertragen. Die Umsetzung ist seit SAS 9.2 mit PROC MCMC möglich. Es gibt im Grunde zwei Modellannahmen, die für die Modellierung der Daten und Effekte in einer Meta-Analyse genutzt werden. Im Fixed-Effects-Modell wird angenommen, dass die einzelnen Studieneffekte dem Behandlungseffekt bis auf einen Stichprobenfehler entsprechen. Die Annahmen im Random-Effects-Modell beruhen darauf, dass die Studieneffekte eine zusätzliche Heterogenität aufweisen und von dem Behandlungseffekt abweichen können.

Zusätzlich zu diesen Modellannahmen können in der bayesianischen Statistik von den zu untersuchenden Studien unabhängige externe Informationen und Annahmen auf un-

bekannte Parameter über a-priori-Verteilungen einfließen. Dabei ist es prinzipiell möglich, diese zusätzlichen Annahmen nicht-informativ oder informativ zu wählen und damit die bayesianische Analyse objektiv oder subjektiv durchzuführen.

Auf den ersten Blick sind keine Vorteile gegenüber einer frequentistischen Meta-Analyse basierend auf dem gleichen Modell sichtbar. Es stellt sich daher die Frage, ob es nicht grundsätzlich einfacher ist, eine frequentistische Meta-Analyse durchzuführen, wenn zusätzliche Annahmen nicht-informativ sind. Dadurch vermeidet man eine mögliche Verzerrung durch die a-priori-Verteilungen, da, wie Lambert et al. [6] bereits gezeigt hat, sogar vermeintlich nicht-informative Verteilungen und Annahmen Informationen tragen können. Es ist bei der objektiven bayesianischen Meta-Analyse unabhängig vom zugrunde liegenden Modell schwierig bzw. nahezu unmöglich, tatsächlich nicht-informative a-priori-Verteilungen zu bestimmen.

Sollen in die Analysen zusätzliche Informationen einfließen, so werden informative a-priori-Verteilungen definiert, die zu einer subjektiven Meta-Analyse führen. Auch hier ist es schwierig, geeignete a-priori-Verteilungen zu bestimmen. Bisher existieren keine Guidelines und nur wenige Publikationen, an denen man sich bei der Definition der Verteilungen orientieren kann. In den Ergebnissen zeigte sich, dass bei dem vorliegenden Beispiel vor allem die Credible-Intervalle des Behandlungseffektschätzers im Random-Effects-Modell und somit die Inferenzen von den informativen a-priori-Verteilungen beeinflusst werden.

Allgemein und insbesondere bei subjektiven bayesianischen Meta-Analysen sollten a-priori-Verteilungen begründet und zusammen mit den Ergebnissen berichtet werden, um eine eventuelle Abhängigkeit zwischen zusätzlichen Annahmen und den geschätzten Parametern zu erkennen. Außerdem ist zu bedenken, dass vor allem in kleinen Meta-Analysen wie im vorliegenden Beispiel eine definite bayesianische Auswertung nicht möglich ist. Je mehr Daten erhoben werden und somit Studien in die Auswertung einfließen, desto genauer und eindeutiger werden jedoch die Ergebnisse und auch die Inferenzen. Bei einer ausreichend großen Anzahl an Studien sollte dies sogar nahezu unabhängig von den a-priori-Annahmen sein. Dies ist darin begründet, dass erhobene Daten und deren Likelihood ein größeres Gewicht im Vergleich zu den a-priori-Verteilungen erhalten.

Die bayesianische Meta-Analyse bietet durch die Möglichkeit, informative a-priori-Verteilungen zu definieren, Vorteile gegenüber der frequentistischen Auswertung, die jedoch mit zunehmendem Umfang der Meta-Analyse nach und nach vernachlässigbar werden. Bei kleinen Meta-Analysen kommt es durch die zusätzlichen Informationen zu einer Verzerrung in Richtung dieser Annahmen. Im Falle einer nicht-informativen bayesianischen Meta-Analyse ist jedoch kein direkter Vorteil gegenüber einer frequentistischen Analyse ersichtlich.

Literatur

- [1] J. Davey, R. M. Turner, M. J. Clarke, und J. P. Higgins, „Characteristics of meta-analyses and their component studies in the Cochrane Database of Systematic Reviews: a cross-sectional, descriptive analysis“, *BMC Med. Res. Methodol.*, Bd. 11, Nr. 1, S. 160, 2011.
- [2] N. J. Talley, V. Meineche-Schmidt, P. Paré, M. Duckworth, P. Räisänen, A. Pap, H. Kordecki, und V. Schmid, „Efficacy of omeprazole in functional dyspepsia: Double-blind, randomized, placebo-controlled trials (the Bond and Opera studies)“, *Aliment. Pharmacol. Ther.*, Bd. 12, Nr. 11, S. 1055–1065, 1998.
- [3] D. J. Spiegelhalter, J. P. Myles, D. R. Jones, und K. R. Abrams, „Bayesian methods in health technology assessment: a review.“, *Health Technol. Assess.*, Bd. 4, Nr. 38, S. 1–130, 2000.
- [4] J. Vanderplas, „Frequentism and Bayesianism: A Python-driven Primer“, *Arxiv*, Nr. Scipy, S. 1–9, 2014.
- [5] J.M. Bernardo, „Comment on article by Gelman“, *Bayesian Anal.*, Bd. 3, Nr. 3, S. 451–454, 2008.
- [6] P. C. Lambert, A. J. Sutton, P. R. Burton, K. R. Abrams, und D. R. Jones, „How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS“, *Stat. Med.*, Bd. 24, Nr. 15, S. 2401–2428, 2005.
- [7] J. B. Kadane, „Comment on Article by Gelman“, *Bayesian Anal.*, Bd. 3, Nr. 3, S. 455–458, 2008.
- [8] A. J. Sutton und K. R. Abrams, „Statistical Methods in Medical Research“, *Stat. Methods Med. Researc*, Bd. 10, S. 277–303, 2001.
- [9] SAS Institute Inc., „Sas/Stat 9.3 User’s Guide“, Cary, NC, 2011.
- [10] F. Chen, „Bayesian modeling using the MCMC procedure“, *Proc. SAS Glob. Forum 2008*, S. 1–22, 2009.
- [11] T. C. Smith, D. J. Spiegelhalter, und A Thomas, „Bayesian approaches to random-effects meta-analysis: a comparative study.“, *Stat. Med.*, Bd. 14, Nr. 24, S. 2685–2699, 1995.