

**Aus der**  
**Klinik und Poliklinik für Psychiatrie und Psychotherapie**  
**der Ludwig- Maximilians- Universität München**

**Vorstand: Herr Prof. Dr. med. H.- J. Möller**

**Zur Prädiktion des Erfolgs einer Lithiumprophylaxe bei bipolar affektiven  
Störungen. Der Einfluss klinischer, demographischer und psychosozialer  
Faktoren auf das Risiko eines stationären Rezidivs**

Dissertation  
zum Erwerb des Doktorgrades der Humanbiologie  
an der Medizinischen Fakultät der  
Ludwig- Maximilians- Universität zu München

vorgelegt von  
Nikolaus Kleindienst

aus  
Lindenberg

2004

Mit Genehmigung der Medizinischen Fakultät  
der Universität München

Berichterstatter: Prof. Dr. rer. nat. R. R. Engel

Mitberichterstatter: Prof. Dr. Dr. M. Albus  
Prof. Dr. D. Hölzel

Mitbetreuung durch den  
promovierten Mitarbeiter: Prof. Dr. med. W. Greil

Dekan: Prof. Dr. med. Dr. h. c. K. Peter

Tag der mündlichen Prüfung: 29. 6. 2004

## Inhaltsverzeichnis

<b>Inhaltsverzeichnis</b> .....	<b>3</b>
<b>0 ZUSAMMENFASSUNG</b> .....	<b>7</b>
<b>1 EINLEITUNG</b> .....	<b>9</b>
<b>2 THEORETISCHER HINTERGRUND</b> .....	<b>11</b>
<b>2.1 Diagnostik und Verlauf bipolar affektiver Störungen</b> .....	<b>11</b>
<b>2.2 Allgemeines zur medikamentösen Therapie bipolar affektiver Störungen</b> .....	<b>13</b>
2.2.1 Akuttherapie, Erhaltungstherapie und Rezidivprophylaxe .....	13
2.2.2 Medikamentöse Rezidivprophylaxe bei bipolaren Störungen .....	14
<b>2.3 Carbamazepin</b> .....	<b>16</b>
<b>2.4 Lithium</b> .....	<b>17</b>
2.4.1 Allgemeines zu Lithium .....	17
2.4.2 Lithium in der Behandlung bipolarer Störungen .....	17
2.4.2.1 Langzeitbehandlung bipolarer Störungen mit Lithium .....	19
2.4.2.2 Zur Kontroverse um die prophylaktische Wirksamkeit von Lithium .....	19
2.4.2.3 Primärarbeiten zur Prädiktion der prophylaktischen Wirksamkeit von Lithium .....	21
2.4.2.4 Integrative Arbeiten zur Prädiktion der prophylaktischen Wirkung von Lithium .....	23
<b>3 FRAGESTELLUNG</b> .....	<b>27</b>
<b>4 METAANALYSEN ZUR PRÄDIKTION DES ERFOLGS EINER LITHIUMPROPHYLAXE</b> .....	<b>29</b>
<b>4.1 Methodik</b> .....	<b>29</b>
4.1.1 Lokalisation der Studien.....	29
4.1.2 Charakterisierung und Auswertung der Studien .....	31
4.1.2.1 Charakterisierung und Rating der Studien.....	31
4.1.2.2 Zusammenfassung der Studien, Effektstärkemaße .....	33
4.1.3 Berechnung des Effektstärkemaßes $r$ aus den Primärstudien .....	34
4.1.3.1 Korrelationskoeffizienten in den Primärstudien .....	34
4.1.3.2 4-Felder-Tafeln in den Primärstudien.....	34
4.1.3.3 Kontingenztafeln in den Primärstudien .....	34
4.1.3.4 t-Tests, F-Tests, Mittelwerte und Standardabweichungen in den Primärstudien .....	35
4.1.3.5 Regressionsanalysen in den Primärstudien.....	35
4.1.3.6 Sonstige Tests oder unzureichende Angaben zur Statistik in den Primärstudien .....	36
4.1.4 Aggregation der Effektstärken.....	36
4.1.5 Darstellung der Studienergebnisse – Forestplot und Treepplot.....	39
<b>4.2 Eingeschlossene Studien</b> .....	<b>41</b>
<b>4.3 Ergebnisse der Metaanalyse zu demographischen Variablen</b> .....	<b>51</b>
4.3.1 Geschlecht .....	51
4.3.2 Alter .....	53
4.3.3 Familienstand .....	55
<b>4.4 Ergebnisse der Metaanalyse zu klinischen Variablen</b> .....	<b>58</b>
4.4.1 Diagnostik und Psychopathologie .....	58
4.4.1.1 Bipolar-II-Störung.....	58

4.4.1.2 Manische Indexepisode .....	60
4.4.1.3 Depressive Indexepisode.....	61
4.4.1.4 Gemischte Indexepisode.....	61
4.4.1.5 Zyklische Indexphase .....	62
4.4.1.6 Psychotische Symptome .....	63
4.4.1.7 Stimmungskongruente psychotische Symptome.....	64
4.4.1.8 Stimmungsinkongruente psychotische Symptome während der Indexepisode.....	64
4.4.1.9 Klassische euphorische Manie .....	66
4.4.1.10 Reizbarkeit, Agitiertheit.....	67
4.4.1.11 Aggressivität, Feindseligkeit .....	67
4.4.1.12 Akustische Halluzinationen.....	67
4.4.1.13 Denkstörungen .....	67
4.4.1.14 Ängstlichkeit .....	68
4.4.1.15 Depressivität, Verlangsamung.....	68
4.4.1.16 Schwere der Indexepisode.....	69
4.4.1.17 Länge der Indexepisode .....	70
4.4.1.18 Vorliegen einer sekundären affektiven Erkrankung .....	71
4.4.2 Krankheitsanamnese.....	72
4.4.2.1 Anzahl bisheriger Krankheitsphasen .....	72
4.4.2.2 Bisherige Anzahl stationärer Krankheitsphasen.....	74
4.4.2.3 Anzahl depressiver Krankheitsphasen.....	75
4.4.2.4 Anzahl manischer Krankheitsphasen.....	76
4.4.2.5 Hohe Phasenfrequenz und rapid cycling .....	77
4.4.2.6 Krankheitsdauer .....	81
4.4.2.7 Alter bei Ersterkrankung .....	83
4.4.2.8 Alter bei Erstbehandlung.....	85
4.4.2.9 Verlaufsmuster MDI.....	87
4.4.2.10 Verlaufsmuster DMI.....	88
4.4.2.11 Verlaufsmuster CC.....	90
4.4.2.12 Isolierte Krankheitsphasen .....	91
4.4.2.13 Irreguläres Verlaufsmuster .....	92
4.4.2.14 Freies Intervall .....	93
4.4.2.15 Verlaufscharakteristika nach J. Angst .....	93
4.4.2.16 Abruptes vs. schleichendes Auftreten der Manie .....	94
4.4.2.17 Anteil manischer Episoden.....	94
4.4.2.18 Hoher Anteil gemischter Episoden.....	95
4.4.2.19 Typ der ersten Episode.....	96
4.4.2.20 Länge des ersten freien Intervalls.....	97
4.4.2.21 Psychotische Merkmale während mindestens einer Episode.....	98
4.4.2.22 Stimmungsinkongruente psychotische Merkmale während mindestens einer Episode.....	99
4.4.2.23 Saisonales Muster.....	99
4.4.2.24 Suizidversuch .....	99
4.4.2.25 Durchschnittliche Morbidität in den Jahren vor der Prophylaxe .....	100
4.4.3 Psychiatrische Komorbidität .....	101
4.4.3.1 Substanzmissbrauch .....	101
4.4.3.2 Alkoholmissbrauch.....	102
4.4.3.3 Angst- oder Panikstörung .....	102
4.4.3.4 Persönlichkeitsstörung.....	103
<b>4.5 Ergebnisse der Metaanalyse zu psychosozialen Variablen.....</b>	<b>104</b>
4.5.1 Soziale Schicht .....	104
4.5.2 Arbeitslosigkeit .....	105
4.5.3 Soziale Unterstützung (social support).....	106
4.5.4 Familienatmosphäre (expressed emotions und affective style).....	107
4.5.5 Adäquate Freizeitaktivitäten.....	107
4.5.6 Stress .....	108
4.5.7 Kritische Lebensereignisse (life events).....	109
4.5.8 Ängstlichkeit .....	110
4.5.9 Neurotizismus.....	111
4.5.10 Extraversion .....	112

4.5.11 Psychotizismus .....	113
4.5.12 Zwanghaftigkeit.....	114
4.5.13 Depressive Persönlichkeit .....	114
4.5.14 Histrionische Persönlichkeit.....	114
4.5.15 Zylothyme Persönlichkeit.....	114
4.5.16 Somatisierungstendenzen .....	115
4.5.17 Extrapunitivität, Intrapunitivität und Dominanz.....	115
4.5.18 Soziale Erwünschtheit (Normorientiertheit).....	115
4.5.19 Compliance .....	116
<b>4.6 Responseprädiktoren – Zusammenfassung der Metaanalyse.....</b>	<b>117</b>
<b>5 DATEN UND MESSINSTRUMENTE .....</b>	<b>120</b>
<b>5.1 Beschreibung der MAP-Studie .....</b>	<b>120</b>
5.1.1 Generelle Charakterisierung der MAP-Studie.....	120
5.1.2 Diagnostische Gruppen und Studienmedikation.....	120
5.1.3 Rekrutierung der Studienpatienten .....	121
5.1.4 Stabilisierungsphase .....	122
5.1.5 Beobachtungsphase .....	123
<b>5.2. Erhebungsinstrumente.....</b>	<b>126</b>
5.2.1 Diagnostik .....	126
5.2.2 Verlaufsbeurteilung .....	126
5.2.3 Drop-outs .....	127
5.2.4 AMDP-100.....	127
5.2.5 BPRS .....	128
5.2.6 BRMeS und BRMaS .....	128
5.2.7 BfS.....	128
5.2.8 GAF.....	128
5.2.9 MPT.....	128
5.2.10 KK.....	129
5.2.11 Globaleinschätzung der Compliance .....	130
5.2.12 SIS.....	130
5.2.13 CFI .....	131
5.2.14 Detaillierte Krankheitsanamnese.....	131
5.2.15 AMDP-Anamnese .....	132
5.2.16 Aktueller sozialer Status .....	133
5.2.17 Weitere Erhebungsinstrumente .....	133
<b>6 ERGEBNISSE AUS DER MAP-STUDIE.....</b>	<b>134</b>
<b>6.1 Studienpatienten.....</b>	<b>134</b>
<b>6.2 Therapieabbruch - Häufigkeit und Gründe.....</b>	<b>135</b>
6.2.1 Zeitpunkt des Therapieabbruchs .....	136
<b>6.3 Response-Prädiktion für Lithium.....</b>	<b>137</b>
6.3.1 Auswertungsstrategien.....	137
6.3.2 Univariate Analysen für einzelne Prädiktoren.....	141
6.3.2.1 Anzahl stationärer Episoden.....	142
6.3.2.2 Krankheitsdauer .....	143
6.3.2.3 Stimmungsinkongruente Wahnhalte in der Anamnese .....	144
6.3.2.4 Verlaufsmuster MDI.....	145
6.3.2.5 Verlaufsmuster DMI.....	146
6.3.2.6 Verlaufsmuster CC.....	147
6.3.2.7 Komorbide Persönlichkeitsstörung .....	148
6.3.2.8 Neurotizismus.....	149
6.3.2.9 Deprimiertheit oder Antriebsarmut .....	150

6.3.2.10 Arbeitslosigkeit .....	151
6.3.2.11 Isoliertes Auftreten der Episoden .....	152
6.3.2.12 Alter bei Ersterkrankung .....	153
6.3.2.13 Phasenfrequenz .....	154
6.3.2.14 Stress .....	155
6.3.2.15 Expressed emotions.....	156
6.3.2.16 Life events.....	157
6.3.2.17 Krankheitskonzepte.....	158
6.3.2.18 Soziale Unterstützung.....	159
6.3.2.19 Länge des ersten freien Intervalls.....	160
6.3.2.20 Zusammenfassung der univariaten Analysen.....	162
6.3.3 Multivariate Analysen .....	162
6.3.3.1 Logistische Regressionsmodelle.....	162
6.3.3.2 Logistische Regression mit vollem Variablensatz .....	163
6.3.3.3 Logistische Regression mit weitgehend den vollständig vorhandenen Variablen .....	164
6.3.3.4 Logistische Regression nach Datenimputation .....	166
6.3.3.5 Weitere multivariate Analysen, Jackknife-Methode.....	166
6.3.3.6 Multivariate Auswertungen für die Carbamazepingruppe .....	167
6.3.3.7 Zwischenfazit zu den multivariaten Analysen .....	167
6.3.4 Prädiktion des Verlaufs aufgrund einer Lithium-Response-Skala.....	168
6.3.4.1 Definition der Lithium-Response-Skala (LRS) .....	168
6.3.4.2 Ergebnisse bei Verwendung der Lithium-Response-Skala (LRS).....	171
6.3.4.3 Ergebnisse für Carbamazepingruppe.....	172
6.3.5 Zusammenfassung der empirischen Ergebnisse für Lithium .....	172
<b>7 DISKUSSION .....</b>	<b>174</b>
<b>8 LITERATURVERZEICHNIS .....</b>	<b>182</b>
<b>DANKSAGUNG.....</b>	<b>201</b>

## 0 Zusammenfassung

Die vorliegende Arbeit beinhaltet drei wesentliche Schritte. Zunächst wird eine systematische quantitative Literaturübersicht zur Prognose der Lithiumresponse durchgeführt. In einem zweiten Schritt werden auf Grundlage dieser systematischen Literaturübersicht verschiedene Prognoseinstrumente entwickelt. Diese Prognoseinstrumente umfassen eine für die klinische Anwendung optimierte Lithium-Response-Skala (LRS), sowie eine Reihe multivariater Verfahren. Im letzten Hauptabschnitt der vorliegenden Arbeit wird die praktische Eignung und prognostische Kraft dieser Prognoseinstrumente anhand eines umfangreichen Datensatzes evaluiert.

Im Literaturteil wurden in einer Serie von Metaanalysen insgesamt 21 (von 68 untersuchten) Variablen als Prädiktoren für die Lithiumresponse identifiziert. Die Mehrzahl dieser Prädiktoren war in den Übersichtsarbeiten der letzten Jahre nicht aufgeführt. Durch den systematischen Ansatz der vorliegenden Arbeit war es möglich, die zum Teil noch nicht erschlossene und teilweise in Vergessenheit geratene Primärliteratur zur Responseprädiktion zugänglich zu machen und zu aggregieren. Hierdurch ergaben sich Erkenntnisse, die klar über den bisherigen Forschungsstand hinausgehen. Auf Grundlage der Metaanalysen wurden die folgenden Variablen als Risikofaktoren identifiziert:

- Hohe Anzahl stationärer Episoden
- Hohe Phasenfrequenz
- Stimmungsinkongruente Wahninhalte in der Anamnese
- Verlaufsmuster DMI
- Verlaufsmuster CC
- Länge des ersten freien Intervalls
- Krankheitsdauer
- Persönlichkeitsstörung
- Stress
- High expressed emotions
- Hoher Neurotizismus
- Arbeitslosigkeit
- Life events
- Depressivität oder Verlangsamung

Demgegenüber erwiesen sich die folgenden Variablen anhand der durchgeführten Metaanalysen als protektive Faktoren:

- Verlaufsmuster MDI

- Isolierte Krankheitsepisoden
- Höheres Alter bei Ersterkrankung
- Hohe soziale Schicht
- Gute Compliance
- Soziale Unterstützung
- Dominanz

In einem nächsten Schritt wurde versucht, diese 21 klinischen und psychosozialen Variablen, die sich bei der Literaturübersicht als Prädiktoren erhärtet hatten, in geeigneter Weise zu verknüpfen. Leitendes Ziel war hierbei, eine für die klinische Anwendung optimierte Prädiktion der Lithiumresponse zu erreichen. Neben gängigen multivariaten Modellen (Cox-Regression, logistische Regression, Diskriminanzanalyse) wurde ein speziell für die vorliegende Problematik abgestimmtes Prognoseinstrument, die Lithium-Response-Skala (LRS), konstruiert.

Schließlich wurden im empirischen Teil dieser Arbeit prognostische Kraft und praktische Eignung der verschiedenen Prognoseinstrumente anhand eines umfangreichen unabhängigen Datensatzes vergleichend evaluiert. Im Ergebnis hat sich die Verwendung der Lithium-Response-Skala (additiv univariater Ansatz) mit Abstand am besten bewährt. Bei diesem Ansatz werden die verstreuten univariaten Einzelinformationen, die für sich genommen eine nur geringe prädiktive Kraft haben, gebündelt und so wird eine deutlich höhere Prognosegüte erreicht (Trefferquote von 75% in einem großen Datensatz). Hervorzuheben ist, daß diese Trefferquote anhand eines unabhängigen Datensatzes ermittelt wurde und somit eine realistische Schätzung für die prognostische Kraft der LRS darstellt. Hinsichtlich der Quote korrekter Prognosen war die LRS gegenüber multivariaten Standardverfahren deutlich im Vorteil – die Trefferquote etwa der Diskriminanzanalyse nach Durchführung einer Kreuzvalidierung lag wesentlich niedriger. Auch die einfache Handhabung und Interpretation sowie der flexible Umgang mit fehlenden Werten spricht eindeutig für die Lithium-Response-Skala (LRS). Dieses im Rahmen dieser Arbeit konstruierte Prognoseinstrument hat sich zumindest bei der Testung anhand einer unabhängigen Studie bewährt. Eine Erweiterung der LRS etwa durch biologische und genetische Faktoren wäre vielversprechend, da die Ergänzung der LRS durch weitere Prädiktoren mit einer Verbesserung der prognostischen Genauigkeit verbunden sein könnte.



## 1 Einleitung

Bipolare affektive Störungen sind in hohem Maße rezidivierende Erkrankungen. So liegt nach derzeitigen Schätzungen das Rezidivrisiko nach Ablauf einer manischen Episode bei mindestens 90% (Angst 1992, Solomon et al. 1995). Die Rezidivraten für bipolare Patienten ohne voll ausgeprägte Manie sind weniger gut untersucht, dürften aber ähnlich hoch liegen (Ayuso-Gutiérrez & Ramoz-Brieva 1982, Angst 1986, Vieta et al. 1997). Angesichts der schwerwiegenden, potenziell lebensbedrohlichen Risiken, die mit einer affektiven Episode verbunden sind, liegt es nahe, möglichst früh mit einer Rezidivprophylaxe zu beginnen.

Allerdings sind sowohl die derzeit etablierten Langzeittherapien (Lithium, Carbamazepin oder Valproat) als auch die gut untersuchten neueren Optionen zur Langzeitbehandlung wie Olanzapin oder Lamotrigin mit Belastungen und Risiken verbunden (Calabrese et al. 2002a, Sussmann 2003). An dieser Stelle sei nur an die Möglichkeit einer lebensbedrohlichen Intoxikation unter Lithium, eines Valproat-assoziierten Leberversagens und an allergische, in seltenen Fällen schwerwiegende Hautveränderungen unter Carbamazepin oder Lamotrigin erinnert (für eine ausführlichere Diskussion vgl. Greil et al. 2000). Diese Belastungen und Risiken sollten in jedem Einzelfall mit dem möglichen Nutzen abgewogen werden. Eine Aufgabe der klinischen Forschung ist es, objektive Kriterien zu definieren, die eine bessere Abwägung des zu erwartenden Nutzens gegen die Belastungen und Risiken erlauben.

Gegenstand der vorliegenden Arbeit ist die Beurteilung der Erfolgsaussichten einer Lithiumtherapie bei Untergruppen manisch-depressiver Patienten. Hiermit ist neben einer individuelleren Kosten-Nutzen-Abwägung eine weitere Zielsetzung verbunden: die Erleichterung differentialtherapeutischer Entscheidungen. Die Charakterisierung von Patienten, die mit hoher bzw. geringer Wahrscheinlichkeit auf Lithium respondieren, wäre ein wichtiger Schritt auf dem Weg zu einer prognosegestützten Differentialtherapie mit dem Ziel, jedem Patienten von Anfang an die individuell aussichtsreichste Therapie zukommen zu lassen.

Die Charakterisierung von Patienten mit guten vs. schlechten Aussichten auf eine erfolgreiche Lithiumprophylaxe ist auch im Kontext der aktuellen kontroversen wissenschaftlichen Diskussion um die generelle Wirksamkeit von Lithium von Interesse. Während Lithium in den meisten aktuellen Therapieleitlinien für bipolare Störungen nach wie vor als Standardtherapie oder Mittel der ersten Wahl in der Langzeittherapie empfohlen wird (z.B. Sachs & Thase 2000, Bowden 2003a), stellen manche Autoren die rezidivprophylaktische Wirkung von Lithium grundsätzlich in Frage (Moncrieff 1997). In diesen gegensätzlichen Positionen spiegeln sich möglicherweise gegensätzliche Erfahrungen mit deutlich unterschiedlichen Settings und Patientengruppen wider. Eine Annäherung dieser (für die

klinische Praxis eminent einflussreichen) Positionen könnte in einer Präzisierung der Umstände liegen, unter denen eine Lithiumtherapie gute Erfolgsaussichten hat. Es ist zu vermuten, daß hierbei neben dem allgemeinen Setting vor allem Patientenvariablen eine entscheidende Rolle zukommt.

Darüber hinaus sind Faktoren, die für das Ansprechen auf eine Lithiumtherapie relevant sind, auch von nosologischem Interesse, da sie Hinweise auf Subgruppen innerhalb der Gruppe manisch-depressiver Erkrankungen geben könnten (Mander 1986). Die wichtigste und unmittelbarste Zielsetzung der vorliegenden Arbeit ist allerdings, einen wissenschaftlichen Beitrag dazu zu leisten, daß bipolare Patienten von Anfang an die individuell für sie wirksamste Langzeittherapie erhalten.

Grundsätzlich stellt sich die Frage, ob eine weitere Charakterisierung von Lithium-Respondern überhaupt Aussicht auf Erfolg hat. Die Erkenntnis, daß es wichtig wäre, mehr über die Erfolgsaussichten einer Lithiumprophylaxe zu wissen ist ja keineswegs neu und dieses Thema wurde bereits in einer ganzen Reihe von Arbeiten untersucht (z.B. Calabrese et al. 1996, Walden et al. 1999, Ketter & Wang 2002). Warum also noch eine weitere Arbeit zur Response-Prädiktion? Es sind vor allem zwei Aspekte, die zumindest den Versuch rechtfertigen, das Thema noch einmal in den Mittelpunkt einer Untersuchung zu stellen. Zum einen sind viele Lithiumstudien zwar publiziert, aber noch gar nicht oder nicht vollständig unter dem Aspekt der Response-Prädiktion ausgewertet. In der vorliegenden Arbeit wird erstmals der Versuch unternommen, alle Langzeitstudien, die eine Verknüpfung von Lithium-Response und klinischen oder psychosozialen Faktoren erlauben, systematisch zu erfassen und die Erkenntnisse aus diesen Studien zu integrieren. Ein zweiter Punkt, der die vorliegende Arbeit spannend macht, ist die Möglichkeit die aus der Literatur gewonnenen Erkenntnisse an einem umfangreichen unabhängigen Datensatz empirisch zu überprüfen. Wie diese Überprüfung ausfallen wird, und ob sich die Lithium-Response auf einer systematischeren Grundlage tatsächlich besser voraussagen lässt, müssen die Daten zeigen. Auf jeden Fall lohnt der Versuch.

## 2 Theoretischer Hintergrund

### 2.1 Diagnostik und Verlauf bipolar affektiver Störungen

In den derzeit maßgebenden Klassifikationssystemen wie DSM-IV und ICD-10 beruht die Abgrenzung bipolar affektiver Störungen von anderen psychiatrischen Erkrankungen primär auf dem Erscheinungsbild der Krankheitsepisoden (Manie, Hypomanie, Depression, Mischzustand) und dem Langzeitverlauf. Eine ätiologische Klassifikation ist demgegenüber weitgehend in den Hintergrund getreten.

Die Übergänge zwischen den verschiedenen Erscheinungsbildern und Verlaufsmustern sind fließend und so ist es schwierig eine klare Abgrenzung zu treffen. Dies spiegelt sich in ausgeprägten diagnostischen Trends der letzten 30-40 Jahre. Nach den grundlegenden Arbeiten vor allem der Arbeitsgruppen um Angst, Perris, Clayton und Winokur in den 60er und frühen 70er Jahren hat sich die Aufspaltung der affektiven Störungen in unipolar depressive und bipolar manisch-depressive Verläufe durchgesetzt (Angst 1966, Perris 1966, Winokur & Clayton 1967). Unipolar manische Verläufe wurden über einige Jahre als eigenständige Kategorie diagnostiziert. Naturalistische Studien zum Langzeitverlauf haben allerdings ergeben, daß fast jeder manische Patient im späteren Verlauf eine Depression erleidet und deshalb – langfristig betrachtet – als bipolar zu diagnostizieren ist. Grundlegend für die Diagnose einer bipolaren Störung nach moderner Diagnostik ist daher das Auftreten einer manischen oder einer gemischten (gleichzeitige manische und depressive Symptomatik) Episode.

In den 70er und 80er Jahren bestand ein klarer Trend zur Ausweitung der bipolar affektiven Störungen insbesondere in der anglo-amerikanischen Diagnostik (vgl. etwa Stoll et al. 1993). Während traditionell das Vorhandensein stimmungsinkongruenter Wahninhalte als Zeichen einer Störung aus dem schizophrenen Formenkreis galt, wurden nunmehr vermehrt bipolare Störungen mit stimmungsinkongruenten Wahninhalten diagnostiziert. Der Trend zur diagnostischen Ausweitung und Differenzierung bipolarer Störungen hält weiter an. Die Hauptverschiebung findet hierbei innerhalb der affektiven Störungen statt (vgl. etwa Marneros 2001, Angst et al. 2003a). Heute besteht die Tendenz auch beim Vorhandensein leichterer manischer Syndrome eine bipolare (nicht mehr unipolare) Störung zu diagnostizieren. Ähnliche Tendenzen bestehen bei der Abgrenzung von Borderline Persönlichkeitsstörungen zu bipolaren Störungen (Deltito et al. 2001). Wie dramatisch sich diese Entwicklungen auf die Prävalenz bipolarer Störungen auswirken, zeigt Abbildung 1. Diese veränderte diagnostische Praxis hat weitreichende therapeutische Konsequenzen – allein schon weil die Studien und

Erfahrungen mit früheren manisch-depressiven Patienten aufgrund der zwischenzeitlichen Verschiebungen nur noch bedingt auf heutige bipolare Patienten übertragbar sind.

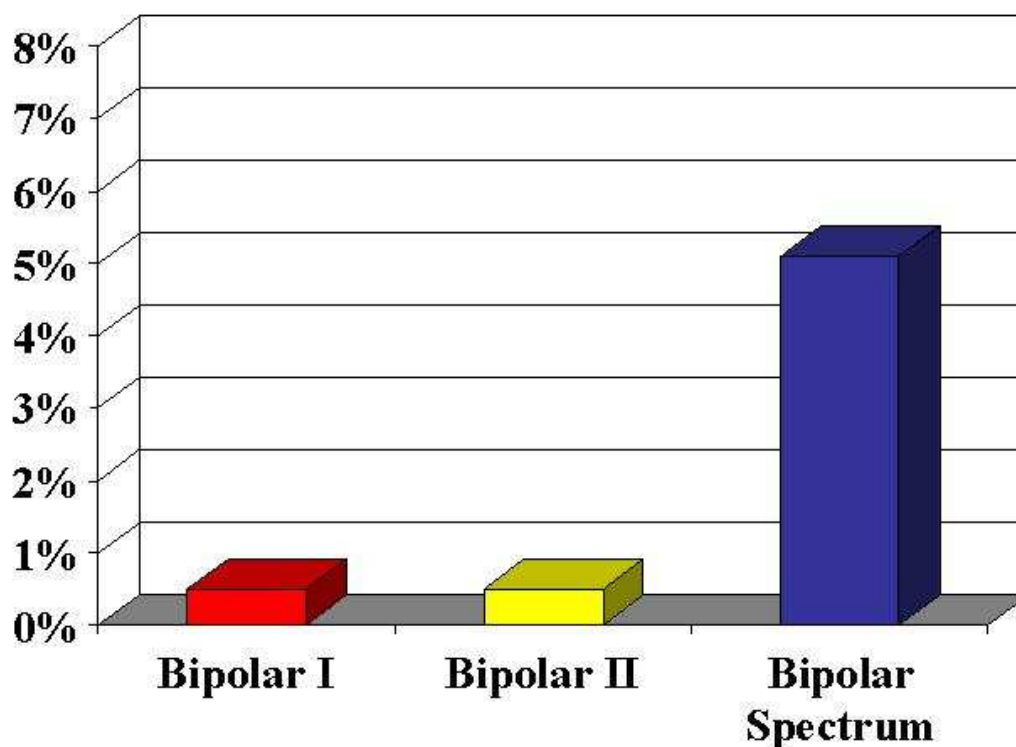


Abbildung 1: Mittlere Lebenszeitprävalenzen (Median) für verschiedene Subtypen bipolar affektiver Störungen (aus Greil & Kleindienst, 2003).

In der Regel verlaufen bipolare Störungen episodisch, d.h. in Krankheitsphasen, die nach einiger Zeit wieder abklingen. Die Mehrzahl der Patienten durchlebt einen rezidivierenden Verlauf mit weitgehend hergestellter Lebensqualität und Leistungsfähigkeit in der Zeit zwischen den Phasen. In der jüngeren Forschung wird auch die Symptomatik nach dem Abklingen einer Krankheitsphase (Intervallsymptomatik) stärker beachtet (Fava et al. 1999, Kleindienst & Greil 2002) und hierbei hat sich gezeigt, daß auch bei bipolaren Störungen ein erheblicher Anteil an Patienten durch Intervallsymptomatik in Lebensqualität und Leitungsfähigkeit beeinträchtigt ist.

Insgesamt ist mit interindividuell sehr unterschiedlichen Verläufen zu rechnen. Diese Heterogenität der Verläufe findet zu einem gewissen Teil ihren Niederschlag in der Diagnostik. So wird nach den für die vorliegende Auswertung zugrunde gelegten Klassifikationssystemen DSM-III-R und DSM-IV zwischen einer Bipolar-I-Störung (definiert durch mindestens eine manische oder gemischte Episode), einer Bipolar-II-Störung

(mindestens eine depressive und eine hypomane, aber bisher keine manische oder gemischte Episode) und weiteren, nicht näher bezeichneten Krankheitsbildern unterschieden. Die Klassifikation weiterer bipolarer Verlaufstypen ist derzeit Gegenstand intensiver Diskussion und Forschung (z.B. Marneros 2001, Akiskal 2003, Angst et al. 2003a,b).

Zur Zeit der Durchführung der MAP-Studie (vgl. Greil et al. 1993, sowie Kap. 5.1) waren die Klassifikations-systeme ICD-9 (WHO) und DSM-III-R maßgebend. Da beim Übergang von DSM-III-R zu DSM-IV bei der Diagnostik bipolarer Störungen nur geringfügige Verschiebungen aufgrund von Symptomen vorgenommen wurden, die durch das verwendete Diagnoseinstrument (SKID-I für DSM-III-R) detailliert erfasst wurden, wird in der vorliegenden Arbeit die Terminologie des DSM-IV zugrunde gelegt. Die Entsprechungen zwischen DSM-III-R und DSM-IV sind in Tabelle 1 dargestellt.

Tabelle 1: Entsprechung der Diagnosen in den Klassifikationssystemen DSM-III-R und DSM-IV.

DSM-III-R		DSM-IV				
Bipolar	$\xrightarrow{\text{entspricht}}$	Bipolar I				
Bipolar NNB*	$\xrightarrow{\text{entspricht}}$	<table style="border: none;"> <tr> <td style="font-size: 3em; vertical-align: middle;">{</td> <td style="padding-left: 10px;">Bipolar II</td> </tr> <tr> <td></td> <td style="padding-left: 10px;">Bipolar NNB *</td> </tr> </table>	{	Bipolar II		Bipolar NNB *
{	Bipolar II					
	Bipolar NNB *					

\*Bipolar NNB = Nicht Näher Bezeichnete Bipolare Störung

## 2.2 Allgemeines zur medikamentösen Therapie bipolar affektiver Störungen

### 2.2.1 Akuttherapie, Erhaltungstherapie und Rezidivprophylaxe

In der Pharmakotherapie bipolarer affektiver Störungen unterscheidet man zwischen Akuttherapie, Erhaltungstherapie und Rezidivprophylaxe (Coppen & Peet 1979, Greil & Schmidt 1985, Kupfer 1991). Dieser Unterscheidung liegt die Vorstellung zugrunde, daß eine antidepressive bzw. antimanische Medikation die Symptomatik unterdrückt, den zugrundeliegenden Krankheitsprozess hingegen weitgehend unbeeinflusst lässt (vgl. Abb. 2). Wenn nach erfolgreicher Therapie der akuten Symptomatik (Akuttherapie) die Behandlung beendet wird, bevor es dem natürlichen Krankheitsverlauf entsprechend zum vollständigen Abklingen der Krankheitsphase gekommen wäre, muss mit dem Wiederauftreten der Symptomatik gerechnet werden (Rückfall). Dies wird durch Untersuchungen bestätigt, in denen die Patienten nach erfolgreicher Akutbehandlung einer depressiven Episode entweder

antidepressiv weiterbehandelt oder auf Placebo umgesetzt wurden. In diesen Studien ergab sich eine deutlich erhöhte Rückfallquote in der Placebogruppe (Übersichten: Prien 1988, Greil & Kleindienst 1997). Dabei muss allerdings berücksichtigt werden, daß durch das abrupte Absetzen der wirksamen Medikation Absetzeffekte aufgetreten sein könnten.

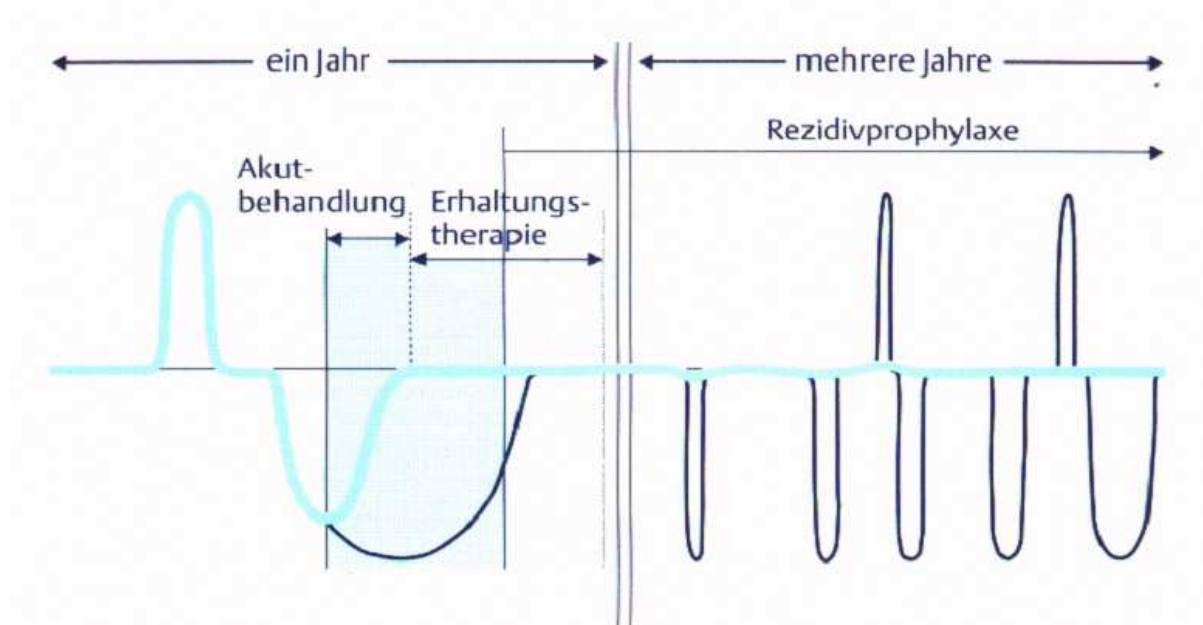


Abbildung 2: Akuttherapie, Erhaltungstherapie und Rezidivprophylaxe bei affektiven Störungen (aus Greil et al. 1996b).

Die medikamentöse Behandlung sollte solange fortgesetzt werden, bis die Krankheitsphase auch unbehandelt abgeklungen wäre (Erhaltungstherapie). Dieser Zeitpunkt lässt sich jedoch nicht bestimmen und muss nach allgemeinen Erfahrungswerten abgeschätzt werden. In der Regel wird empfohlen, an eine erfolgreiche Akuttherapie eine Erhaltungstherapie über mindestens sechs Monate anzuschließen, um ein Rezidiv in die aktuelle Krankheitsphase zu verhindern (etwa Prien 1988).

Während die Erhaltungstherapie dazu dient, einen Rückfall in die gegenwärtige Krankheitsphase zu verhindern, soll die Rezidivprophylaxe das Auftreten zukünftiger Phasen (Rezidive) verhüten.

### 2.2.2 Medikamentöse Rezidivprophylaxe bei bipolaren Störungen

Das Spektrum an Behandlungsmöglichkeiten mit Phasenprophylaktika hat sich in jüngster Zeit stark erweitert. Neben dem klassischen Lithium und den auch als Phasenprophylaktika etablierten Antikonvulsiva Carbamazepin und Valproat (Post et al. 1996, Post et al. 1997,

Bowden et al. 2000) liegen inzwischen positive Befunde für Lamotrigin und für Olanzapin aus großen randomisierten Studien vor (Bowden et al. 2003b, Tohen et al. zur Publikation eingereicht; vgl. Tab. 2). Weitere Behandlungsstrategien sind derzeit Gegenstand intensiver Forschung. Dabei werden insbesondere Kombinationstherapien (vgl. Freeman u. Stoll 1998, Calabrese 2002b) sowie weitere Antiepileptika (z.B. Oxcarbazepin, Gabapentin), Calcium-Blocker (z.B. Nimodipin, Verapamil) und weitere atypische Neuroleptika (z.B. Clozapin, Risperidon, Quetiapin) als erfolversprechende Alternativen diskutiert oder in kontrollierten Therapiestudien untersucht.

Tabelle 2: Mögliche Substanzen zur Rezidivprophylaxe bipolarer affektiver Störungen

Substanz	Bemerkungen
Lithium	<ul style="list-style-type: none"> <li>• Wirksamkeit durch eine Vielzahl kontrollierter Studien nachgewiesen</li> <li>• ausgeprägte antisuizidale Wirkung</li> </ul>
Carbamazepin	<ul style="list-style-type: none"> <li>• Wirksamkeit durch kontrollierte Studien belegt</li> </ul>
Olanzapin	<ul style="list-style-type: none"> <li>• Wirksamkeit durch eine große randomisierte Studie nachgewiesen</li> <li>• ausgeprägte antimanische Wirkung</li> </ul>
Lamotrigin	<ul style="list-style-type: none"> <li>• Wirksamkeit durch eine große randomisierte Studie nachgewiesen</li> <li>• ausgeprägte antidepressive Wirkung</li> <li>• möglicherweise besonders günstig bei rapid cycling</li> </ul>
Valproat	<ul style="list-style-type: none"> <li>• Wirksamkeit aufgrund offener Studien und klinischer Erfahrung angenommen</li> <li>• möglicherweise besonders günstig bei rapid cycling</li> </ul>
Clozapin? Gabapentin? Nimodipin? Verapamil? Topiramal?	<ul style="list-style-type: none"> <li>• Wirksamkeit dieser mit Fragezeichen versehenen Substanzen aufgrund einzelner Fallberichte oder offener Studien vermutet</li> <li>• Kalziumantagonisten (Nimodipin, Verapamil) möglicherweise speziell bei rapid cycling stimmungsstabilisierend</li> <li>• Nimodipin möglicherweise speziell bei ultra rapid cycling stimmungsstabilisierend</li> <li>• Gabapentin möglicherweise zusätzlich antimanisch wirksam</li> <li>• Clozapin möglicherweise auch bei schizoaffektiven Störungen wirksam</li> </ul>

## 2.3 Carbamazepin

Carbamazepin wurde 1957 bei der gezielten Suche nach neuen Antiepileptika entdeckt. Carbamazepin ist eine wenig wasserlösliche Neutralsubstanz, die strukturell chemisch klassischen trizyklischen Antidepressiva wie Imipramin ähnlich ist (s. Abb. 3).

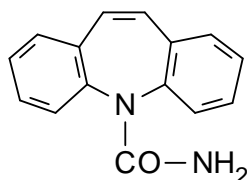


Abbildung 3: Strukturformel von Carbamazepin.

Nachdem sich Carbamazepin als wirksames und relativ sicheres Antiepileptikum erwiesen hatte, wurde es 1963 in der Schweiz unter dem Handels-Namen „Tegretol“ und 1964 in Deutschland unter dem Handels-Namen „Tegretal“ eingeführt.

Bei der Epilepsiebehandlung mit Carbamazepin wurden auch positive Effekte auf psychische Störungen beobachtet, wie sie häufig im Rahmen einer epileptischen Erkrankung auftreten. So wurden unter Carbamazepin eine Verbesserung kognitiver Fähigkeiten und eine Abnahme von Ängstlichkeit und depressiver Verstimmung festgestellt. Inzwischen ist Carbamazepin als Therapie bei einer ganzen Reihe von neurologischen und psychiatrischen Störungen etabliert und zugelassen. Neben der Therapie von Epilepsien (insbesondere von bestimmten Formen der Epilepsie wie von fokalen Epilepsien) wird Carbamazepin bei einer Vielzahl weiterer neurologischer Indikationen eingesetzt. Carbamazepin ist eine wichtige Behandlungsoption unter anderem bei bestimmten Neuralgien (z.B. Trigeminusneuralgien oder Glossopharyngeusneuralgien), bei Migräne oder bei Multipler Sklerose (vgl. Tab. 3).

Auch in der Psychiatriebehandlung ist Carbamazepin fest etabliert – unter anderem in der Akutbehandlung der Manie und in der Rezidivprophylaxe bipolarer und schizoaffektiver Störungen (vgl. Tab. 3).



Tabelle 3: Einsatzgebiete von Carbamazepin

*Psychiatrische Einsatzgebiete*

- Prophylaxe bei bipolaren Störungen
- Prophylaxe bei schizoaffektiven Störungen
- Behandlung leichter und mittelschwerer Manien
- Alkoholentzugs-Syndrom
- Möglicherweise besteht eine leichte antidepressive Wirkung
- Möglicherweise wirksam bei Schizophrenien und Verhaltensstörungen

*Neurologische Einsatzgebiete*

- Epilepsien
- Bestimmte Nervenschmerzen (z.B. bei Trigeminus-Neuralgie oder Diabetes insipidus)
- Mono- oder Polyneuropathien
- Multiple Sklerose
- Migräne, Kopfschmerzen

## 2.4 Lithium

### 2.4.1 Allgemeines zu Lithium

Lithium ist ein in der Natur vorkommendes chemisches Element aus der Reihe der Alkalimetalle. Lithium kommt natürlicherweise in Form von Salzen vor, z.B. als Lithium-Carbonat, Lithium-Azetat oder als Lithium-Sulfat. Diese verschiedenen Lithiumsalze werden auch als Phasenprophylaktika verwendet. Für die Wirkung entscheidend ist das im Salz enthaltene Lithiumion mit seinen den Kationen  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Ca}^{++}$  und  $\text{Mg}^{++}$  ähnlichen biochemischen Eigenschaften.

### 2.4.2 Lithium in der Behandlung bipolarer Störungen

Lithium wirkt in allen Stadien des Krankheitsverlaufs bipolarer Störungen. Sowohl in der Akuttherapie als auch in der Erhaltungstherapie und Phasenprophylaxe. In der Akuttherapie der Manie ist bereits Lithium alleine hochwirksam. Falls kein ausreichender oder ausreichender schneller Behandlungserfolg mit Lithium oder einer anderen Monotherapie erreicht wird, eignet sich Lithium gut zur Kombination mit weiteren Medikamenten.

In der Akutbehandlung der Depression ist Lithium vor allem als Zusatztherapie zu einem Antidepressivum wirksam. Ein hoher Anteil von Patienten, für die ein Antidepressivum allein nicht ausreichend wirksam ist, sprechen gut auf eine Augmentation mit Lithium an. Dies hat darüber hinaus den Vorteil, daß Lithium eine ungewollte Wirkung mancher Antidepressiva verhindern kann: das direkte Umschlagen von einer Depression in eine Manie (vgl. etwa Bottlender et al. 2001).

Etwa 20% der manisch-depressiven Patienten versterben früher oder später an einem Suizid. Lithium scheint im Vergleich zu den anderen stimmungsstabilisierenden Therapien die Suizidrate stärker zu senken (vgl. etwa Goodwin et al. 2003, Müller-Oerlinghausen et al. 2003). Bei einer in vielen Fällen aufgrund eines Suizides tödlich verlaufenden Erkrankung ist diese antisuizidale Wirkung von größter Bedeutung.

Ein vierter wichtiger Pol (vgl. Abb. 4) im Verlauf manisch-depressiver Erkrankungen ist das Auftreten psychotischer Merkmale (z.B. Größenideen oder Wahnvorstellungen). In der Regel bietet Lithium in dieser Hinsicht keinen guten Schutz (vgl. Tab. 4), so daß man in diesen Fällen auf andere Medikamente zurückgreifen muss. Auch eine Kombination von Lithium mit einer antipsychotischen Substanz (z.B. Valproat oder Neuroleptika) bietet sich für diese Fälle an.

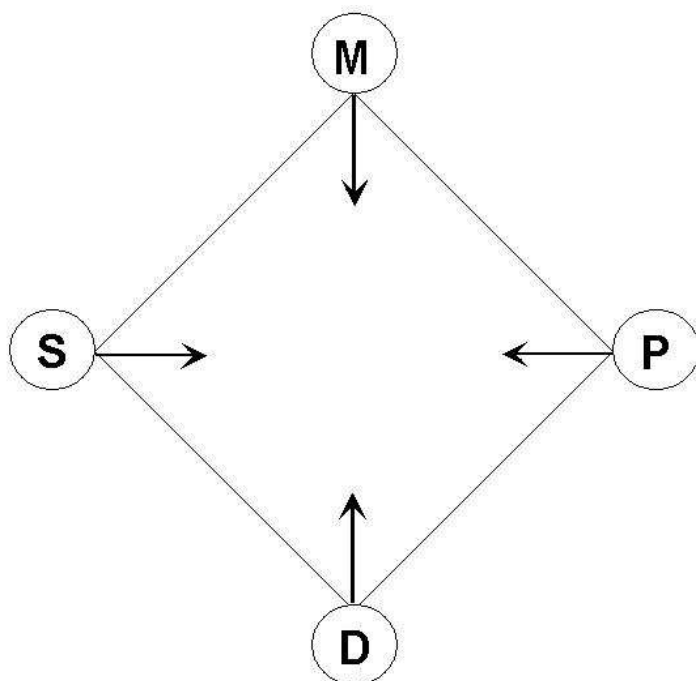


Abbildung 4: Hauptziele bei der Behandlung bipolarer Störungen: M = Manischer Pol, D = Depressiver Pol, S = Suizidalität, P = Psychotische Merkmale (aus Greil & Kleindienst, 2003).

Tabelle 4: Klinische Wirkungen von Lithium in der Behandlung bipolarer Störungen

- |   |
|---|
| <ul style="list-style-type: none"> <li>• Antimanische Wirkung</li> <li>• Verstärkung der antidepressiven Wirkung von Antidepressiva</li> <li>• Prophylaktische Wirkung gegen Depressionen</li> <li>• Prophylaktische Wirkung gegen Manien</li> <li>• Prophylaktische Wirkung gegen Suizide</li> </ul> |
|---|

#### 2.4.2.1 Langzeitbehandlung bipolarer Störungen mit Lithium

Nach Abklingen der Manie oder Depression verringert Lithium die Wahrscheinlichkeit für das Wiederauftreten des Krankheitsbildes – sowohl von Rückfällen als auch von Rezidiven. Die vieldiskutierte These von Coryell et al. (1997), daß Lithium lediglich erhaltungstherapeutische, aber keine rezidivprophylaktische Wirkung haben könnte, hat sich nicht bestätigt. Eine Re-Analyse der von Coryell et al. (1997) vorgestellten Daten zeigte, daß diese Daten – im Gegensatz zu der in der Originalarbeit vorgenommenen Interpretation – mit einer nachhaltigen prophylaktischen Wirksamkeit von Lithium in Einklang stehen (Kleindienst et al. 1999). Dies schließt nicht aus, daß in Einzelfällen auch nach Jahren der Rezidivfreiheit unter Lithium erneut Rezidive auftreten (Post et al. 1993 Maj et al. 1989). Die plausibelste Erklärung für dieses in der Angelsächsischen Literatur unter den Stichworten "late non-response" und "loss of efficacy" diskutierte Phänomen ist nicht ein Wirkverlust, sondern ist vielleicht eher in Schwankungen in der Intensität des Krankheitsprozesses zu vermuten (Kleindienst et al. 1999, Maj 2000).

#### 2.4.2.2 Zur Kontroverse um die prophylaktische Wirksamkeit von Lithium

Da die rezidivprophylaktische Wirksamkeit von Lithium Gegenstand einer aktuellen, kontroversen wissenschaftlichen Diskussion ist, und da diese Diskussion für die Relevanz der vorliegenden Arbeit von großer Bedeutung ist, soll hierauf im folgenden kurz eingegangen werden (für eine ausführlichere Darstellung vgl. Kleindienst & Greil 2003).

Die Wirksamkeit von Lithium in der Rezidivprophylaxe bipolarer Störungen wurde zunächst in den 1970er Jahren durch eine Vielzahl kontrollierter Studien belegt. Die Ergebnisse waren sehr klar und führten zur amtlichen Zulassung von Lithium in den USA und in allen größeren europäischen Staaten. In 8 von 10 dieser klassischen randomisierten Therapievergleichsstudien war Lithium signifikant besser als Placebo. In den beiden anderen Studien (Melia 1970 und Dunner et al. 1976) zeigte sich ein Trend zugunsten von Lithium, aber die Fallzahlen in diesen Studien waren für eine inferenzstatistische Absicherung dieser Befunde zu gering.

Die eindrucksvollen Ergebnisse der kontrollierten Prophylaxestudien haben sich in naturalistischen Studien der folgenden Jahre nicht bestätigt (Darcourt 1995, Post et al. 1996). Dies führte dazu, daß die prophylaktische Wirkung von Lithium Mitte der 1990er Jahre generell in Frage gestellt wurde (Moncrieff 1995, 1997) und auch die klassischen Therapievergleichsstudien aus den 1970er Jahren in methodischer Hinsicht kritisiert wurden. Moncrieff (1995, 1997) argumentiert, daß diese Studien aktuellen methodischen Standards nicht mehr entsprechen und führt drei Hauptkritikpunkte an:

1. Bei etwa der Hälfte der Studien handelt es sich um Absetzstudien
2. Die klassischen Studien wurden nicht als Doppelblindstudien durchgeführt.
3. Die Ergebnisse werden durch eine hohe Rate an Studienabbrechern verfälscht.

Der erste Punkt ist gravierend. In vielen Lithiumstudien der 1970er wurden die auf den Placeboarm randomisierten Patienten abrupt von Lithium auf Placebo umgesetzt, während die Patienten im Lithiumarm weiterhin Lithium erhielten. Spätere Studien (Klein et al. 1981, Greil et al. 1982, Suppes et al. 1991) haben gezeigt, daß das abrupte Absetzen von Lithium speziell das Risiko für manische Rückfälle und Rezidive erhöht, und hierdurch könnte es zu einer erhöhten Rückfallrate in den Placebogruppen gekommen sein. Eine genaue Analyse der Originaldaten aus den frühen Studien (z.B. Schou et al. 1970) zeigt allerdings, daß dieses Phänomen in den klassischen Studien wohl weniger ausgeprägt war als in den späteren Studien. Absetzphänomene scheinen speziell bei Patienten mit atypischen bipolaren Störungen und bei nicht vollständiger Stabilisierung unter Lithium aufzutreten, und weniger bei den klassischen manisch-depressiven Patienten aus den frühen Studien (Schou 1998). Dennoch bleibt dieser Kritikpunkt von Moncrieff (1995, 1997) insgesamt bestehen, wobei zu beachten ist, daß sie nur auf die Absetzstudien zutrifft und damit etwa die Hälfte der frühen Studien nicht berührt.

Eine moderne Therapievergleichsstudie würde in jedem Fall ein rasches Absetzen der Studienmedikation vermeiden, und in dieser Hinsicht entspricht ein Teil der frühen Studien nicht mehr modernen Standards. Bezüglich der beiden anderen Kritikpunkte ist allerdings einzuwenden, daß wir auch heute noch keines dieser Probleme wirklich im Griff haben – weder die Entblindung in Langzeitstudien mit Medikamenten, die wie Lithium ein charakteristisches Nebenwirkungsprofil aufweisen, noch die mögliche Verfälschung der Ergebnisse durch Studienabbrecher. Moncrieff (1995, 1997) fordert für den Wirknachweis eine doppelblinde Placebo-kontrollierte Studie, bei der die Verzerrung durch Studienabbrecher weitgehend eliminiert wurde. Nach den bisherigen Erfahrungen erscheint es allerdings unwahrscheinlich, daß bei einer Beobachtungszeit von mindestens 6 Monaten die Verblindung sowohl in Bezug auf den Patienten als auch auf den behandelnden Arzt aufrecht

erhalten werden kann (Marini et al. 1976, Watkins et al. 1987, Oxtoby et al. 1989, Double 1996). Auch ist damit zu rechnen, daß die drop-out Rate in der Placebogruppe höher ist. Der hierdurch entstehende Bias ist nicht quantifizierbar und damit praktisch nicht eliminierbar, auch nicht durch den Einsatz moderner statistischer Verfahren (hierzu vgl. Kleindienst & Greil 2002). Der Verfasser ist der Ansicht, daß die Ergebnisse einer randomisierten Therapievergleichsstudie auch dann valide sein können, wenn die Studie ohne doppelte Verblindung durchgeführt wurde und wenn die drop-out Rate in der Placebo-Gruppe höher liegt als in der Verum-Gruppe. Dementsprechend wird mit einer Vielzahl anderer Autoren (etwa mit Maj 2000, Pies 2002) die Meinung vertreten, daß die prophylaktische Wirksamkeit von Lithium wissenschaftlich gut belegt ist. Diese Auffassung wird durch eine neuere dreiarmlige (doppelblinde) Therapievergleichsstudie (Lamotrigin vs. Lithium vs. Placebo) gestützt (Bowden et al. 2003), in der Lithium eine signifikante Überlegenheit über Placebo zeigte ( $p = 0.003$ ).

Ein Punkt, auf den die Lithiumkritiker aufmerksam gemacht haben, bleibt allerdings bestehen: Die Ergebnisse neuerer Studien zur Langzeitbehandlung mit Lithium sprechen für eine geringere Effektivität als man aufgrund der Studien aus den 1970er Jahren annehmen konnte. Dies hängt mutmaßlich mit zwei Umständen zusammen: Erstens mit dem wesentlich ungünstigeren Setting einiger späterer Studien (Routinebedingungen in den Feldstudien etwa der 1980er Jahre vs. hochbetreutes universitäres Setting in den 1970ern (vgl. hierzu etwa Guscott & Taylor 1994, Grof 1998, Schou 1998). Und vermutlich auch mit der in Kap. 2.1 besprochenen Veränderung der diagnostischen Praxis. Es ist nicht von vornherein anzunehmen, daß Lithium in einer weit größeren Gruppe bipolarer Patienten ebenso wirksam ist wie in einer bipolaren Kerngruppe (vgl. Gershon u. Soares 1997, Greil et al. 1998, Grof 1998). Dieser Punkt ist im Kontext der vorliegenden Arbeit von besonderem Interesse: Insbesondere durch die Veränderungen der diagnostischen Praxis ist derzeit unklar, inwieweit sich die Ergebnisse der frühen Studien auf Patienten übertragen lassen, die nach heutigen Kriterien als bipolar diagnostiziert werden.

#### 2.4.2.3 Primärarbeiten zur Prädiktion der prophylaktischen Wirksamkeit von Lithium

Der Versuch, den Erfolg einer Lithiumprophylaxe mit bestimmten Prädiktoren in Zusammenhang zu bringen, reicht zurück bis in die frühen Studien zur Rezidivprophylaxe affektiver Störungen. So stellen bereits Baastrup & Schou (1967) eine höhere Rezidivneigung bipolarer Patienten mit stimmungsinkongruenten Wahninhalten fest. Seither wurde eine Vielzahl von Befunden zur Charakterisierung von Respondern und Non-Respondern publiziert. Diese Primärstudien umfassen ein weites Spektrum von möglichen Prädiktoren. In den frühen Studien wurde die Lithiumresponse vor allem mit diagnostischen, klinischen und

demographischen Variablen in Verbindung gebracht (Melia 1970, Van der Velde 1970, Cundall et al. 1972, Persson 1972, Prien et al. 1973, 1974, Dunner & Fieve 1974). Diese Studien dienten primär dem Nachweis der globalen Wirksamkeit und umfassten in der Regel eine nur sehr geringe Zahl von Patienten. Nicht zuletzt aufgrund der geringen Fallzahlen zeigten sich insgesamt nur wenige Unterschiede zwischen Lithiumrespondern und Non-Respondern. Ein, aus diesen im Vergleich zu den neueren Studien diagnostisch heterogenen Einzelstudien der 60er und 70er Jahre interessantes Ergebnis, ist der potenziell starke Einfluss diagnostischer Faktoren auf die Lithiumresponse. Demnach war die klare Diagnose einer primären affektiven Störung für den Erfolg einer Lithiumprophylaxe mit entscheidend (Ananth & Pecknold 1978) – ein Befund, der durch eine Serie von Untersuchungen der Forschungsgruppe um Paul Grof bestätigt wurde (etwa Grof et al. 1983a). Weitere Befunde zur Diagnostik, etwa ein besseres Ansprechen bipolarer im Vergleich zu unipolaren Patienten (Goodwin et al. 1972, Baron et al. 1975, Mendels 1976) haben sich später nicht in dieser Deutlichkeit bestätigt (vgl. etwa Page et al. 1987, Brochier & Gay 1991, Marneros 1999, S. 248) und sind nach wie vor umstritten.

Eine Sonderstellung in der frühen Prädiktorforschung nehmen die Arbeiten zur sog. lithium-ratio (dem Verhältnis der Lithiumkonzentration in Erythrozyten und Plasma) ein. Hierbei handelt es sich um einen, mit Transport und Retention in den Zellen assoziierten, biologischen Marker. Im Anschluss an die Arbeit von Mendels & Frazer (1973) wurde eine Vielzahl gezielter Untersuchungen zum prädiktiven Wert des Verhältnisses dieser lithium-ratio durchgeführt (z.B. Rybakowski & Strzyzewski 1976, Rybakowski et al. 1977, Pandey et al. 1978). Insgesamt erwiesen sich jedoch die Ergebnisse als zu wenig konsistent, um diesen biologischen Marker zweifelsfrei als Prädiktor zur etablieren (Dorus & Pandey 1980). Die weitere Erforschung biologischer Prädiktoren der Lithiumresponse – etwa mittels Dexamethason-Suppressions-Test, der Lautstärkeabhängigkeit akustisch evozierter kortikaler Potentiale oder mit bildgebenden Verfahren – erbrachte vielversprechende, wenn auch nicht durchschlagende Ansätze, die allerdings nicht Gegenstand der vorliegenden Arbeit sind (hierzu vgl. etwa Greden et al. 1983, Hegerl & Mavrogiorgou 1997, Ketter & Wang 2002).

In den Primärstudien der 80er und 90er Jahre wurde vermehrt der Einfluss der Krankheitsanamnese, sowie genetischer und psychosozialer Faktoren auf die Lithiumresponse untersucht. Im Bereich der Krankheitsanamnese standen vor allem zwei Variablen im Zentrum der Aufmerksamkeit – die Anzahl/Frequenz affektiver Episoden und das Verlaufsmuster. Zum prognostischen Wert von Anzahl oder Frequenz affektiver Episoden in der Vorgeschichte liegt eine Vielzahl von Arbeiten vor (z.B. Faravelli et al. 1980, Abou-Saleh 1983, Maj et al. 1986, Mander 1986, Gelenberg et al. 1989, Kulhara et al. 1999). Die Ergebnisse sind widersprüchlich (etwa Abou-Saleh & Coppen, 1986 vs. Yaziki et al. 1999), was sich in diesem Fall am plausibelsten mit der unterschiedlichen Art der outcome-Messung

erklären lässt (vgl. hierzu Kap. 4.4.2.5). Insgesamt hat sich die Auffassung durchgesetzt, daß eine hohe Anzahl oder Frequenz (insbes. ein sogenanntes rapid cycling) prognostisch ungünstig ist (etwa Sarantidis & Waters 1981). Bezüglich des Verlaufsmusters berichteten Kukopulos & Reginaldi (1980a,b) über einen positiven Einfluss des „klassischen“ Musters MDI (d.h. einer Episodenfolge Manie – Depression – freies Intervall) auf die Lithiumresponse. Dieser Befund wurde mehrfach repliziert (insbes. Haag et al. 1986, Grof et al. 1987) und gilt als weitgehend gesichert (vgl. insbes. Faedda et al. 1991). Eine umfassende quantitative Literaturübersicht zu diesen und zu weiteren prädiktiven Faktoren der Krankheitsanamnese wird in Kapitel 4.4.2 gegeben.

Da genetische Faktoren im Rahmen dieser Arbeit nicht untersucht werden, sei an dieser Stelle auf die exzellenten Übersichtsarbeiten von Alda (2001) und von Serretti & Artioli (2003) zum Zusammenhang von Molekulargenetik und Lithiumresponse verwiesen. Eine Literaturübersicht zum Zusammenhang zwischen Response und familiärer Belastung mit psychiatrischen Erkrankungen findet sich in den Übersichtsarbeiten von Coryell & Winokur (1980) und Goodwin & Jamison (1990), sowie in der Originalarbeit von Grof et al. (2002).

Für die Erforschung des Zusammenhangs zwischen psychosozialen Faktoren und Lithiumresponse ergibt sich ein relativ uneinheitliches Bild. Zu einem großen Teil hängt dies mit der außerordentlichen Vielfalt an Erhebungsinstrumenten zusammen (Johnson 1980). Dadurch ist es schwieriger als etwa bei demographischen oder klinischen Variablen, die Ergebnisse verschiedener Studien zu vergleichen. Eine umfassende quantitative Literaturübersicht zur Prädiktion der Lithiumresponse aufgrund der verschiedenen psychosozialen Faktoren wird in Kapitel 4.5 gegeben. Eine Zusammenfassung des aktuellen Wissensstands aufgrund der quantitativ auswertbaren Primärliteratur findet sich gegen Ende von Kapitel 4.6.

#### 2.4.2.4 Integrative Arbeiten zur Prädiktion der prophylaktischen Wirkung von Lithium

Seit den späten 70er und vor allem seit den 80er Jahren kamen zu diesen Einzelbefunden zusammenfassende Bewertungen der Forschungslage und die Zusammenfassung einzelner Variablen in komplexen Prognosemodellen hinzu. Insgesamt überwiegt eine zurückhaltende Bewertung der Möglichkeit, den Erfolg einer Rezidivprophylaxe vorauszusagen (z.B. Sutter 1983, Brochier & Gay 1991). Einige Autoren schließen sogar, dass es „offensichtlich keine Prädiktoren für den möglichen Therapieerfolg ... einer Lithiumprophylaxe gibt“ (Mander 1986, S. 40). Diese zurückhaltenden und pessimistischen Einschätzungen sind allerdings nicht unwidersprochen geblieben. Eine Reihe von Autoren vertreten den Standpunkt, dass die Vielzahl negativer und nicht replizierbarer Ergebnisse wesentlich durch unzureichende

Methodik bei der Durchführung und Auswertung der Studien mitbedingt sei, und daß „das Ansprechen auf eine Lithiumprophylaxe voraussagbar ist, sofern eine geeignete Methodik verwendet wird“ (Grof et al. 1993, S.16). Sie argumentieren, dass nur ein multivariates Modell der komplexen Natur psychiatrischer Erkrankungen gerecht werden kann (Grof et al. 1983a, Goodwin & Jamison 1990, Deister & Marneros 1993). Diese Befunde werden gestützt durch die empirischen Studien von Grof et al. (1983a,b), Kusalic & Engelsmann (1998) und Kato (1999). In allen drei Studien ergaben sich auf der Grundlage multivariater Vorhersagemodelle sehr hohe Übereinstimmungsraten zwischen prognostiziertem und tatsächlichem Verlauf unter einer Lithiumprophylaxe. Einschränkend ist jedoch anzumerken, daß sich bei einer kritischen Beurteilung dieser Studien deutliche Unterschiede hinsichtlich der Methodik zeigen. So sind die hohen Trefferquoten von 78.9% (Kusalic & Engelsmann, 1998) und 93.5% (Kato 1999) sehr zurückhaltend zu interpretieren, da in beiden Studien keine Kreuzvalidierung durchgeführt wurde. Dieser methodische Mangel wiegt umso schwerer als beide Studien nur jeweils ca. 30 Patienten umfassen und in beiden Fällen eine Vielzahl unabhängiger Variablen in die Modellgleichungen eingeschlossen wurde. In dieser Situation wirkt sich das Fehlen einer Kreuzvalidierung besonders stark aus und führt in der Regel zu einer sehr hohen Trefferquote – auch, falls die unabhängigen Variablen keine oder nur geringe prognostische Kraft besitzen. Diese Gefahr eines statistischen Artefakts beim Einsatz multivariater Verfahren ohne Kreuzvalidierung bleibt in der aktuellen methodischen Diskussion weitgehend unbeachtet. Dieser Punkt wird daher im empirischen Teil dieser Arbeit nochmals aufgegriffen, untersucht und diskutiert (vgl. Kap. 6.3.3.5). Im Gegensatz zu den Studien von Kusalic & Engelsmann (1998) und Kato (1999) ergab sich in der Studie von Grof et al. (1983a,b) ein hoher Anteil korrekt prognostizierter Verläufe (83%) bei einer unabhängigen Stichprobe. Leider wird aufgrund der Originalarbeiten nicht klar, welches Verfahren zur Prognose benutzt wurde (auch nicht aufgrund der ausführlicheren Darstellung in Grof et al. 1983c). Immerhin wird deutlich, daß sich die Prognose im wesentlichen auf drei unabhängige Variablen stützt: 1. die Diagnose einer primären affektiven Störung, 2. ein hohes Ausmaß an Remissionsqualität, 3. eine geringe Phasenfrequenz in der Vorgeschichte. Aufgrund dieser Ergebnisse wird die bereits in früheren Studien berichtete Bedeutung einer Diagnose einer primären affektiven Störung für die Lithiumresponse noch einmal bestätigt. Dieses offenbar wichtige Ergebnis ist in der Diskussion der letzten zwei Jahrzehnte in den Hintergrund getreten. Dies hängt vermutlich damit zusammen, daß seit der Verwendung einer strikteren Diagnostik (insbesondere seit Verwendung des 1980 erschienenen DSM-III) fast ausschließlich Patienten mit primären affektiven Störungen in Studien zur Lithiumresponse eingeschlossen werden. Dieser Entwicklung folgend, bewegt sich auch die vorliegende Arbeit ausschließlich innerhalb des Rahmens primärer affektiver Störungen. Auf die prädiktive Kraft der beiden verbleibenden Variablen aus den multivariaten Analysen von Grof et al. (1983a,b) und zu deren Bedeutung aus heutiger Sicht wird in den Kapiteln 4.4.2.12, 4.4.2.5 und 7 nochmals eingegangen.



Als Zwischenfazit lässt sich festhalten, daß auch mit den in früheren Publikationen *prima vista* außerordentlich erfolgreichen multivariaten Prognosemodellen noch kein Durchbruch bei der Prognose der Lithiumresponse gelungen ist. Derzeit muss weiter als offen gelten, ob die Prognose durch den Einsatz komplexer statistischer Verfahren wesentlich verbessert werden kann. Bedenken bestehen hierbei auch aufgrund modelltheoretischer Überlegungen. So weisen Preisig & Angst (1991) darauf hin, dass der Einsatz multivariater Prognosemodelle eigentlich nur dann erfolgversprechend wäre, wenn die kausalen Zusammenhänge im Wesentlichen bekannt wären – dies ist aufgrund des unzureichenden Kenntnisstandes derzeit kritisch zu sehen.

Die insgesamt kontroverse Forschungslage zur Prädiktion der Lithiumresponse ist offenbar auch durch eine sehr selektive Rezeption der Primärstudien bedingt. Zwar liegt eine Vielzahl von Übersichtsarbeiten zu Prädiktion der Lithiumresponse aufgrund klinischer, demographischer oder psychosozialer Variablen vor (insbesondere Carroll 1979, Abou-Saleh & Copen 1986, Brochier & Gay 1991, Grof et al. 1993, Calabrese et al. 1996, Walden et al. 1999, Ketter & Wang 2002), doch bei kritischem Lesen ergibt sich meist der Eindruck, daß eine große Zahl von Primärarbeiten nicht rezipiert wurde. Dieses Phänomen ist im Allgemeinen gut bekannt und keineswegs spezifisch für die Literaturlage zur Prädiktion der Lithiumresponse (Mulrow 1987, Antman et al. 1992, Mulrow 1995). Auch, um einer sich hieraus möglicherweise ergebenden Verzerrung entgegenzuwirken, wurden in den letzten Jahren vermehrt systematische Übersichtsarbeiten zur Prädiktion der Lithiumresponse aufgrund einzelner Variablen durchgeführt (Faedda et al. 1991, Viguera et al. 2000, Baethge et al. 2003, Tondo et al. 2003). Im Rahmen dieser Arbeit sind insbesondere die auf Lithiumresponse fokussierten Arbeiten von Faedda et al. (1991) zum Verlaufsmuster in der Vorgeschichte und von Viguera et al. (2000) zur Prädiktion aufgrund des Geschlechts von Interesse. Die Arbeit von Faedda et al. (1991) ist eine sorgfältige und hohen methodischen Standards genügende Metaanalyse von 5 Primärstudien, aufgrund der sich eindeutig ergibt, daß das „klassische“ Verlaufsmuster Manie – Depression – freies Intervall (MDI) mit einer günstigen Prognose für eine Lithiumprophylaxe assoziiert ist. Offenbar wurden alle relevanten Primärstudien berücksichtigt und im Abstand von nunmehr 12 Jahren ist dem nur wenig hinzuzufügen. Sinnvoll erscheint es allenfalls, diese hervorragende Analyse durch zwei neuere Primärstudien zu aktualisieren und somit die Effektstärkeschätzung auf eine breitere Basis zu stellen. Die Metaanalyse von Viguera et al. (2000) zur Prädiktion der Lithiumresponse aufgrund des Geschlechts weist hingegen nicht die gleiche hohe Qualität auf. Zum einen wurde aufgrund eines sehr schematischen Vorgehens bei der Lokalisation der Originalarbeiten etwa die Hälfte der Primärstudien übersehen. Zum anderen hätte es die große Anzahl von Primärstudien erlaubt, eine homogenere Auswertung vorzunehmen. Insbesondere wäre eine Unterscheidung nach akuttherapeutischem und prophylaktischem Ansprechen wünschenswert

gewesen. Hierzu ist andererseits zu bemerken, daß der Zusammenhang zwischen akuttherapeutischer und prophylaktischer Response wenig untersucht und weitgehend unklar ist (Dick 1983). Anzumerken ist auch, daß die Primärarbeiten, in denen die Lithiumresponse mit dem Geschlecht in Verbindung gebracht werden kann, in der Regel nicht als solche gekennzeichnet und damit nur schwer lokalisierbar sind. Ein Beispiel hierfür ist die Arbeit von Upadhyaya et al. (1985). Diese Arbeit bezieht sich primär auf biochemische Korrelate der Lithiumresponse, enthält aber eine Graphik, aufgrund der eine 4-Felder-Tafel zum Zusammenhang von Geschlecht und Response erstellt werden kann. Insgesamt ist die Lokalisation von Primärstudien zur Prädiktion der Lithiumresponse ungleich schwieriger als etwa bei Therapievergleichsstudien – die ja in aller Regel als solche gekennzeichnet sind. Es wäre wünschenswert all diese systematisch nur schwer erfassbaren Befunde zur Responseprädiktion einer breiteren Rezeption zugänglich zu machen. Hierzu wäre es notwendig, eine möglichst breite und umfassende Recherche nach empirischen Studien durchzuführen, für welche die Möglichkeit besteht, daß sich die Lithiumresponse mit möglichen Prädiktoren in Zusammenhang bringen lässt.

Insgesamt lässt sich der aktuelle Forschungsstand zur Responseprädiktion wie folgt zusammenfassen:

- Es liegt eine Vielzahl von Einzelbefunden zur Charakterisierung von Lithiumrespondern vor.
- Viele dieser Einzelbefunde sind systematisch nur schwer zugänglich, werden kaum rezipiert oder sind weitgehend in Vergessenheit geraten.
- Nur für einige wenige Variablen liegen gesicherte Erkenntnisse über deren Prädiktionskraft vor. Insbesondere ist unklar, welche Variablen aufgrund der aktuellen Forschungslage als Prädiktoren anzusehen sind und welche nicht.
- Der Therapieerfolg einer Lithiumprophylaxe lässt sich mit Hilfe einzelner Variablen nur sehr begrenzt voraussagen.
- Ob der Einsatz multivariater Modelle zu einer Verbesserung der Prognose führt, ist unklar.

### 3 Fragestellung

Für die vorliegende Arbeit sind zwei Leitfragen zentral:

1. Welche klinischen, demographischen und psychosozialen Variablen können bei einer systematischen Literaturübersicht derzeit als Responseprädiktoren für eine Lithiumprophylaxe gelten?
2. Lässt sich aus diesen Responseprädiktoren ein im klinischen Einsatz hilfreiches Prognoseinstrument konstruieren?

Zur Beantwortung der ersten Leitfrage werden die folgenden in Kapitel 4 dargestellten Teilschritte durchgeführt:

- Zunächst wird versucht, alle Studien zu lokalisieren, die eine quantitative Aussage über den Zusammenhang zwischen der Response auf eine Lithiumprophylaxe und klinischen, demographischen oder psychosozialen Variablen erlauben.
- Für jede mögliche Prädiktorvariable werden die Ergebnisse aller Primärstudien im Rahmen einer Metaanalyse quantitativ aggregiert.
- Das Ziel ist, für jede bisher untersuchte Prädiktorvariable eine Effektstärkeschätzung ihrer prädiktiven Kraft anzugeben. Die Punktschätzung der Effektstärke wird durch die Angabe eines Konfidenzintervalls ergänzt.

Die Beantwortung der zweiten Leitfrage beinhaltet im wesentlichen die beiden folgenden (in Kapitel 6 ausgeführten) Teilschritte:

- Die Beschreibung geeigneter Strategien zur Responseprädiktion auf Grundlage der durchgeführten Metaanalysen.
- Die Evaluation dieser Strategien aufgrund einer umfangreichen empirischen Studie.

Die thematische Einschränkung der Arbeit auf klinische, demographische und psychosoziale Variablen ist wesentlich durch die Datenlage der empirischen Studie bedingt, auf deren Grundlage die Evaluation der Prognosestrategien erfolgt. In dieser Studie wurde eine Vielzahl dieser Variablen gezielt erhoben. Zu biologischen und genetischen Variablen liegen in dieser Studie hingegen nur wenige Daten vor. Es wäre in jedem Fall wünschenswert, die vorliegende Arbeit durch eine Analyse biologischer und genetischer Prädiktoren zu ergänzen. Beim gegenwärtigen Forschungsstand, und mehr noch – bei möglicherweise weiteren Fortschritten insbesondere der molekulargenetischen Charakterisierung von Respondern in den nächsten Jahren – wäre die Einbeziehung aller bekannten Prädiktoren zur individuellen Prognosestellung vermutlich mit einer genaueren Prognose verbunden, als es derzeit möglich ist. Die

vorliegende Arbeit versteht sich als Beitrag, manisch-depressiven Patienten möglichst ohne Umwege und Rückschläge die individuell optimale Behandlung zukommen zu lassen.

## 4 Metaanalysen zur Prädiktion des Erfolgs einer Lithiumprophylaxe

### 4.1 Methodik

#### 4.1.1 Lokalisation der Studien

Basis der Suche von Primärstudien, in denen die Lithiumresponse mit Kovariaten in Zusammenhang gebracht werden, war eine Literaturrecherche (Medline, Embase, Psychlit). Um das Risiko für den Ausschluss relevanter Artikel gering zu halten, wurden bei der Rechner-gestützten Literatursuche sehr weite Auswahlkriterien gewählt (zur zentralen Bedeutung einer möglichst vollständigen Erfassung aller Primärstudien für die Validität einer Metaanalyse vgl. etwa Dickersin et al. 1995). Durchgeführt wurde eine Volltextsuche mit folgenden Suchbegriffen: (lithium and response) and (affective or depress\* or mani\* or bipolar) and (human in tg). Die letzte Klammer dient dem Ausschluss tierexperimenteller Studien, die Asterisken am Ende der Worte (z.B. depress\*) dienen dazu, sowohl Studien mit Begriffen wie "depression" als auch "depressive" einzuschließen. Berücksichtigt wurden Arbeiten aus dem Zeitraum von 1966 – 3/2003. Da die erste systematische Studie zur prophylaktischen Wirksamkeit von Lithium 1967 publiziert wurde (Baastrup & Schou 1967), wurden mit diesem zeitlichen Fenster keine relevanten Arbeiten ausgeschlossen. Die maschinelle Suche von Kandidatenstudien wurde durch eine ausführliche Handsuche ergänzt. Hierbei wurden insbesondere Übersichtsarbeiten zur Wirksamkeit von Lithium und – soweit verfügbar – Kongressbände (bzw. CDs) berücksichtigt. Alle Literaturverzeichnisse der Übersichts- und Originalarbeiten zur Responseprädiktion wurden auf möglicherweise relevante Arbeiten durchgesehen. Um die Metaanalyse auf eine möglichst breite empirische Basis zu stellen und um einem (letztlich unvermeidbaren) Publikationsbias entgegenzuwirken, wurde eine Vielzahl von Studienleitern und Forschern im Bereich bipolarer Störungen angeschrieben und nach nicht publizierten Studien befragt. Insgesamt wurden Arbeiten in englischer, deutscher, französischer, niederländischer, italienischer und spanischer Sprache berücksichtigt. Eine dem Titel nach vielversprechende Arbeit in russischer Sprache (Minsker et al. 1977) wurde aufgrund fehlender Russischkenntnisse mit Unterstützung von Frau Giegling erarbeitet, erfüllte aber schließlich nicht die endgültigen Einschlusskriterien (die Auswertung lies keine separaten Analysen für bipolare Patienten zu).

In einem zweiten Schritt wurden die Arbeiten gesichtet, und in die Metaanalyse eingeschlossen, falls sie die folgenden Kriterien erfüllten:

1. Mindestbeobachtungsdauer von 6 Monaten
2. Separat ausgewiesene Ergebnisse für bipolar affektive Störungen
3. Separat ausgewiesene Ergebnisse für Patienten mit Lithiumtherapie

4. Quantifizierbarer Zusammenhang zwischen Lithiumresponse und mindestens einem klinischen, demographischen oder psychosozialen Faktor
5. Keine Mehrfachpublikation

Zu 1) Die Metaanalyse zielt primär auf Prädiktion der prophylaktischen Wirksamkeit von Lithium. Allerdings ist eine sichere Unterscheidung zwischen Erhaltungstherapie und Rezidivprophylaxe nicht möglich (2.2.1) und daher wird – wie allgemein üblich – durch das 6-Monats-Kriterium eine rein operationale Abgrenzung getroffen.

Zu 2) Durch dieses Kriterium wird die diagnostische und therapeutische Trennung zwischen unipolarer Depression und bipolarer Störung vollzogen, die sich im Anschluss an die Arbeiten von Angst (1966), Perris (1966) und Winokur & Clayton (1967) durchgesetzt hat (vgl. Kap. 2.1). Da diese Trennung mit der Einführung von Lithium als Phasenprophylaktikum parallel lief, gehen durch dieses Kriterium leider mehrere Arbeiten für die Metaanalyse verloren. Auf der anderen Seite wird dadurch eine größere Präzision gewonnen und eine größere Nähe mit der aktuellen therapeutischen Praxis erreicht. Des weiteren wird durch die Beschränkung auf die Gruppe der bipolaren Patienten eine klarere Interpretation der Prädiktorvariablen erreicht, die mit der Diagnose assoziiert sind. So ist in gemischten Stichproben der Einfluss von Variablen wie Ersterkrankungsalter oder Geschlecht (unipolare Patienten erkranken in der Regel später und sind häufiger weiblich) nur schwer interpretierbar, weil nicht klar ist, ob der Effekt direkt mit dem Erkrankungsalter bzw. Geschlecht zusammenhängt, oder ob vielleicht doch die Diagnose der entscheidende Faktor ist.

Im Gegensatz zu Stichproben mit Beteiligung von unipolaren Patienten war der Ausschluss von Stichproben mit einem gewissen Anteil von schizoaffektiven Patienten nicht immer praktikabel. Zwar wurden Kollektive mit primär schizoaffektiven Patienten ausgeschlossen, aber die Abgrenzung zwischen schizoaffektiver Störung und bipolarer Störung ist so stark vom diagnostischen System (insbes. ICD vs. DSM) und der Auflage des diagnostischen Systems abhängig, daß eine strikte Beschränkung auf bipolar affektive Stichproben kaum durchführbar ist und an dieser Stelle eine gewisse Unschärfe in Kauf genommen werden musste.

Leider war nicht in jedem Fall nachvollziehbar, ob anerkannte diagnostische Kriterien zur Diagnostik (insbesondere ICD, DSM, RDC, Feighner-Kriterien) verwendet wurden. Nach heutigem Standard wäre dies ein gravierender methodischer Mangel und spräche gegen eine professionell durchgeführte Studie. In den 60er und 70er Jahren war die Forschungspraxis eine andere und daher wurden auch Studien ohne klar definierte operationalisierte Diagnostik aufgenommen.

Zu 3) Aus der aktuellen Forschung zeichnet sich ab, daß Lithium nur in einer Subgruppe bipolarer Patienten hochwirksam ist, während andere Phasenprophylaktika ein möglicherweise breiteres Wirkspektrum aufweisen. Eine klarere Konturierung der Gruppe der Lithium-Responder wäre in der aktuellen Diskussion um einen differenzierten Einsatz verschiedener Mood-Stabilizer von Bedeutung. Eine Beschränkung der Metaanalyse auf Patienten mit einer strikten Monotherapie ist hingegen kaum praktikabel, da Polypharmazie in der Behandlung bipolarer Störungen die Regel (vgl. etwa Frye et al. 2000) und zudem in vielen Studien nur unzureichend dokumentiert ist. Ausgeschlossen wurden allerdings Patienten, von denen bekannt war, daß sie mit einem weiteren Mood-Stabilizer behandelt wurden (Valproat, Carbamazepin, Lamotrigin, Topiramid, Gabapentin, Olanzapin, Clozapin).

Zu 4) Prinzipiell wurden alle Studien berücksichtigt, die mindestens eine klinische, demographische oder psychosoziale Variable, mit der prophylaktischen Wirksamkeit von Lithium in Zusammenhang bringen. Nicht berücksichtigt wurden Therapie- und Response-assoziierte Parameter wie Dosierung, initiales Ansprechen auf die Therapie, sowie biologische und genetische Parameter. Die Ergebnisdokumentation musste in irgendeiner Form eine Quantifizierung dieses Zusammenhangs erlauben. Da bei extrem ungleichen Zellbesetzungen in 4-Felder Tabellen das für die Metaanalysen verwendete Effektstärkemaß  $r$  gegen 0 tendiert (auch wenn tatsächlich ein kausaler Zusammenhang vorhanden ist), wurde in diesen 2x2-Kreuztabellen gefordert, daß jede Zelle mit mindestens 5% der gültigen Fälle besetzt ist. Dieses Kriterium dient der Vermeidung eines Bias, der durch die Verwendung wenig sensitiver Messinstrumente entstehen könnte. Bei möglichen Prädiktoren sehr hoher oder sehr geringer Inzidenz wäre ein anderes Vorgehen (allerdings auch mit einem anderen Effektstärkemaß) indiziert.

Zu 5) Bei Überschneidungen zwischen den Patientenpopulationen verschiedener Studien wurde die Studie mit der größten Fallzahl in die Metaanalyse eingeschlossen und die anderen Studien ausgeschlossen.

#### **4.1.2 Charakterisierung und Auswertung der Studien**

##### 4.1.2.1 Charakterisierung und Rating der Studien

Die Studien wurden hinsichtlich der für die weitere Interpretation wesentlichen Hauptmerkmale charakterisiert und tabellarisch zusammengefasst. Obwohl offensichtliche Qualitätsunterschiede zwischen den Studien vorlagen (z.B. Fall-Kontroll-Studien vs. Kohortenstudien) wurde auf ein Qualitätsrating der verschiedenen Studien verzichtet. Die üblicherweise verwendeten Ratingsysteme (bspw. Dardennes et al. 1995) erscheinen nur sehr

bedingt geeignet zumal diese Ratingsysteme sich in der Regel primär für die Evaluierung von Therapievergleichsstudien eignen (z.B. ist nicht erkennbar, warum eine Studie mit formaler Power-Analyse im Zusammenhang dieser Arbeit stärker gewichtet werden sollte). Insgesamt ist auch damit zu rechnen, daß einige methodische Mängel der Studien (z.B. fehlende Verblindung) einen geringeren Einfluss auf die Ergebnisse haben als etwa bei einer Therapievergleichsstudie und so wurde aufgrund der außerordentlich großen Schwierigkeiten einer fundierten Quantifizierung von Qualitätsunterschieden hierauf verzichtet. An methodischen Unterschieden zwischen den Studien fällt im Zusammenhang der vorliegenden Metaanalyse vor allem die Art der outcome-Messung ins Gewicht. Ein einheitliches outcome-Kriterium wie "affektives Rezidiv in der Beobachtungszeit" würde zu klareren und homogeneren Ergebnissen führen. Die Beschränkung auf eine Klasse von outcome-Kriterien hätte allerdings zum Ausschluss etwa der Hälfte der Studien geführt und daher wurden alle Studien mit klar definierten und quantifizierbaren outcome-Kriterien eingeschlossen. Die Mischung verschiedener outcome-Kriterien stellt zweifellos einen gewissen methodischen Mangel der vorliegenden Metaanalysen bei der Aggregation und zusammenfassenden Interpretation der Primärstudien dar. Das Problem der Heterogenität der Studien wird verstärkt durch unterschiedliche Beobachtungsdauern, unterschiedliche Messinstrumente bei der Erfassung mancher Variablen und dem langen Zeitraum von über 35 Jahren in dem die lokalisierten Studien publiziert wurden. Aufgrund dieser Überlegungen wurde von vornherein mit starker Variabilität in den Effektstärken gerechnet und bei der Aggregation der Daten über die Studien ein entsprechender metaanalytischer Ansatz gewählt ("random-effects model", vgl. Kap. 4.1.4).

Der in Metaanalysen von Therapievergleichsstudien wichtige Aspekt unterschiedlicher Behandlung von drop-outs in den Originalarbeiten fällt im Rahmen dieser Studie hingegen nicht ins Gewicht. Zwar wurden in den Primärstudien verschiedene Methoden zur Behandlung unvollständiger Beobachtungen (z.B. Survival-Analysen vs. ATP-Analysen) verwendet. Spezielle Verfahren zur Behandlung von drop-outs beziehen sich jedoch nie auf die Analyse des Einflusses klinischer, demographischer oder psychosozialer Variablen, so daß sämtliche in der vorliegenden Arbeit durchgeführten Analysen zur Prädiktion der Lithiumresponse auf Primäranalysen according to protocol (ATP) beruhen. Das generelle Problem der unvollständigen und fehlenden Beobachtungen bleibt natürlich bestehen – es pflanzt sich von den Primärstudien auf die Ebene der Metaanalyse fort. Auf dieser Ebene scheint es allerdings nicht mehr möglich, dieses Problem zu entschärfen.

Auch methodische Unterschiede bei der statistischen Auswertung in den Originalarbeiten spielen für die Ergebnisse in der vorliegenden Metaanalyse eine nur untergeordnete Rolle, weil die Berechnung der Effektstärken aus den Primärstudien meist auf deskriptiven Maßen beruhte.



#### 4.1.2.2 Zusammenfassung der Studien, Effektstärkemaße

In der medizinischen und psychologischen Literatur werden aktuell mehrere Effektstärkemaße (ES) zur Messung der Stärke eines Zusammenhangs parallel verwendet (vgl. etwa Laupacis et al. 1988, Rosenthal 1991). Dies begründet sich zum Teil aus Forschungstraditionen, zum Teil aus sachlichen Gründen. So eignen sich in der Grundlagenforschung eher an der Kausalität orientierte ES-Maße wie das "relative Risiko", im klinischen Alltag hingegen eher ES-Maße, in denen im weitesten Sinn ökonomische Aspekte eine Rolle spielen wie die "number needed to treat" (zur Definition und Diskussion verschiedener ES-Maße vgl. Rosenthal 1991, Sinclair & Bracken 1994). In den meisten Kontexten (z.B. im Kontext dieser Arbeit) sind beide Aspekte von Bedeutung und würden daher ein zweidimensionales ES-Maß erfordern. Dieses Vorgehen wäre jedoch zum einen unüblich und würde zum anderen weit höhere Anforderungen an die Ergebnisdarstellung der Primärstudien stellen. Daher wurde traditionell ein eindimensionales ES-Maß ausgewählt und in dieser Metaanalyse verwendet, der als Effektstärkemaß interpretierte Korrelationskoeffizient  $r$ . Dieses Maß hat unter anderem folgende Vorzüge:

- 1) Hohe Vertrautheit bei Forschern und Anwendern
- 2) Anschaulichkeit und generell gute Interpretierbarkeit
- 3) "Genügsamkeit", d.h. es sind nur wenige Angaben aus den Primärstudien nötig
- 4) "Flexibilität", d.h. aus fast jeder quantitative Ergebnisdarstellung kann ein  $r$  berechnet werden

Zu 1) Da das Effektstärkemaß  $r$  exakt dem üblichen Korrelationskoeffizienten nach Bravais-Pearson entspricht, besitzt es nicht nur bei Forschern, sondern auch bei Lesern und Anwendern von Metaanalysen ein hohes Maß an Vertrautheit. Die Wichtigkeit dieses Punktes wird auch durch empirische Untersuchungen zur Interpretation von Effektstärkemaßen in der medizinischen Anwendung belegt. Demnach werden Arbeiten, die weniger vertraute Effektstärkemaße wie etwa die odds-ratio verwenden, sehr häufig missverstanden und falsch interpretiert (vgl. Sinclair & Bracken 1994).

Zu 2) Das Effektstärkemaß kann auf verschiedene Weise interpretiert und veranschaulicht werden. Neben der Interpretation als Korrelation weist Rosenthal (1991) auf eine der klinischen Denkweise sehr entgegenkommenden Veranschaulichung hin:  $r$  kann interpretiert werden als eine Verbesserung der Erfolgsaussichten von  $50\% - \frac{r}{2}$  auf  $50\% + \frac{r}{2}$ . Demnach entspricht beispielsweise ein  $r$  von 0.14 einer Verbesserung der Erfolgsaussichten von 43% auf 57%. Das  $r$  von  $0.14 = 14\%$  entspricht demnach gerade der Verbesserung der

Erfolgsaussichten. Dieser Zusammenhang gilt exakt, falls die Erfolgsaussichten symmetrisch um die 50% liegen, andernfalls näherungsweise.

Zu 3-4) Daß nur sehr wenige Angaben aus den Primärstudien genügen und quantitative Angaben aus den Primärstudien fast immer in die Effektstärke  $r$  übersetzt werden können, wird im folgenden Abschnitt (4.1.3.) ausführlich illustriert. Dieser Punkt ist ein großer Vorzug gegenüber Effektstärkemaßen, die (wie etwa das relative Risiko) an eine speziellere Ergebnisdarstellung in den Primärstudien gebunden sind, und daher einen höheren Anteil von Studien aufgrund fehlender Angaben von der Analyse ausschließen.

### 4.1.3 Berechnung des Effektstärkemaßes $r$ aus den Primärstudien

#### 4.1.3.1 Korrelationskoeffizienten in den Primärstudien

Falls in der Primärstudie der Zusammenhang zwischen Prädiktorvariable und Response als Korrelationskoeffizient nach Bravais-Pearson oder nach Spearman, oder eine Punkt-biseriale Korrelation angegeben war, wurde dieser Korrelationskoeffizient ohne Umrechnung als Effektstärkemaß  $r$  interpretiert. Weitere Korrelationskoeffizienten (z.B. tetrachorische Korrelation, Kendall's  $\tau$ ) kamen in den hier ausgewerteten Primärstudien nicht vor.

#### 4.1.3.2 4-Felder-Tafeln in den Primärstudien

Aus jeder 4-Felder-Tafel lässt sich ein  $\chi^2$ -Wert berechnen, der wie üblich in das ES-Maß  $r$  überführt wurde:

$$r = \pm \sqrt{\frac{\chi^2}{n}} \quad (\text{F1})$$

Diese Definition entspricht exakt der Definition des  $\phi$ -Koeffizienten. Das Vorzeichen ergibt sich aus der Art der Kodierung. Bei der Kodierung wurde, wie in der gesamten Arbeit, darauf geachtet, daß ein positives  $r$  für einen positiven Prädiktor steht und ein negatives  $r$  für einen Risikofaktor.

#### 4.1.3.3 Kontingenztafeln in den Primärstudien

Um eine klarere Interpretation für die einzelnen Merkmalsausprägungen zu erhalten, und um die Richtung des Zusammenhangs angeben zu können, wurden größere Kontingenztafeln (mit  $k$  Kategorien,  $k > 2$ ) in  $k-1$  4-Felder-Tafeln unterteilt. Z.B. wurde Familienstand nicht global

beurteilt, sondern es wurden neue, dichotome Variablen definiert (z.B. ledig vs. nicht ledig) und in 4-Felder-Tafeln beurteilt. Für diese präziser gefassten Variablen wurde die ES wieder nach der oben gegebenen Formel (F1) berechnet.

#### 4.1.3.4 t-Tests, F-Tests, Mittelwerte und Standardabweichungen in den Primärstudien

Aus den Fallzahlen pro Gruppe ( $n_1$ ,  $n_2$ ), sowie den Mittelwerten ( $M_1$ ,  $M_2$ ) und Standardabweichungen ( $S_1$ ,  $S_2$ ) lässt sich der übliche t-Wert berechnen:

$$t = \frac{M_1 - M_2}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (\text{F2})$$

Dieser t-Wert kann durch Umkehrung der bei der inferenzstatistischen Testung von r gängigen t-Verteilungs-Approximation in das ES-Maß r umgerechnet werden:

$$r = \pm \sqrt{\frac{t^2}{t^2 + (n_1 + n_2 - 2)}} \quad (\text{F3})$$

Das Vorzeichen ergibt sich aus der Kodierung und dem Vorzeichen des t-Wertes.

F-Werte lassen sich über den Zusammenhang zwischen F- und t-Verteilung in t-Werte übersetzen:

$$t = \pm \sqrt{F} \quad (\text{F4})$$

und anschließend über Formel (F3) in das ES-Maß r überführen.

#### 4.1.3.5 Regressionsanalysen in den Primärstudien

Die Regressionskoeffizienten  $b$  bei der Regression von Y auf X stehen über die Standardabweichungen  $S_x$  und  $S_y$  in direktem Zusammenhang mit dem Korrelationskoeffizienten nach Bravais-Pearson und damit mit dem ES-Maß r:

$$r = b \frac{S_x}{S_y} \quad (\text{F5})$$

#### 4.1.3.6 Sonstige Tests oder unzureichende Angaben zur Statistik in den Primärstudien

In allen anderen Fällen, in denen jedoch neben der Fallzahl noch ein p-Wert angegeben wurde, wurde der p-Wert über die t-Verteilung in einen t-Wert übersetzt und nach Formel (F3) in das ES-Maß  $r$  übertragen. Das Vorzeichen ergab sich hier aus der in der Originalarbeit berichteten Richtung des Zusammenhangs. In einem Fall wurden anstatt von p-Werten Konfidenzintervalle berichtet. Hierbei wurden aus dem allgemeinen Zusammenhang zwischen Intervallschätzung zum Vertrauensgrad  $\gamma = 1 - \alpha$  und Hypothesentestung zum Signifikanzniveau  $\alpha$  der zweiseitige p-Wert berechnet, in den entsprechenden t-Wert übertragen und dann wieder über die Formel (F3) in das ES-Maß  $r$  überführt. An dieser Stelle sei noch einmal betont, daß in der gesamten Arbeit die Kodierung der Variablen so gewählt wurde, daß ein positives  $r$  für eine positive Prädiktion steht (positive Korrelation zwischen Prädiktorvariable und outcome) und ein negatives  $r$  für einen Risikofaktor.

#### 4.1.4 Aggregation der Effektstärken

Die ES-Maße wurden zunächst für jede Variable separat für alle Studien  $i = 1, \dots, m$  berechnet und dann für jede Variable über alle Studien zu einem gemittelten ES-Maß  $\bar{r}$  aggregiert. Die arithmetische Mittelung erfolgte aus inferenzstatistischen Überlegungen grundsätzlich auf Ebene der z-transformierten Werte von  $r$ :

$$z(r_i) = \ln \sqrt{\frac{(1 + r_i)}{(1 - r_i)}} \quad (\text{F6})$$

Da aufgrund der Heterogenität der Studiencharakteristika auch eine starke Heterogenität der ES-Maße zu erwarten war, wurde die Aggregation grundsätzlich nach dem Modell zufälliger Effekte ("random-effects model", DerSimonian & Laird 1986) durchgeführt. Bei diesem Ansatz geht in die relative Gewichtung der Studien neben der Fallzahl noch ein Korrekturfaktor ein, durch den "Ausreißerstudien" weniger stark gewichtet werden, da ein Teil des Ergebnisses individuellen Studiencharakteristika zugerechnet wird. Dieser Korrekturfaktor verschwindet bei relativ homogenen ES-Maßen und wurde daher generell für alle Variablen verwendet. In jedem Falle wurde jedoch eine formelle Prüfung der Heterogenität mittels der in Formel (F9) angegebenen  $\chi^2$ -Statistik (nach DerSimonian & Laird 1986, S. 182) durchgeführt und – falls kein Anhalt für eine überzufällige Heterogenität vorlag – zusätzlich eine Aggregation der Studien nach dem Modell mit festen Effekten durchgeführt (Gewichtung allein auf Basis der Fallzahlen). Die Unterschiede in den Ergebnissen zwischen festen und zufälligen Effekten waren in diesen Fällen jedoch so gering, daß im Sinne einer einheitlichen Darstellung in der gesamten Arbeit nur Ergebnisse berichtet werden, die auf dem

"random-effects model" beruhen. Bei diesem Ansatz erfolgt die Zusammenfassung der ES-Maße aus den einzelnen Studien aufgrund der nach DerSimonian & Laird (1986, S. 183) modifizierten Formel (F7).

$$\bar{z}(r) = \frac{\sum_{i=1}^m w_i^* z(r_i)}{\sum_{i=1}^m w_i^*} \quad (\text{F7})$$

$$\text{mit } w_i^* = \left( \frac{1}{n_i - 3} + \max \left( 0, \frac{Q_w - (k - 1)}{\sum_{i=1}^m (n_i - 3) - \frac{\sum_{i=1}^m (n_i - 3)^2}{\sum_{i=1}^m (n_i - 3)}} \right) \right)^{-1} \quad (\text{F8})$$

$$\text{mit } Q_w = \sum_{i=1}^m (n_i - 3) \left( z(r_i) - \frac{\sum_{i=1}^m (n_i - 3) z(r_i)}{\sum_{i=1}^m (n_i - 3)} \right)^2 \quad (\text{F9})$$

$Q_w$  ist approximativ  $\chi^2$ -verteilt mit  $m-1$  Freiheitsgraden und ist die übliche zur Prüfung der Heterogenität von ES-Maßen (etwa der ES-Maße  $z(r_i)$ ) verwendete Statistik.

$\bar{z}(r)$  ist approximativ normalverteilt und kann daher über die 95%-Konfidenzintervalle für normalverteilte Variablen approximativ auf Signifikanz geprüft werden. Um eine bessere Interpretierbarkeit des zusammenfassenden ES-Maßes zu erreichen, werden der gemittelte z-Wert  $\bar{z}(r)$  und das zugehörige Konfidenzintervall jedoch zunächst in ein gemitteltes ES-Maß  $\bar{r}$  mit Konfidenzintervall  $KI(\bar{r})$  zurücktransformiert. Dies geschieht mit Hilfe der Formeln (F10) und (F11).

$$\bar{r} = \tanh(\bar{z}(r)) \quad (\text{F10})$$

$$KI(\bar{r}) = \tanh \left( \frac{\bar{z}(r) \pm 1,96}{\sqrt{\sum_{i=1}^m (n_i - 3)}} \right) \quad (\text{F11})$$

Die Metaanalyse wurde in SAS (Version 8.02) realisiert. Da die notwendigen Programme nicht implementiert waren und selbst programmiert wurden mussten (in SAS-IML), wurden alle Programme extensiv anhand von Testdatensätzen auf mögliche Programmierfehler getestet. Aufgrund dieser Testung und aufgrund der vollständigen Übereinstimmung aller Ergebnisse bei der Kreuztestung mit einem von Prof. Engel in Excel programmierten Auswertungsprogramms können ergebnisrelevante Programmierfehler nahezu ausgeschlossen werden.

Multivariate metaanalytische Verfahren unter Einbeziehung potenzieller Moderatorvariablen wurden im Rahmen dieser Arbeit nicht durchgeführt. Dieses Verfahren ist vor allem aussichtsreich, falls entweder ausreichendes Vorwissen über den Einfluss von Moderatorvariablen vorhanden ist (hypothesengeleitetes Vorgehen) oder falls ausreichend Primärstudien vorhanden sind, die bezüglich möglicher Moderatorvariablen klassifizierbar sind (exploratives Vorgehen). Im aktuellen Forschungsstadium erschien es adäquater, zunächst empirische Belege zur Prädiktion der Wirksamkeit von Lithium zu lokalisieren und systematisch zusammenzufassen. Für bestimmte, relativ gut untersuchte potenzielle Responseprädiktoren (z.B. Phasenfrequenz) wäre in einem weiteren Schritt eine stärker inhaltlich fundierte metaanalytische Untersuchung, in der mögliche Moderatorvariablen (z.B. Art der outcome-Messung) berücksichtigt werden, von großer Relevanz für Forschung und klinische Praxis.

#### 4.1.5 Darstellung der Studienergebnisse – Forestplot und Treeplot

Für jede der untersuchten Variablen wurden die Ergebnisse aus den Primärstudien in einem sogenannten Forestplot (Schwarzer et al. 2000) dargestellt. Hierbei wird das 95%-Konfidenzintervall der Effektstärkeschätzung für jede der Primärstudien als Balken dargestellt (vgl. Abb. 5).

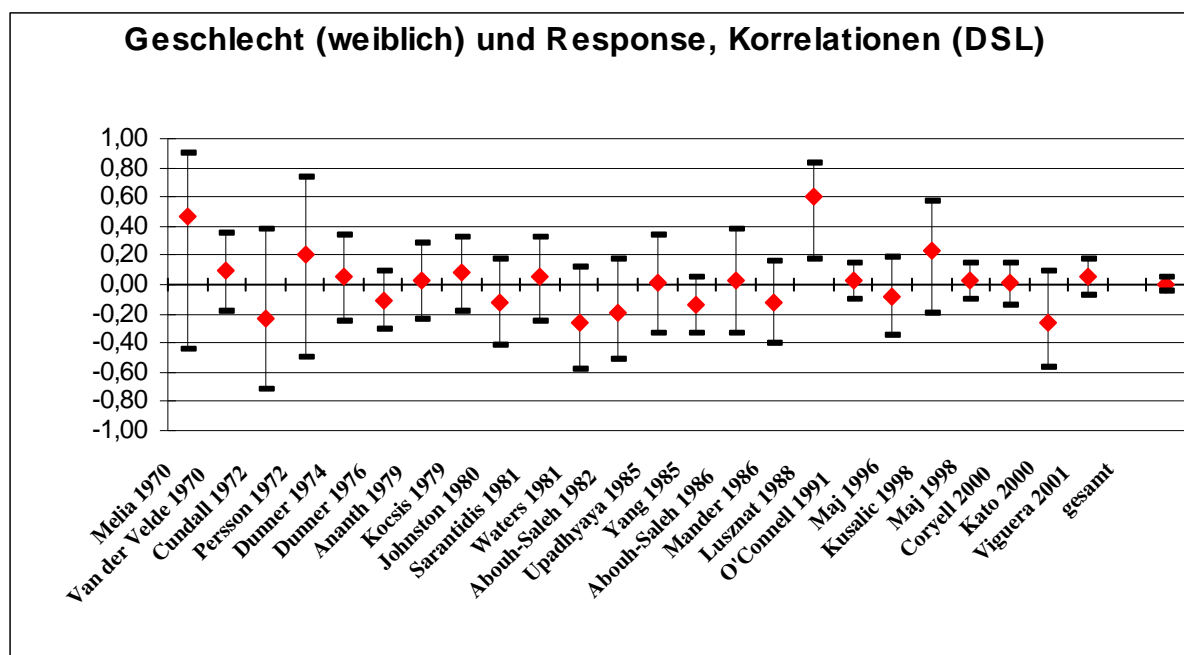


Abbildung 5. Beispiel eines Forestplots.

Die 95%-Intervalle aus den einzelnen Studien werden (nach Publikationsjahr sortiert) nebeneinander dargestellt. Zusätzlich findet sich am rechten Ende der Abbildung das 95%-Intervall der aus allen Studien aggregierten Effektstärkeschätzung. In der Regel ist dieses Konfidenzintervall der aggregierten Effektstärkeschätzung enger (präziser) als die Konfidenzintervalle aus den Einzelstudien. Aufgrund der Zusammenfassung der Einzelergebnisse nach dem Modell zufälliger Effekte (vgl. Kap. 4.1.4) kann es bei starker Inhomogenität in den Primärstudien in Ausnahmefällen allerdings auch zum gegenteiligen Effekt kommen - zu einem breiteren Konfidenzintervall für die aggregierte Schätzung kommen (vgl. etwa Kap. 4.4.1.2).

Falls für eine mögliche Prädiktorvariable 8 oder mehr Primärstudien vorlagen, wurde anhand eines Funnelplots untersucht, ob es Hinweise für einen Publikationsbias gibt. Bei der im Rahmen dieser Arbeit verwendeten Variante des Funnelplots werden dabei die Effektstärke (x-Achse) gegen die Fallzahl der Studien (y-Achse) graphisch dargestellt (vgl. z.B. Abb. 6). Falls ein Publikationsbias vorliegt, ist zu erwarten, daß signifikante Ergebnisse und

Ergebnisse aus großen Studien bevorzugt veröffentlicht werden (vgl. Schwarzer et al. 2000). Dies wird bei einer ausreichenden Anzahl publizierter Studien im Funnelplot durch eine Abweichung des empirischen vom erwarteten Funnelplot sichtbar. Unter der Nullhypothese, daß kein Publikationsbias vorliegt, ist im Funnelplot eine trichterförmige Verteilung der Effektstärken zu erwarten – die Verteilung der Effektstärken sollte im Idealfall einem auf der Mündung stehenden Trichter (engl. „funnel“) ähneln (zu den verteilungstheoretischen Hintergründen und zu einer eingehenderen Diskussion des Funnelplots vgl. etwa Schwarzer et al. 2000). Bei Vorliegen eines Publikationsbias wird eine Abweichung von dieser Trichterform erwartet, die sich im „Fehlen“ von Studien geringer Effektstärke und Fallzahl äußert. Kurz gesagt ist der Funnelplot also eine informelle, graphische Methode, um im Rahmen einer Metaanalyse einen Publikationsbias aufzuspüren. Einschränkend ist zu bemerken, daß im Rahmen dieser Arbeit dieses Instrument nur von eingeschränkter Aussagekraft ist. Zum einen wegen der relativ geringen Zahl von Primärstudien für die einzelnen Prädiktorvariablen, und zum anderen aufgrund der starken Heterogenität der Primärstudien. Durch die geringe Zahl von Primärstudien ist die Aussagekraft der Funnelplots generell eingeschränkt. Bei starker Heterogenität der Primärstudien ist auch unter der Nullhypothese („kein Publikationsbias“) mit Abweichungen von der Trichtergestalt zu rechnen und damit die Beurteilung des Funnelplots zusätzlich erschwert.



## 4.2 Eingeschlossene Studien

Insgesamt wurden durch die kombinierte Datenbank- und Handsuche ca. 2000 potenziell relevante Studien lokalisiert und auf Eignung für die Metaanalyse geprüft. Hiervon erfüllten 58 Studien alle Einschlusskriterien (ca. 3%). Dieser geringe Anteil eingeschlossener Primärstudien hängt vor allem mit der Verwendung sehr weiter Einschlusskriterien zusammen, die verwendet wurden, um eine möglichst breite und – soweit möglich – unverzerrte Datenbasis zu bekommen (vgl. Kap. 4.1.1). Diese 58 eingeschlossenen Primärstudien wurden tabellarisch bezüglich der wichtigsten Studienmerkmale charakterisiert (vgl. Tab. 5).

Tabelle 5: Kurzcharakteristik der eingeschlossenen Studien.

Autoren und Publikations-jahr	Diagnose	n <sup>1</sup>	Studien-dauer (Jahre)	Studien-typ <sup>2</sup>	Evaluations-kriterium <sup>3</sup>	Bemerkungen
Abouh-Saleh 1983	bipolar I (RDC)	57	unklar ("lithium prophylaxis")	FKS	affektives Rezidiv	
Abouh-Saleh & Coppen 1986	bipolar	33	1-2	FKS	Ø Morbidity Index	Untergruppenauswertung der 33 bipolaren Patienten
Ananth et al. 1979	bipolar	59	2	FKS	Rezidiv oder psychotrope Zusatzmedikation	
Aronoff et al. 1970	bipolar	12	Ø 2	teils KHS, teils RCT	affektives Rezidiv	Untergruppenauswertung der 12 bipolaren Patienten
Ayuso-Gutiérrez & Ramos-Brieva 1982	bipolar I und II (Feighner-Kriterien)	100	Ø 9,9	FKS	Ø Zeit bis zum nächsten Rezidiv	

Autoren und Publikationsjahr	Diagnose	n <sup>1</sup>	Studien-dauer (Jahre)	Studien-typ <sup>2</sup>	Evaluations-kriterium <sup>3</sup>	Bemerkungen
Baldessarini et al. 2000	bipolar I (n=218) und bipolar II (n=142) nach DSM-IV	360	Ø 13.3	FKS	Vielzahl von Evaluationskriterien (u.a. affektives Rezidiv)	Überschneidung des Patientenpools mit Tondo et al. 1998, 2001 und mit Viguera et al. 2001
Bouman et al. 1986	bipolar (Feighner-Kriterien)	56	mehrere Jahre (Details unklar)	KHS	affektives Rezidiv	Untergruppenauswertung der 56 bipolaren Patienten
Cazzullo et al. 1979	bipolar	26	unklar ("long-term treatment")	unklar	affektives Rezidiv	Untergruppenauswertung der 26 bipolaren Patienten
Coryell et al. 2000	Manie oder schizoaffektive Störung, manischer Typ (RDC)	181	≥ 0.5	FKS <sup>4</sup>	Morbidity Index	Aufgrund fehlender und widersprüchlicher Angaben konnte nur ein Teil der Prädiktoren ausgewertet werden
Cundall et al. 1972	bipolar	12	1	RCT, doppelblind	affektives Rezidiv	Untergruppenauswertung der 12 bipolaren Patienten
Denicoff et al. 1997	bipolar (DSM-III-R)	42	1	RCT, doppelblind	CGI	Auswertung der Behandlungszeit mit Lithium

Autoren und Publikationsjahr	Diagnose	n <sup>1</sup>	Studien-dauer (Jahre)	Studien-typ <sup>2</sup>	Evaluations-kriterium <sup>3</sup>	Bemerkungen
Dunner et al. 1976	bipolar I (Feighner-Kriterien)	96	2	teils FKS, teils KHS	Zeit bis zum ersten Rezidiv	Untergruppenauswertung der 96 mit Lithium behandelten Patienten. Die Auswertung bezieht sich auf den Zeitraum von 2 Jahren.
Dunner & Fieve 1974	bipolar I und II (Feighner-Kriterien)	55	0.5-5.5	FKS	affektives Rezidiv	
Engström et al. 1997	bipolar (DSM-III-R)	98	Ø 10-13 (je nach Untergruppe)	FKS	Phasen-frequenz	
Faedda et al. 1989, 1991	bipolar, saisonales Muster (DSM-III-R)	50	1	FKS	affektives Rezidiv	
Faravelli et al. 1980	bipolar	26	3-8	KHS	affektives Rezidiv und prä-post-Vergleiche der Phasen-frequenz	Untergruppenauswertung der 26 bipolaren Patienten
Franchini et al. 1999	bipolar (DSM-IV)	171	4-8	unklar	prä-post-Vergleiche der Phasen-frequenz	Untergruppenauswertung der 171 bipolaren Patienten
Gasperini et al. 1993	bipolar (DSM-III-R)	123	2	FKS	Phasen-frequenz	Untergruppenauswertung der 123 bipolaren Patienten

Autoren und Publikationsjahr	Diagnose	n <sup>1</sup>	Studien-dauer (Jahre)	Studien-typ <sup>2</sup>	Evaluations-kriterium <sup>3</sup>	Bemerkungen
Gelenberg et al. 1989	bipolar (DSM-III-R)	94	3	RCT	affektives Rezidiv	
Grof et al. 1987	bipolar affektiv oder bipolar schizoaffektiv (RDC)	50	2	unklar	prä-post-Vergleiche der Phasenfrequenz	
Haag et al. 1986, 1987	bipolar affektiv (RDC)	48	> 1 Jahr oder 1 Zykluslänge	FKS	prä-post-Vergleiche der Hospitalisierung	Untergruppenauswertung der 48 bipolaren Patienten
Hanus & Zapletalek 1984	bipolar affektiv	59	1	FKS	prä-post-Vergleiche	Untergruppenauswertung der 59 bipolaren Patienten
Hartong et al. 2003	bipolar (DSM-III-R)	44	2	RCT	affektives Rezidiv	
Johnston et al. 1980	bipolar	44	≥ 1	KHS	affektives Rezidiv	
Kato 2000	bipolar (DSM-III-R)	32	Ø 4.8	FKS	affektives Rezidiv	
Kocsis et al. 1979	bipolar	61	0.5-12.1	FKS	prä-post Vergleiche der Phasenfrequenz	
Kukopulos & Reginaldi 1980a,b	manisch depressive Psychose (ICD-9)	301	> 1	FKS	prä-post-Vergleiche der Schwere, Länge und Frequenz affektiver Rezidive	

Autoren und Publikations-jahr	Diagnose	n <sup>1</sup>	Studien-dauer (Jahre)	Studien-typ <sup>2</sup>	Evaluations-kriterium <sup>3</sup>	Bemerkungen
Kulhara et al. 1999	manisch depressive Psychose (ICD-9)	118	2-27	teils KHS, teils FKS	prä-post-Vergleiche der Frequenz affektiver Rezidive	
Kusalic & Engelsmann 1998	bipolare Störung (DSM-III-R)	29	2	KHS	Rezidiv oder klinisch bedeutsame Intervall-symptomatik	
Lusznat et al. 1988	bipolar I (durch BRMaS bestätigte klinische Diagnose)	20	1	RCT, doppel-blind	Schwellen-werte auf Depressions- und Manieskala	
Maj et al. 1986	bipolar (DSM-III)	23	2	KHS	affektives Rezidiv	Auswertung war nur für die Untergruppe mit niedrigem Lithiumspiegel (0.3 mEq/l bis 0.5 mEq/l) möglich
Maj et al. 1989, Maj 1990	bipolar (DSM-III)	118	2	unklar	prä-post Vergleiche der durchschnittlichen Morbidität	
Maj et al. 1996	bipolar (RDC)	65	5	KHS	Anzahl affektiver Rezidive (0 vs. $\geq 2$ )	Alle Patienten waren bei Einschluss bereits mindestens 5 Jahre stabil unter Lithium.

Autoren und Publikationsjahr	Diagnose	n <sup>1</sup>	Studien-dauer (Jahre)	Studien-typ <sup>2</sup>	Evaluations-kriterium <sup>3</sup>	Bemerkungen
Maj et al. 1998		247	5	KHS	affektives Rezidiv (Gruppe A der Originalarbeit vs. Gruppe B+C)	Untergruppenauswertung der in „Tabelle 1“ beschriebenen Patienten (247 von 402)
Maj et al 1985	bipolar (DSM-III)	68	2	KHS	affektives Rezidiv	
Mander 1986	bipolar (DSM-III)	48	2-14	FKS	affektives Rezidiv	Untergruppenauswertung der 48 Patienten mit Lithiumprophylaxe
Melia 1970	bipolar (n=6), rezidivierend manisch (n=1)	7	2	RCT	affektives Rezidiv	Untergruppenauswertung der 7 bipolaren Patienten des RCT
Murray et al. 1983	bipolar	121	bis zu 2	KHS	Hospitalisierungsrate	Untergruppenauswertung der 121 bipolaren Patienten
O'Connell et al. 1991	bipolar (DSM-III)	248	1	FKS	GAF-score	t-Werte und damit r's waren aufgrund fehlender Angaben z.T. nur näherungsweise berechenbar
Okuma 1993	bipolar (DSM-III-R)	108	2	FKS	50% Reduktion der Krankenzzeit	Untergruppenauswertung der 108 Patienten mit Lithiumprophylaxe

Autoren und Publikationsjahr	Diagnose	n <sup>1</sup>	Studien-dauer (Jahre)	Studien-typ <sup>2</sup>	Evaluations-kriterium <sup>3</sup>	Bemerkungen
Persson 1972	bipolar	10	2	KHS	affektives Rezidiv	Untergruppenauswertung ("Part I, Period B")
Prien et al. 1973, 1974	bipolar I	91	2	RCT, doppelblind	affektives Rezidiv	
Prien et al. 1984		42	2	RCT, doppelblind	affektives Rezidiv	Untergruppenauswertung der 42 (von 117) bipolaren Patienten mit Lithiumprophylaxe. Anteil der depressiven Indexepisoden aus der Gesamtgruppe der bipolaren Patienten geschätzt
Rosenthal et al. 1979	bipolar I (RDC)	66	bis zu 3.7	teils RCT, teils FKS	affektives Rezidiv	Die verwendeten Auswertungen beziehen sich auf die 32. Woche.
Sarantidis & Waters 1981	bipolar affektiv und schizoaffektiv (RDC)	46	2	FKS	prä-post Vergleich der Hospitalisierungszeit	
Schürhoff et al. 2000	bipolar (DSM-IV)	97	> 1	KHS	affektives Rezidiv	

Autoren und Publikationsjahr	Diagnose	n <sup>1</sup>	Studien-dauer (Jahre)	Studien-typ <sup>2</sup>	Evaluations-kriterium <sup>3</sup>	Bemerkungen
Small et al. 1991	manische Episode (DSM-III-R)	24	2	RCT	prä-post Vergleich des CGI	Untergruppenauswertung der 24 auf Lithium randomisierten Patienten
Stallone et al. 1973	bipolar	25	2	RCT	Psychotrope Zusatzmedikation	Untergruppenauswertung der 25 auf Lithium randomisierten Patienten
Stefos et al. 1996	bipolar I (n=14) und bipolar II (n=7) nach RDC	21	4	FKS für die ersten 3 Jahre, KHS für das vierte Jahr	affektives Rezidiv	
Tondo et al. 1998	bipolar I und II (DSM-IV)	317	Ø 6,35	KHS	< 50% der Zeit in Krankheitsphasen	Überschneidung des Patientenpools mit Baldessarini et al. 2000, Tondo et al. 2001, Viguera et al. 2001
Tondo et al. 2001	bipolar I und II (DSM-IV)	360	≥ 1	FKS	u.a. Zeit krank unter Lithium und prä-post Vergleiche der Zeit in Krankheitsphasen	Überschneidung des Patientenpools mit Baldessarini et al. 2000, Tondo et al. 1998, Viguera et al. 2001



Autoren und Publikationsjahr	Diagnose	n <sup>1</sup>	Studien-dauer (Jahre)	Studien-typ <sup>2</sup>	Evaluations-kriterium <sup>3</sup>	Bemerkungen
Upadhyaya et al. 1985	bipolar	35	≥ 2	FKS	prä-post Vergleiche der Häufigkeit affektiver Episoden	
Van der Velde et al. 1970	bipolar	55	3	KHS	affektives Rezidiv	Untergruppenauswertung der 55 Patienten, die stabilisiert werden konnten
Viguera et al. 2001	bipolar I (n=218) und bipolar II (n=142) nach DSM-IV	360	Ø 4.5	FKS	affektives Rezidiv	Überschneidung des Patientenpools mit Baldessarini et al. 2000 und Tondo et al. 1998, 2001
Waters et al. 1982	bipolar (DSM-III)	29	1	RCT	affektives Rezidiv	
Yang 1985	bipolar (DSM-III)	101	2	FKS	prä-post Vergleiche der Häufigkeit affektiver Episoden	Stichprobe chinesischer Patienten. Dichotomisierung der Ergebnisse in „Tabelle 1“ (responders vs. non-responders); dabei wurden die Standardabweichungen aus der Gruppe der non-responders geschätzt.

Autoren und Publikationsjahr	Diagnose	n <sup>1</sup>	Studien-dauer (Jahre)	Studien-typ <sup>2</sup>	Evaluations-kriterium <sup>3</sup>	Bemerkungen
Yazici et al. 1999	bipolar I und II (DSM-IV)	141	2	FKS	prä-post Vergleiche affektiver Episoden und des Ø Morbidity-Index	
Young et al. 1993	bipolar I und II (RDC)	81	1	KHS	Hospitalisierung oder Behandlung mit Antidepressiva oder Neuroleptika	

<sup>1</sup>Die Fallzahl n bezieht sich auf die Patienten, für die Einschlusskriterien der Metaanalyse erfüllt waren (z.B. auf alle bipolaren Patienten einer Studie).

<sup>2</sup>RCT = Randomized Controlled Trial (Randomisierte Therapievergleichsstudie), KHS = Kohortenstudie (prospektives Design), FKS = Fall-Kontroll-Studie (retrospektives Design).

<sup>3</sup>Kriterium, aufgrund dessen in der Primärstudie beurteilt wurde, ob ein Patient auf die Lithiumprophylaxe reagiert hat oder nicht.

<sup>4</sup>Aufgrund des response-korrelierten Einschlusskriteriums  $\geq 6m$  unter Lithium als FKS klassifiziert.

## 4.3 Ergebnisse der Metaanalyse zu demographischen Variablen

### 4.3.1 Geschlecht

Insgesamt war der Zusammenhang zwischen Geschlecht und prophylaktischer Wirkung von Lithium für 1779 Patienten in 24 Studien, die alle Einschlusskriterien erfüllten, quantifizierbar. Wie in Abb. 6 dargestellt zeigte sich keinerlei Zusammenhang zwischen Geschlecht und Response ( $\bar{r} = -0.002$  für weibliches Geschlecht, KFI = [-0.046, 0.049],  $p = 0.94$ ). Die empirische Heterogenität der Effektstärken zwischen den Studien ist mit der Homogenitätsannahme vereinbar ( $\chi^2 = 21.6$ ,  $p = 0.54$ ). Der Funnelplot gibt keinen Anlass, einen Publikationsbias anzunehmen (vgl. Abb. 7).

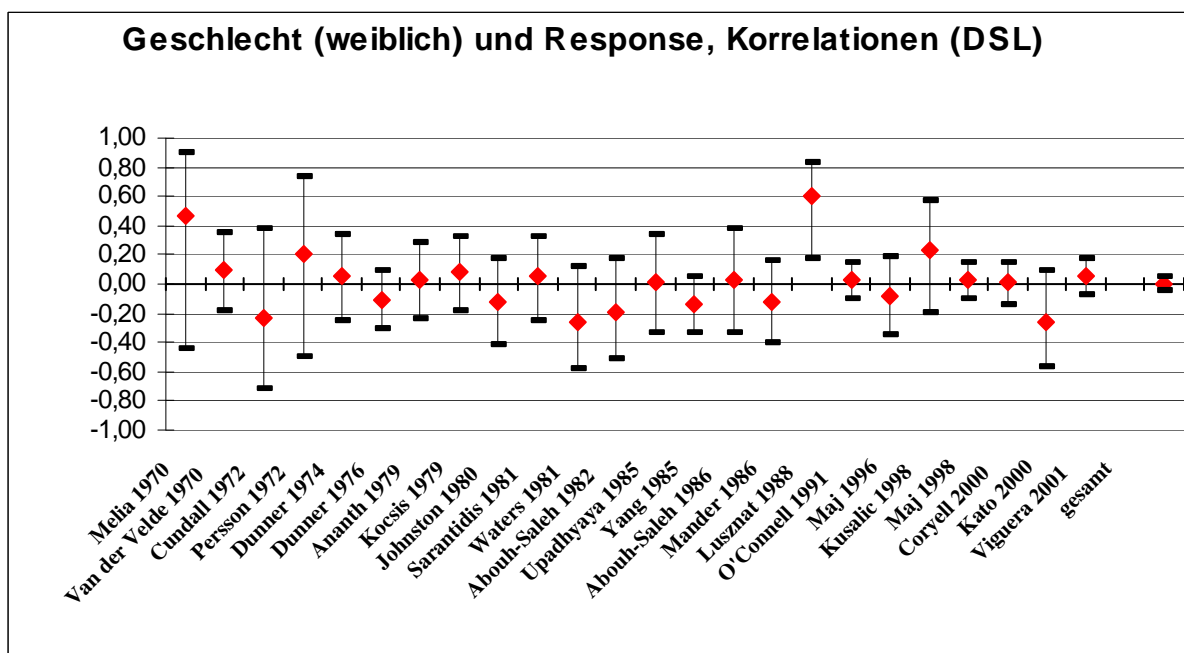


Abbildung 6. Punkt-biseriale Korrelation  $r$  zwischen weiblichem Geschlecht und prophylaktischer Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

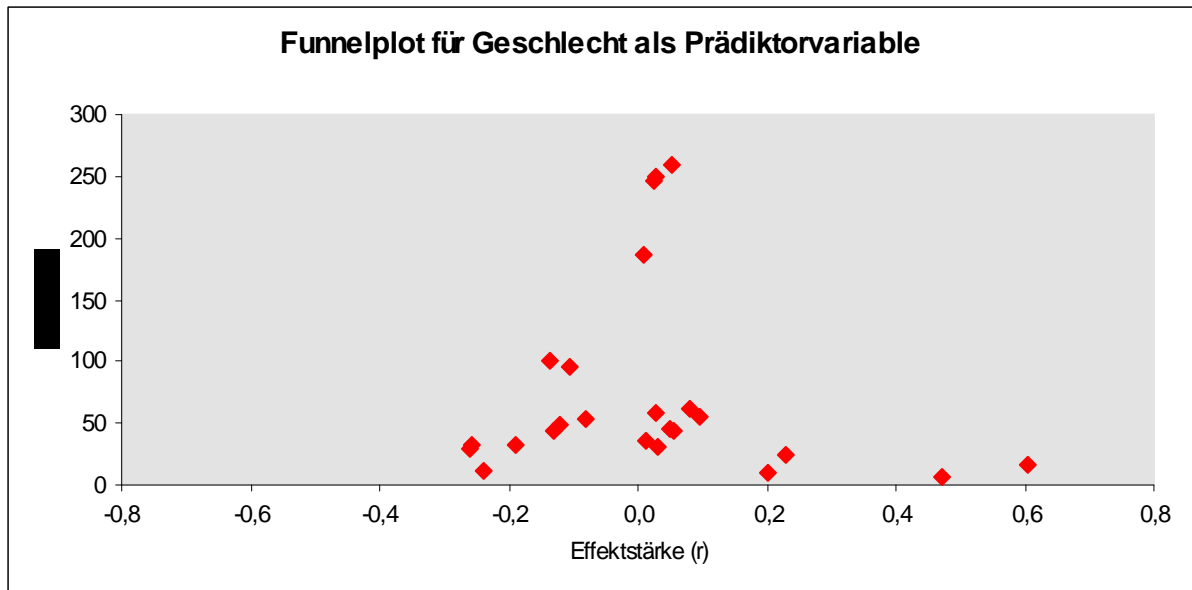


Abbildung 7. Funnelplot für die Variable Geschlecht: Effektstärke  $r$  in Abhängigkeit von der Fallzahl  $n$  in den einzelnen Stichproben.

Aufgrund der hohen Anzahl in die Metaanalyse eingeschlossener Patienten und aufgrund der insgesamt hohen Einheitlichkeit der Ergebnisse der Primärstudien lässt sich mit einiger Sicherheit ausschließen, daß das Geschlecht einen starken Einfluss auf die Lithiumresponse hat (das Konfidenzintervall zum Vertrauensgrad 0,95 = [-0,046, 0,049])

### 4.3.2 Alter

Insgesamt gingen die Daten von 1291 Patienten aus 17 Studien in die Metaanalyse zum Zusammenhang zwischen Alter und therapeutischem Ansprechen auf eine Lithiumprophylaxe ein (vgl. Abb. 8). Bei einer metaanalytischen Zusammenfassung dieser Studien zeigt sich kein statistisch bedeutsamer Zusammenhang zwischen Alter bei Studienbeginn und Therapieresponse ( $\bar{r} = -0.035$ , KFI =  $[-0.140, 0.071]$ ,  $p = 0.52$ ). Die empirische Heterogenität der Effektstärken zwischen den Studien ist statistisch nicht mit der Homogenitätsannahme vereinbar ( $\chi^2 = 46.4$ ,  $p = 0.00009$ ). Die Heterogenität der Primärstudien wurde durch Aggregation nach dem random-effects model (DerSimonian & Laird 1986, vgl. Kap. 4.1.2.2) berücksichtigt. Eine eingehende Untersuchung der Quellen dieser Heterogenität in einer stärker inhaltlich ausgerichteten Analyse wurde im Rahmen dieser Arbeit nicht durchgeführt. Aufgrund des Funnelplots gibt es keinen Anhalt für einen Publikationsbias (vgl. Abb. 9).

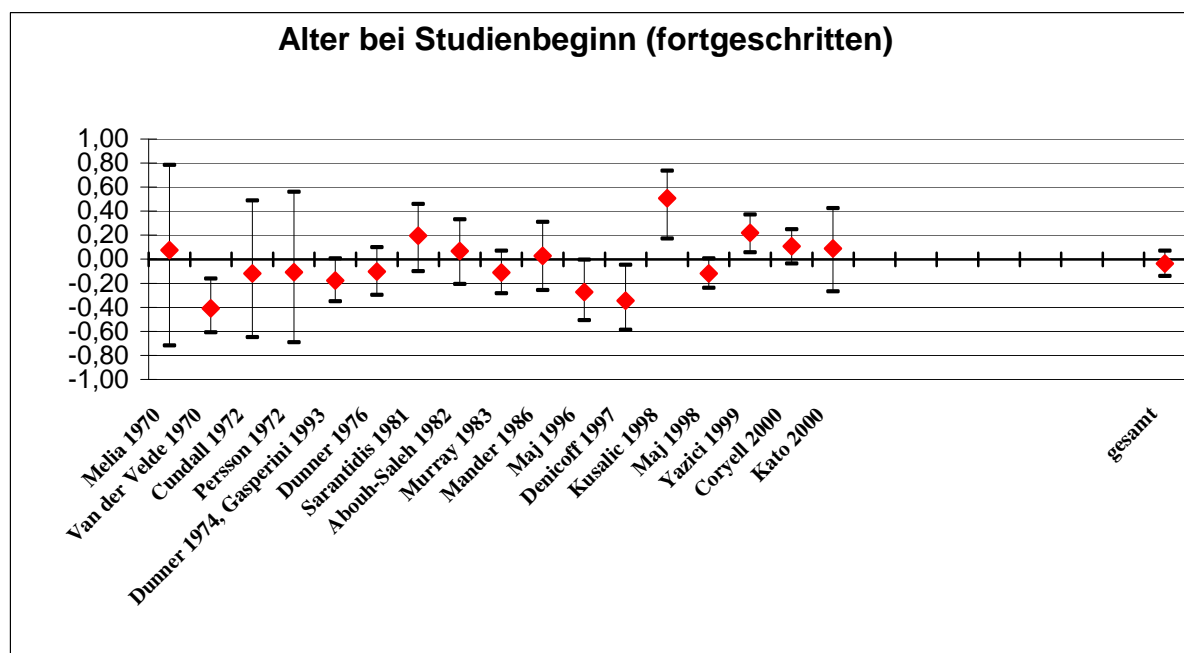


Abbildung 8. Korrelation  $r$  zwischen Alter bei Studienbeginn und prophylaktischer Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

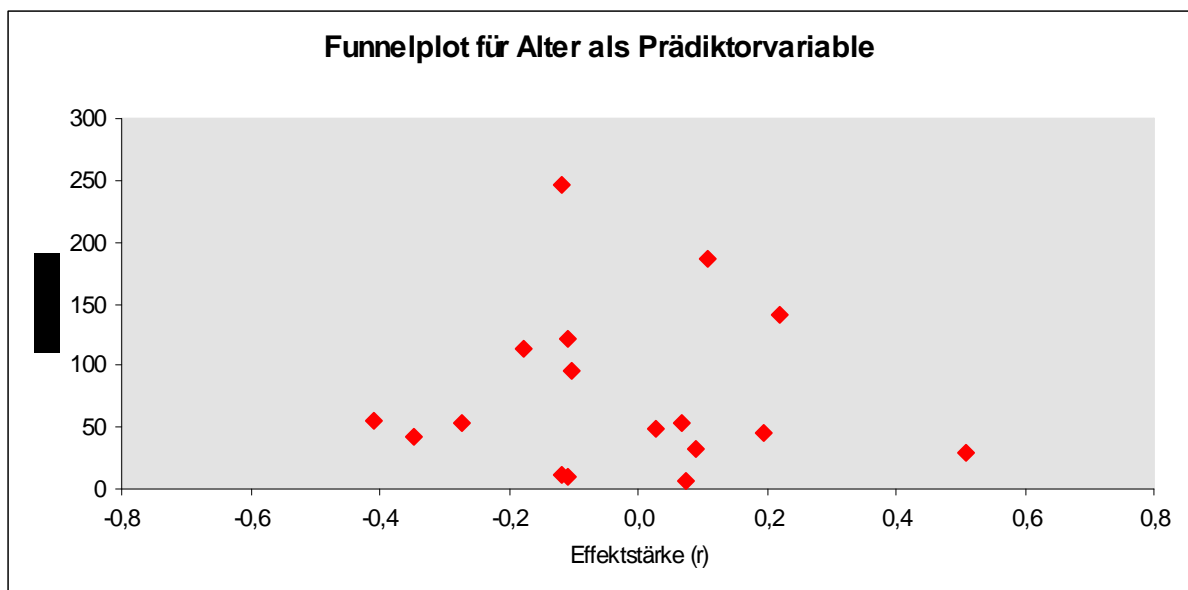


Abbildung 9. Funnelplot für die Variable Alter bei Studienbeginn: Effektstärke  $r$  in Abhängigkeit von der Fallzahl  $n$  in den einzelnen Stichproben.

### 4.3.3 Familienstand

Um präzisere und besser quantifizierbare Aussagen zu erlauben, wurde die Variable Familienstand in drei Hauptvariablen aufgeteilt:

Familienstand – ledig (ja/nein)

Familienstand – verheiratet (ja/nein)

Familienstand – getrennt/geschieden/verwitwet (ja/nein)

Familienstand – ledig

Aufgrund der Primärstudien sind nur begrenzte metaanalytische Aussagen über den Zusammenhang zwischen Familienstand (ledig) und therapeutischem Ansprechen auf eine Lithiumprophylaxe möglich. Insgesamt wird in 5 Studien berichtet, daß dieser Zusammenhang untersucht wurden, aber nur in 2 dieser 5 Studien (389 von 882 Patienten) wird das Ergebnis in einer quantifizierbaren Form berichtet. Aufgrund der beiden quantifizierbaren Studien zum Familienstand (ledig) ergab sich im Modell mit zufälligen Effekten ein nicht signifikanter Trend der dafür spräche, daß ein lediger Familienstand eher ungünstig für den weiteren Verlauf unter einer Lithiumprophylaxe wäre ( $\bar{r} = -0.134$ , KFI = [-0.272, 0.008],  $p = 0.06$ , vgl. Abb. 10). Aufgrund der genannten Einschränkungen sollte dieses Ergebnis allerdings mit Zurückhaltung interpretiert werden.

Die empirische Heterogenität der Effektstärken zwischen den beiden Studien ist statistisch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 1.932$ ,  $p = 0.16$ ). Da in den drei übrigen Studien berichtet wurde, daß kein signifikanter Zusammenhang zwischen ledigem Familienstand und Lithiumresponse bestand, ist zudem mit einem erheblichen Bias, der zu einer Überschätzung der Effektstärke führt, zu rechnen.

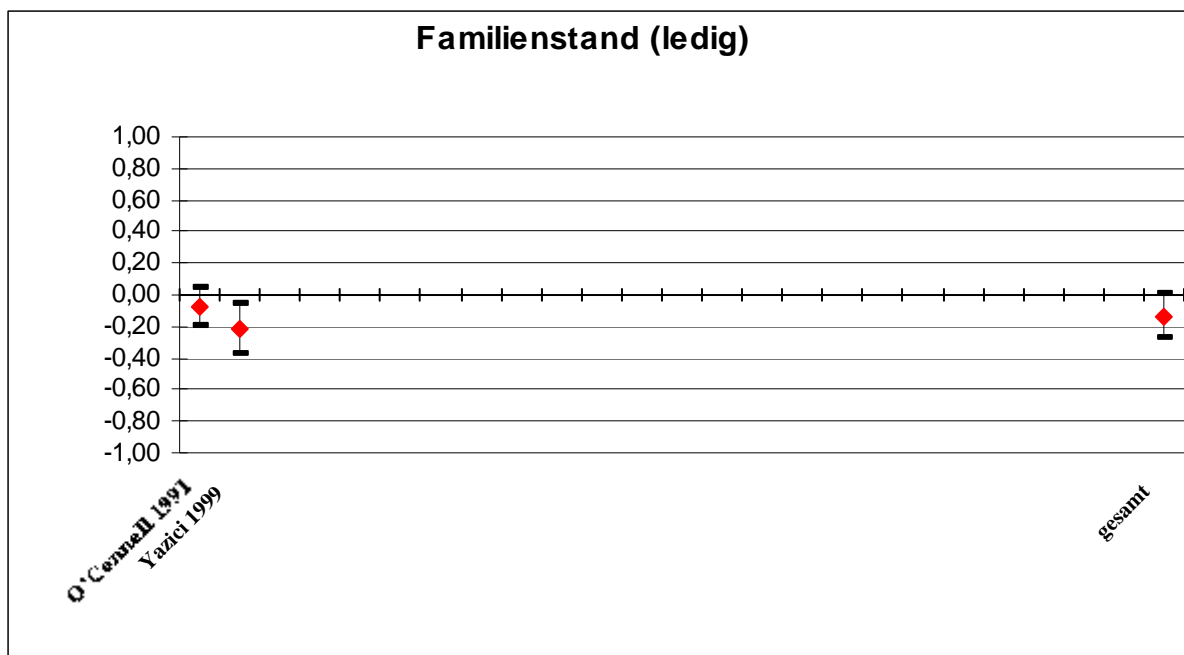


Abbildung 10. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen Familienstand (ledig) und prophylaktischer Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### Familienstand – verheiratet

Mit dem eben angesprochenen Problem, einem Bias der quantitativ auswertbaren publizierten Ergebnisse, ist hier in noch verstärktem Maße zu rechnen. Insgesamt wird in 3 Studien (665 Patienten) der Zusammenhang zwischen verheiratetem Familienstand und Lithiumresponse untersucht, doch nur in einer Studie (mit 248 Patienten) wird das Ergebnis in einer quantifizierbaren Form dargestellt (O'Connell et al. 1991). Auch hier wurde in den übrigen Studien (Denicoff et al. 1997, Tondo et al. 2001) berichtet, daß kein signifikanter Zusammenhang zwischen Familienstand (verheiratet) und Lithiumresponse bestand. Deshalb ist mit einem erheblichen, zu einer Überschätzung der Effektstärke führenden Bias zu rechnen.

In der Studie, die den Zusammenhang zwischen verheiratetem Familienstand und Lithiumresponse exakt berichtet (O'Connell et al. 1991), fand sich ein signifikant besseres Ansprechen bei verheirateten Patienten ( $r = 0.151$ , KFI = [0.027, 0.270],  $p = 0.02$ ). Da hier allerdings mit einem bedeutsamen Bias zu rechnen ist, der dieses Ergebnis komplett in Frage stellt, wird dieses Ergebnis mit großer Vorsicht interpretiert. Dies heißt insbesondere, daß im folgenden *nicht* davon ausgegangen wird, daß sich diese Variable in dieser Metaanalyse als Prädiktor erhärtet hat – vielmehr wird die Situation für die Variable Familienstand im Sinne einer konservativen Auswertungsstrategie weiterhin als unklar betrachtet.



### Familienstand – getrennt/geschieden/verwitwet

Die Datenlage aufgrund der Primärstudien für den Einfluss der Variable „Familienstand – getrennt/geschieden/verwitwet“ auf das Ansprechen auf eine Lithiumtherapie ist dieselbe wie für den Einfluss Variable „Familienstand – verheiratet“. Auch hier lässt nur die Studie von O’Connell et al. (1991) eine Quantifizierung dieses Zusammenhangs zu, während die anderen Studien in nicht näher quantifizierbarer Weise „nicht signifikante“ Zusammenhänge berichten. In dieser Studie fand sich kein signifikanter Zusammenhang zur Lithiumresponse ( $r = -0.067$ , KFI =  $[-0.189, 0.059]$ ,  $p = 0.30$ ). Insgesamt ergab sich also kein Anhalt für einen Zusammenhang zwischen der Variable „Familienstand – getrennt/geschieden/verwitwet“ und dem Ansprechen auf eine Lithiumtherapie, bei einer allerdings sehr schmalen empirischen Basis.

## 4.4 Ergebnisse der Metaanalyse zu klinischen Variablen

### 4.4.1 Diagnostik und Psychopathologie

#### 4.4.1.1 Bipolar-II-Störung

Der Zusammenhang zwischen der Diagnose einer bipolar-II-Störung und der prophylaktischen Wirkung von Lithium war für 1023 Patienten aus 8 Studien, die alle Einschlusskriterien erfüllten, quantifizierbar. In 7 dieser 8 Studien war dieser quantitative Zusammenhang direkt erfassbar. In der achten Studie, der Studie von Ayuso-Gutiérrez & Ramos-Brieva (1982) wurde von einem schlechteren Verlauf der bipolar-II-Patienten ( $p < 0.01$ ) berichtet. Da diese Aussage Information über das Ausmaß und die Richtung des Zusammenhangs zwischen der Diagnose einer bipolar-II-Störung und Lithiumresponse enthält, wurde sie in eine Schätzung der Effektstärke  $r$  übersetzt. Aufgrund der Art, wie die Ergebnisse in der Arbeit von Ayuso-Gutiérrez & Ramos-Brieva (1982) berichtet werden, ist anzunehmen, daß "p < 0.01" bedeutet, daß der p-Wert im Bereich von 0.001 bis 0.01 liegt. Der Mittelpunkt dieses Intervalls ( $p = 0.0055$ ) wurde als Schätzwert für den p-Wert angesetzt und hieraus wurde schließlich ein Effektstärkemaß  $r$  berechnet (zur Berechnung von  $r$  auf Grundlage eines p-Wertes vgl. Kap. 4.1.3).

Wie in Abb. 11 dargestellt zeigen die 8 untersuchten Studien ein sehr heterogenes Bild. Gemäß zwei dieser Studien (Kukopulos et al. 1980a,b und Ayuso-Gutiérrez & Ramos-Brieva 1982) war eine bipolar-II-Störung signifikant mit einem ungünstigen Verlauf assoziiert, während die Studien von Yaziki et al. (1999) und von Tondo et al. (2001) einen signifikant besseren Verlauf für bipolar-II-Patienten berichten. Dementsprechend zeigt der Test auf Heterogenität der Effektstärken eine klare Verletzung der Homogenitätsannahme an ( $\chi^2 = 28.0$ ,  $p = 0.0002$ ). Bei einer quantitativen Zusammenfassung aller Studien (random effects model) zeigt sich kein Zusammenhang zwischen der Diagnose einer Bipolar-II-Störung und der Response ( $\bar{r} = 0.011$ ,  $KFI = [-0.125, 0.147]$ ,  $p = 0.87$ ). Der Funnelplot ist bei insgesamt nur 8 Studien gewiss nicht sehr aussagekräftig, gibt aber keinen Anlass, einen Publikationsbias anzunehmen (vgl. Abb. 12). Allenfalls der hohe Anteil signifikanter, aber stark differierender Ergebnisse an den publizierten Studien könnte als Hinweis auf einen Publikationsbias (hier: bevorzugte Publikation signifikanter Ergebnisse) gewertet werden.

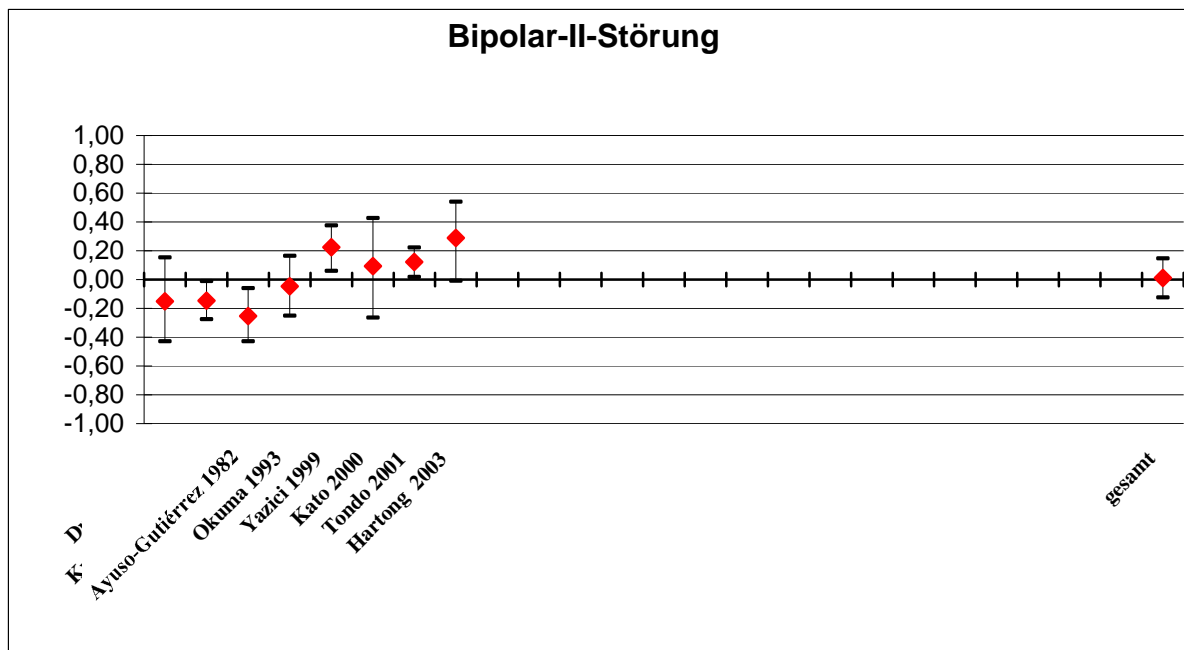


Abbildung 11. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen einer Bipolar-II-Störung und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

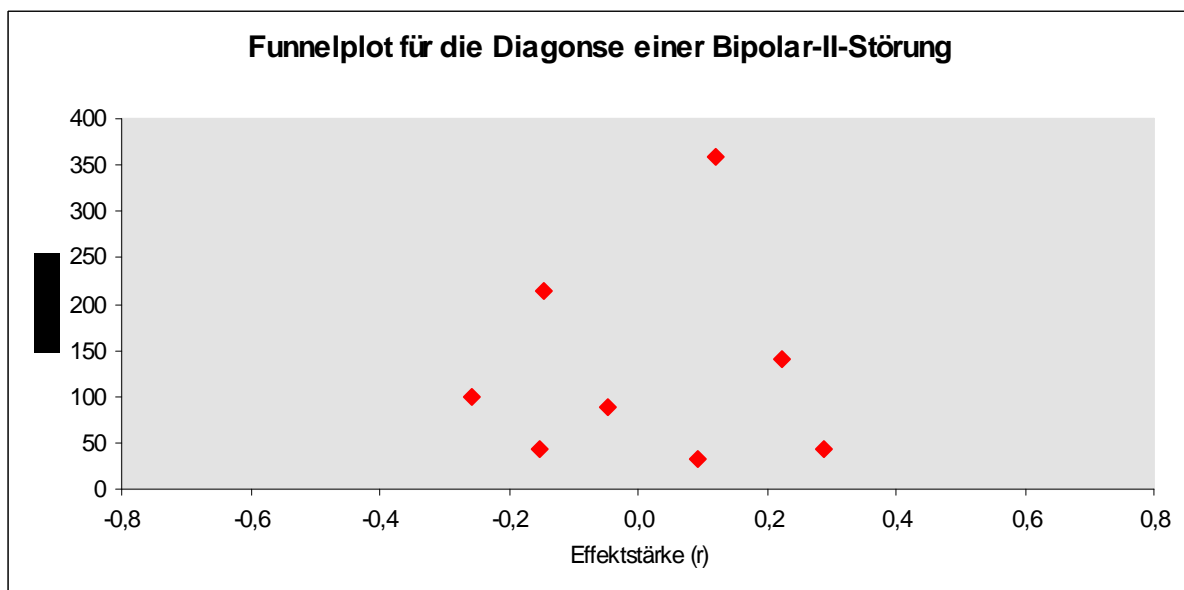


Abbildung 12. Funnelplot für die Diagnose einer Bipolar-II-Störung: Effektstärke  $r$  in Abhängigkeit von der Fallzahl  $n$  in den einzelnen Stichproben.

#### 4.4.1.2 Manische Indexepisode

Der Einfluss eines manischen Zustandsbildes während der Indexepisode auf die Lithiumresponse wurde in 4 Primärstudien mit insgesamt 514 Patienten untersucht. Auch hier fällt eine außerordentlich hohe empirische Heterogenität der in diesen Studien berichteten Ergebnisse auf (vgl. Abb. 13). Der Test auf Heterogenität der Effektstärken weist dementsprechend wieder auf eine klare Verletzung der Homogenitätsannahme hin ( $\chi^2 = 37.5$ ,  $p = 0.00000004$ ). Bei einer quantitativen Zusammenfassung aller Studien nach dem random effects model zeigt sich kein Einfluss einer manischen Indexepisode auf den weiteren Verlauf ( $\bar{r} = -0.084$ ,  $KFI = [-0.379, 0.226]$ ,  $p = 0.60$ ). Aufgrund der extrem starken Heterogenität ergibt sich bei der Zusammenfassung aller Studien nach dem random effects model (DerSimonian-Laird) ein breiteres Konfidenzintervall als in den einzelnen Studien. Die Metaanalyse führte hier (ausnahmsweise) nicht zu einer Präzisierung der Aussage über die Effektstärke, sondern zu einer Aufdeckung der großen, sich in Heterogenität der Primärstudien äußernden Unsicherheit der Forschungslage.

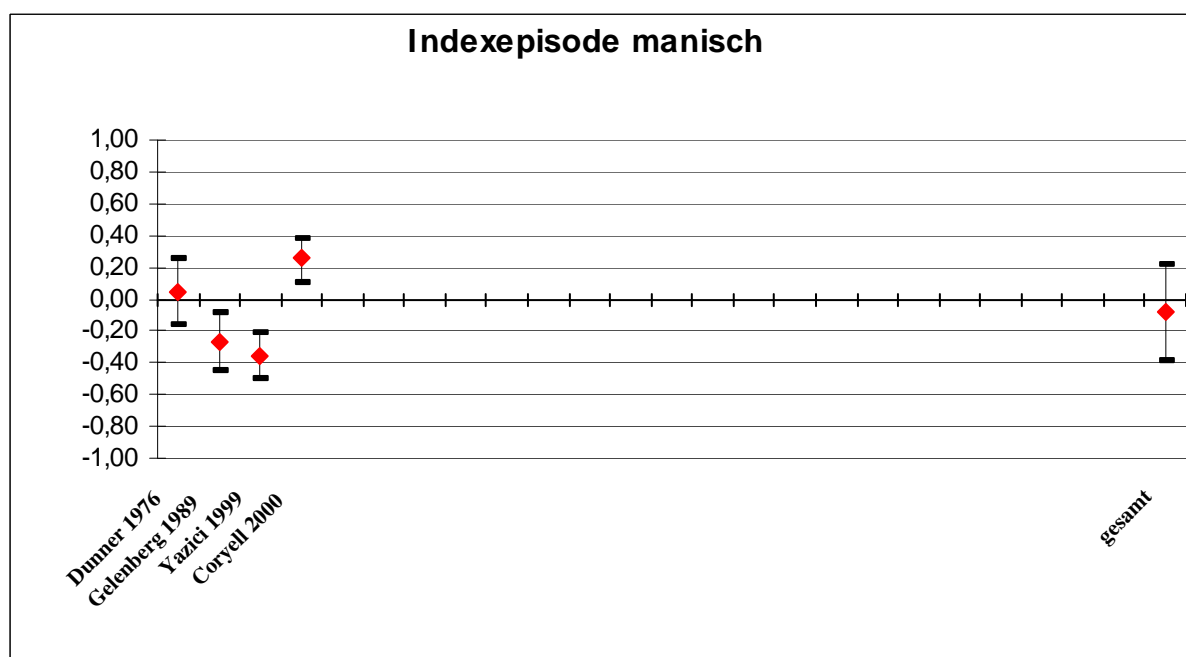


Abbildung 13. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen einer manischen Indexepisode und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.1.3 Depressive Indexepisode

Insgesamt wurden 4 Primärstudien mit insgesamt 459 Patienten zum Zusammenhang zwischen einem depressiven Zustandsbild während der Indexepisode und dem darauffolgenden Ansprechen auf eine Lithiumprophylaxe lokalisiert. Wie in Abb. 14 dargestellt ergab sich bei einer quantitativen Zusammenfassung dieser Studien kein Zusammenhang zwischen depressiver Ausprägung der Indexepisode und einem günstigen Verlauf unter Lithium ( $\bar{r} = 0.003$ , KFI = [-0.127, 0.132],  $p = 0.97$ ). Die empirische Heterogenität der Effektstärken zwischen den Studien ist statistisch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 5.37$ ,  $p = 0.15$ ).

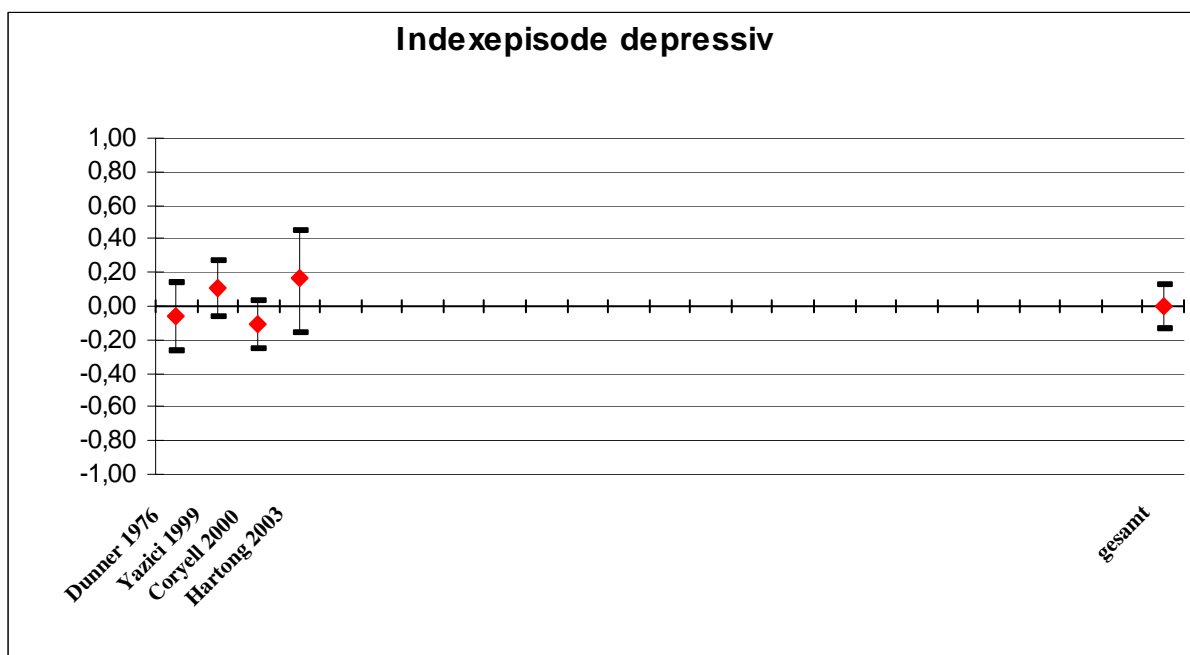


Abbildung 14. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen einer depressiven Indexepisode und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.1.4 Gemischte Indexepisode

Zur Prädiktion der Lithiumresponse bei bipolaren Patienten aufgrund eines manisch-depressiven Mischzustands während der Indexepisode wurde nur die an  $n = 186$  Patienten durchgeführte Studie von Coryell et. al. (2000) lokalisiert. Im Ergebnis zeigte sich in dieser Studie ein tendenziell negativer Einfluss des Mischzustands auf das Ansprechen auf eine Lithiumprophylaxe ( $r = -0.138$  (KFI = [-0.276, 0.006],  $p = 0.06$ ).

#### 4.4.1.5 Zyklische Indexphase

In zwei Primärstudien (Dunner et al. 1976, Coryell et al. 2000) wurde an insgesamt  $n = 277$  Patienten untersucht, ob eine zyklische Indexphase (d.h. ein direkter Wechsel von einer Depression in eine Manie oder umgekehrt) einen Einfluss auf die Lithiumresponse hat. Bei einer quantitativen Zusammenfassung der beiden Studien zeigte sich im Ergebnis kein statistisch signifikanter Einfluss des direkten Syndromwechsels in der Indexphase auf das Ansprechen auf eine Lithiumprophylaxe ( $r = -0.097$  (KFI =  $[-0.277, 0.089]$ ,  $p = 0.31$ , vgl. Abb. 15). Die empirische Heterogenität der beiden Effektstärken ist statistisch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 2.21$ ,  $p = 0.14$ ).

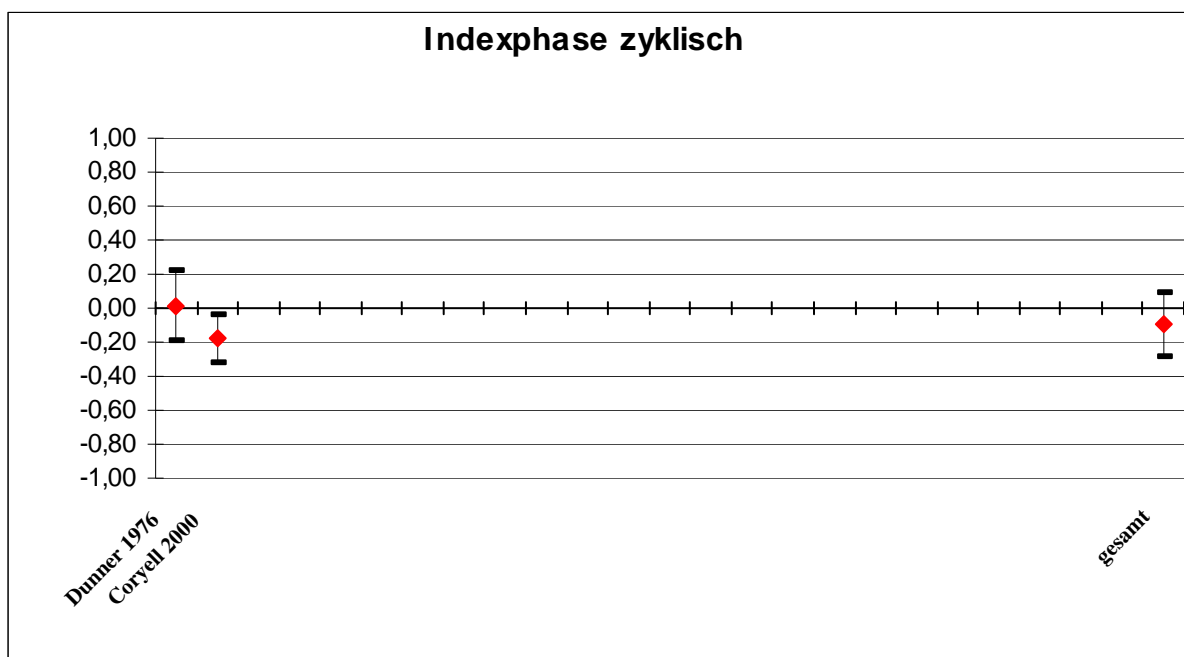


Abbildung 15. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen einer zyklischen Indexphase und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.1.6 Psychotische Symptome

Der Einfluss von psychotischen Symptomen und/oder Wahnvorstellungen (nicht nach Stimmungskongruenz unterschieden) während der Indexepisode auf den Langzeitverlauf unter Lithium wurde in 5 Primärstudien mit insgesamt 863 Patienten untersucht. Bei einer metaanalytischen Zusammenfassung aller Studien (random effects model) zeigte sich kein signifikanter Einfluss psychotischer Symptomatik auf den weiteren Verlauf ( $\bar{r} = -0.149$ , KFI =  $[-0.325, 0.038]$ ,  $p = 0.12$ ).

Zu beachten ist allerdings eine visuell (vgl. Abb. 16) wie statistisch erhebliche Heterogenität ( $\chi^2 = 26.0$ ,  $p = 0.00003$ ). Möglicherweise hängt die Heterogenität der Studienergebnisse in diesem Fall damit zusammen, daß unter psychotischen Symptomen in den verschiedenen Studien unterschiedliches gemeint war. In den beiden folgenden Kapiteln (4.4.1.7 – 4.4.1.8) wird daher nach der Art psychotischer Symptome (stimmungskongruente psychotische Symptome) differenziert.

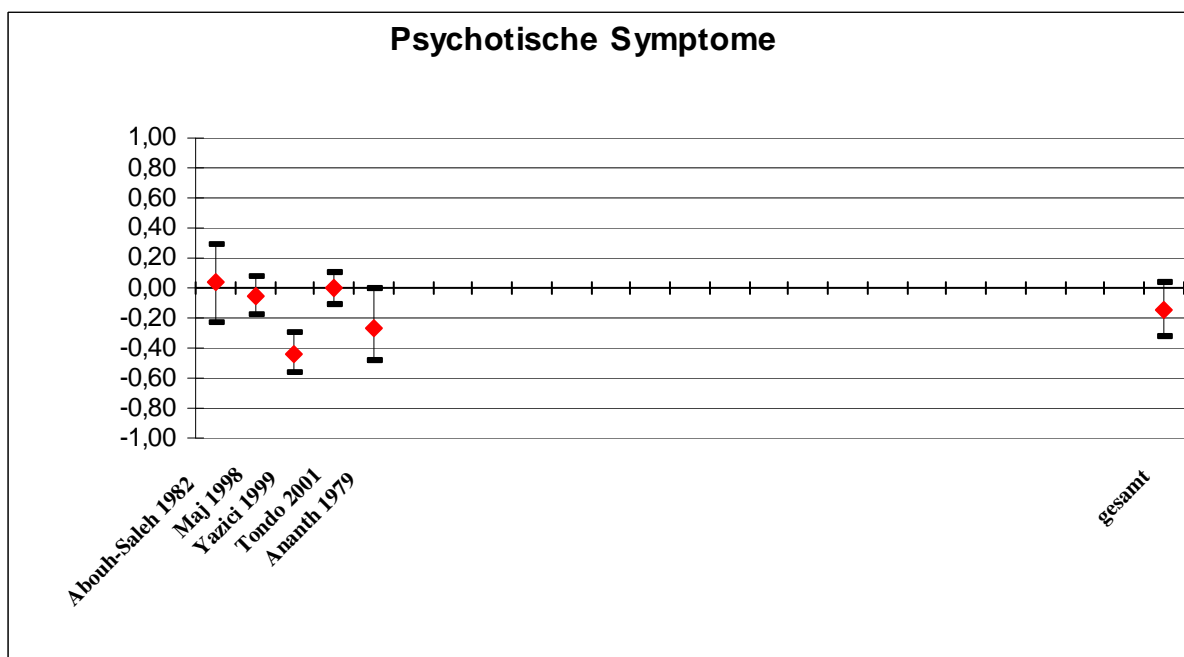


Abbildung 16. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen psychotischer Symptome während der Indexepisode und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.1.7 Stimmungskongruente psychotische Symptome

Stimmungskongruente (auch synthyme) psychotische Symptome zeichnen sich dadurch aus, daß sie zur affektiven Gestimmtheit des Patienten passen, also mit der Affektlage des Patienten kongruent sind. Beispiele wären ein Größenwahn während einer manischen Episode oder ein Schuldwahn während einer depressiven Episode.

In der einzigen Primärstudie, in der explizit der Zusammenhang zwischen stimmungskongruenten psychotischen Symptomen und dem Ansprechen auf eine Lithiumprophylaxe untersucht wird (Maj et al. 1986) fand sich bei  $n = 23$  Patienten kein signifikanter Zusammenhang zwischen stimmungskongruenter psychotischer Symptomatik und Lithiumresponse ( $r = -0.098$ , KFI =  $[-0.328, 0.490]$ ,  $p = 0.66$ ).

#### 4.4.1.8 Stimmungsinkongruente psychotische Symptome während der Indexepisode

Stimmungsinkongruente psychotische Symptome oder offensichtlich in diesem Sinne als „atypischer Symptomatik“ bezeichnete Symptome wurde in 4 Primärstudien mit insgesamt 202 Patienten als Prädiktorvariable untersucht. Bei einer quantitativen Zusammenfassung dieser Studien (vgl. Abb. 17) zeigte sich ein Trend, wonach das Vorliegen atypischer Symptomatik ungünstig für den weiteren Verlauf unter Lithium war ( $\bar{r} = -0.237$ , KFI =  $[-0.478, 0.037]$ ,  $p = 0.09$ ). Die empirische Heterogenität der Effektstärken zwischen den Studien ist statistisch nicht mit der Homogenitätsannahme vereinbar ( $\chi^2 = 10.6$ ,  $p = 0.01$ ). Die Heterogenität der Primärstudien wurde durch Aggregation nach dem random-effects model (DerSimonian & Laird 1986, vgl. Kap. 4.1.2.2) berücksichtigt. Eine eingehende Untersuchung der Quellen dieser Heterogenität in einer stärker inhaltlich ausgerichteten Analyse wurde im Rahmen dieser Arbeit nicht durchgeführt.



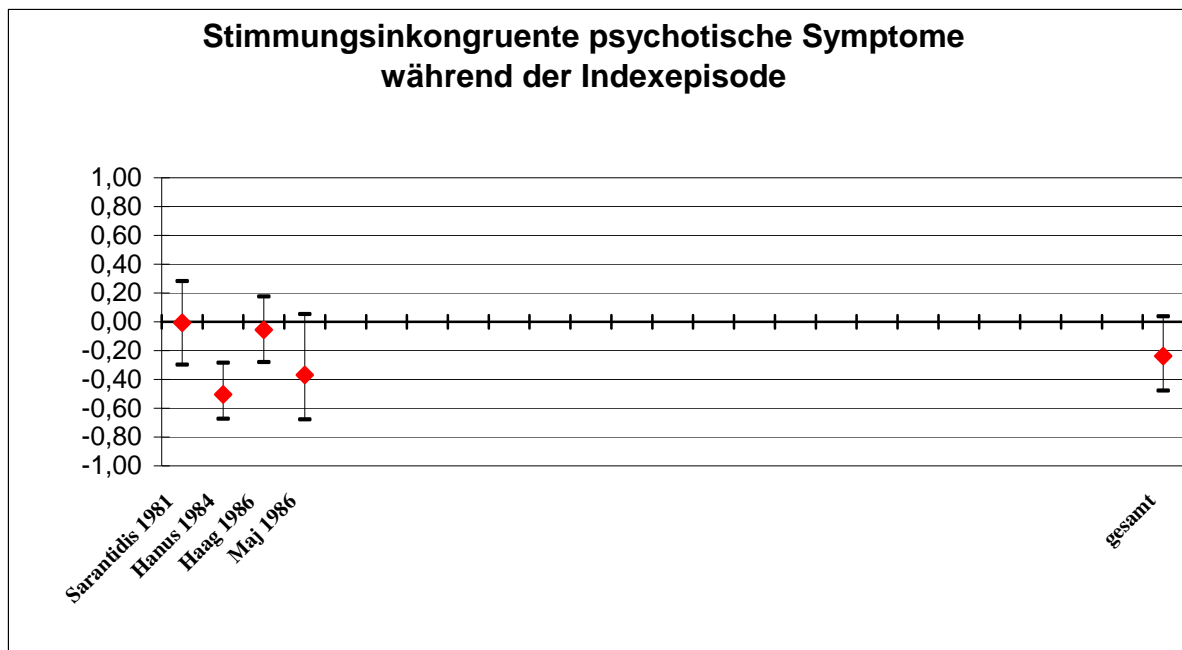


Abbildung 17. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen stimmungsinkongruenter Symptome während der Indexepisode und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

In zwei Primärstudien (Faravelli et al. 1980, Okuma 1993) wurde der Begriff „atypische Symptomatik“ in einem wohl abweichenden Sinn gebraucht, und deshalb wurden diese beiden Studien nicht mit den übrigen 4 Studien zusammengefasst. Faravelli et al. 1980 verstehen unter atypischer Symptomatik Symptome wie Zwanghaftigkeit, Verwirrtheit oder Wahnvorstellungen. Zwar wurden in dieser Primärstudie keine separaten Ergebnisse für bipolare Patienten ausgewiesen, doch ist es aufgrund der Ergebnisdarstellung möglich, die Effektstärke zum Zusammenhang  $r$  zwischen atypischer Symptomatik und Lithiumresponse zumindest abzuschätzen. Gemäß dieser Abschätzung liegt die Punktschätzung für  $r$  im Bereich von -1.00 bis -0.13.

Okuma et al. (1993) verstehen unter atypischer Symptomatik "atypical symptoms such as alteration of consciousness". In dieser Studie zeigte sich kein eindeutiger Zusammenhang zwischen atypischer Symptomatik und einem günstigen Verlauf unter Lithium ( $\bar{r} = 0.063$ , KFI = [-0.127, 0.249],  $p = 0.52$ ).

#### 4.4.1.9 Klassische euphorische Manie

Zum Einfluss eines klassischen euphorischen Zustandsbildes (Ideenflucht, Größenideen, vorwiegend euphorische Stimmung, ...) während einer manischen Indexepisode auf die Lithiumresponse wurden 3 Studien mit insgesamt  $n = 105$  Patienten lokalisiert (Ananth et al. 1979, Luszmat et al. 1988, Kusalic & Engelsmann 1998). Auch hier zeigt sich bereits visuell eine relativ starke Heterogenität der Effektstärken (vgl. Abb. 18), die auch zu einer formellen Ablehnung der Homogenitätsannahme führt ( $\chi^2 = 7.02$ ,  $p = 0.03$ ). Bei einer quantitativen Zusammenfassung aller Studien nach dem random effects model zeigt sich kein Einfluss eines klassischen euphorischen Zustandsbildes während der Indexepisode auf den weiteren Verlauf ( $\bar{r} = -0.006$ ,  $KFI = [-0.394, 0.384]$ ,  $p = 0.98$ ). Aufgrund der starken Heterogenität ist die Zusammenfassung der Studien ohne eine eingehende Prüfung der Heterogenitätsquellen allerdings nur von bedingter Aussagekraft.

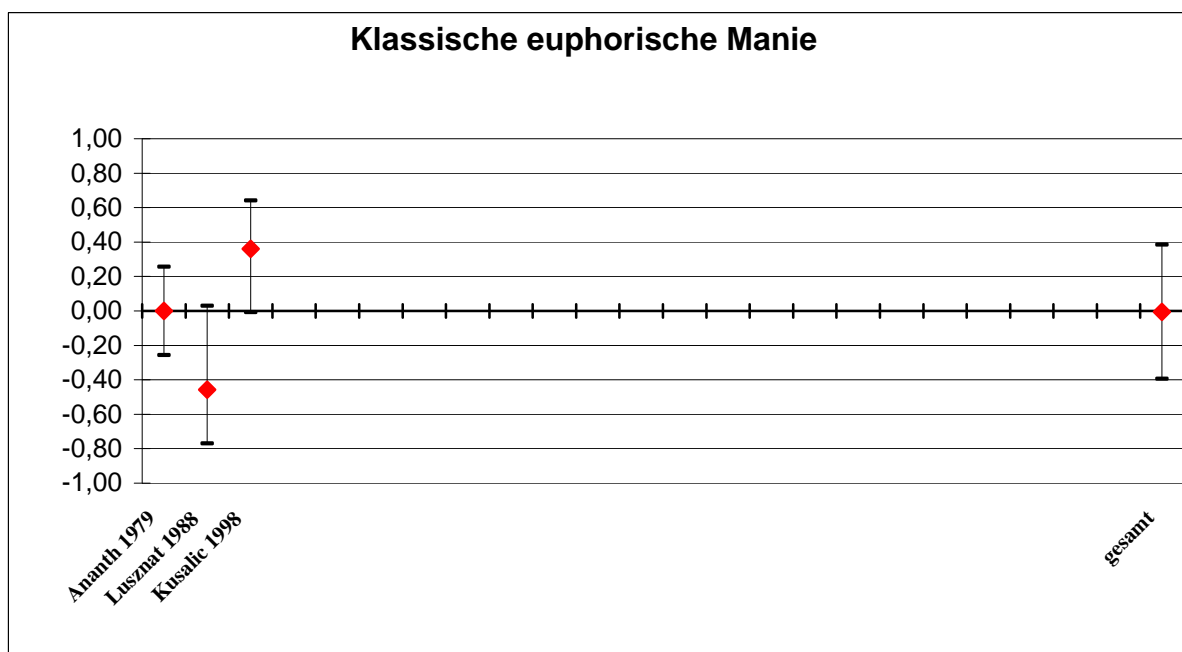


Abbildung 18. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen einer klassischen euphorischen Manie während der Indexepisode und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.1.10 Reizbarkeit, Agitiertheit

In der Studie von Ananth et al. (1979) wird eine Reihe psychischer Symptome, die in Kapiteln 4.4.1.10 – 4.4.1.15 besprochen werden, mit der Lithiumresponse in Zusammenhang gebracht. Leider wird aus der Originalarbeit nicht klar, zu welchem Zeitpunkt der Erkrankung die Psychopathologie erhoben wurde. Dies stellt eine gewisse Einschränkung für die Interpretierbarkeit der Ergebnisse dar. Die Studie von Ananth et al. (1979) ist eine an  $n = 59$  durchgeführte retrospektive Prädiktionsstudie mit einer Beobachtungszeit von 2 Jahren.

Reizbarkeit und Agitiertheit standen in gemäß dieser Studie in keinem signifikanten Zusammenhang mit einem Rezidiv unter Lithium ( $\bar{r} = 0.168$ , KFI = [-0.092, 0.407],  $p = 0.20$ ).

#### 4.4.1.11 Aggressivität, Feindseligkeit

Aufgrund der Studie von Ananth et al. (vgl. Kap. 4.4.1.10) ergab sich kein signifikanter Zusammenhang zwischen Aggressivität und Feindseligkeit mit einem Rezidiv unter Lithium ( $\bar{r} = -0.092$ , KFI = [-0.340, 0.168],  $p = 0.49$ ). Weitere Prophylaxestudien zum Zusammenhang zwischen Aggressivität und Feindseligkeit und Lithiumresponse wurden nicht lokalisiert.

#### 4.4.1.12 Akustische Halluzinationen

Aufgrund der Studie von Ananth et al. (vgl. Kap. 4.4.1.10) ergab sich kein signifikanter Zusammenhang zwischen akustischen Halluzinationen und einem Rezidiv unter Lithium ( $\bar{r} = 0.004$ , KFI = [-0.252, 0.260],  $p = 0.98$ ). Weitere Prophylaxestudien zum Zusammenhang zwischen akustischen Halluzinationen und Lithiumresponse wurden nicht lokalisiert.

#### 4.4.1.13 Denkstörungen

Aufgrund der Studie von Ananth et al. (vgl. Kap. 4.4.1.10) ergab sich kein signifikanter Zusammenhang zwischen Denkstörungen und einem Rezidiv unter Lithium ( $\bar{r} = -0.168$ , KFI = [-0.407, 0.092],  $p = 0.20$ ). Weitere Prophylaxestudien zum Zusammenhang zwischen Denkstörungen und Lithiumresponse wurden nicht lokalisiert.

#### 4.4.1.14 Ängstlichkeit

Aufgrund der Studie von Ananth et al. (vgl. Kap. 4.4.1.10) ergab sich kein Zusammenhang zwischen Ängstlichkeit und einem Rezidiv unter Lithium ( $\bar{r} = -0.059$ , KFI = [-0.310, 0.201],  $p = 0.66$ ). Weitere Prophylaxestudien zum Zusammenhang zwischen Ängstlichkeit und Lithiumresponse wurden nicht lokalisiert.

#### 4.4.1.15 Depressivität, Verlangsamung

Depressivität und Verlangsamung stellen gemäß der Studie von Ananth et al. (vgl. Kap. 4.4.1.10) einen Risikofaktor für ein Rezidiv unter Lithium dar ( $\bar{r} = -0.307$ , KFI = [-0.522, -0.055],  $p = 0.02$ ). Weitere Prophylaxestudien zum Zusammenhang zwischen Denkstörungen und Lithiumresponse wurden nicht lokalisiert. Falls nur wenige oder gar nur eine Studie zu einem Zusammenhang publiziert ist und in dieser Studie ein signifikantes Ergebnis berichtet wird, ist prinzipiell immer mit einem Publikationsbias zu rechnen – etwa, weil signifikante Ergebnisse leichter publizierbar sind. Da im konkreten Fall keine empirisch belegbaren Anhaltspunkte für einen Bias vorliegen, wird Depressivität und Verlangsamung im folgenden als Risikofaktor für einen ungünstigen Verlauf unter Lithium betrachtet.

#### 4.4.1.16 Schwere der Indexepisode

Zum Einfluss der Schwere der Indexepisode auf den Langzeitverlauf unter Lithium wurden 4 Primärstudien lokalisiert. In zwei dieser Studien (Aronoff et al. 1970, Small et al. 1991) wurde lediglich angegeben, daß kein Zusammenhang zwischen Schwere der Indexepisode und der Lithiumresponse besteht. Die beiden anderen Studien zeigen sehr unterschiedliche Ergebnisse. Während Yazici et al. (1999) einen signifikant negativen Einfluss der Schwere der Indexepisode berichten, zeigt sich in der Studie von Mander (1986) allenfalls die gegenläufige Tendenz. Dementsprechend zeigt der Test auf Heterogenität der Effektstärken eine deutliche Verletzung der Homogenitätsannahme an ( $\chi^2 = 12.9$ ,  $p = 0.00002$ ). Bei einer quantitativen Zusammenfassung der insgesamt  $n = 189$  Patienten aus den beiden Studien mit quantifizierbaren Ergebnissen nach dem random effects model zeigt sich kein signifikanter Einfluss des Schweregrads der Indexepisode auf den weiteren Verlauf ( $\bar{r} = -0.084$ ,  $KFI = [-0.379, 0.226]$ ,  $p = 0.60$ , vgl. Abb. 19). Aufgrund der extrem starken Heterogenität ergibt sich bei der Zusammenfassung aller Studien nach dem random effects model (DerSimonian-Laird) ein breiteres Konfidenzintervall als in den einzelnen Studien. Das Ergebnis der Metaanalyse ist aufgrund der divergierenden Ergebnisse in den Primärstudien nur mit Vorsicht interpretierbar und weist primär auf eine unklare Forschungslage hin.

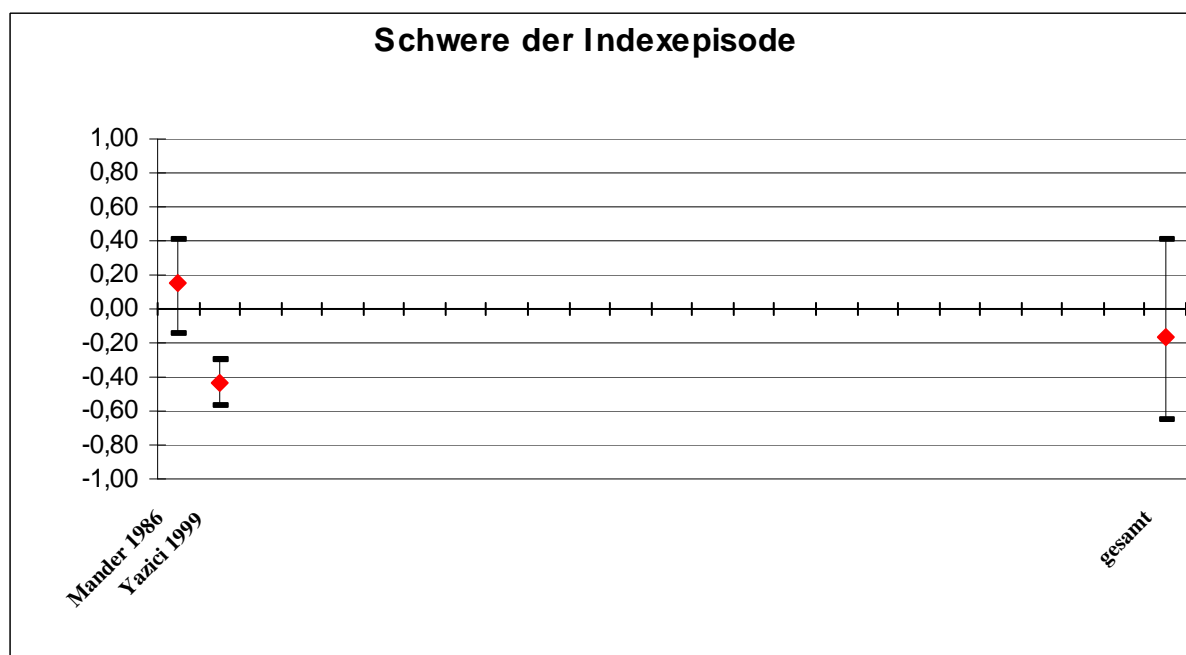


Abbildung 19. Korrelation  $r$  zwischen der Schwere der Indexepisode und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.1.17 Länge der Indexepisode

Der Zusammenhang zwischen Länge der Indexepisode und Lithiumresponse wurde in 3 Primärstudien untersucht. In einer dieser Studien (Aronoff et al. 1970) wurde allerdings nur angegeben, daß kein Zusammenhang zwischen diesen beiden Variablen besteht. Bei einer quantitativen Zusammenfassung (random effects model) der insgesamt  $n = 388$  Patienten aus den beiden anderen Studien zeigt sich kein signifikanter Einfluss der Länge der Indexepisode auf den weiteren Verlauf ( $\bar{r} = -0.173$ , KFI =  $[-0.385, 0.056]$ ,  $p = 0.14$ , vgl. Abb. 20). Die empirische Heterogenität der Effektstärken zwischen den Studien ist statistisch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 2.44$ ,  $p = 0.12$ ).

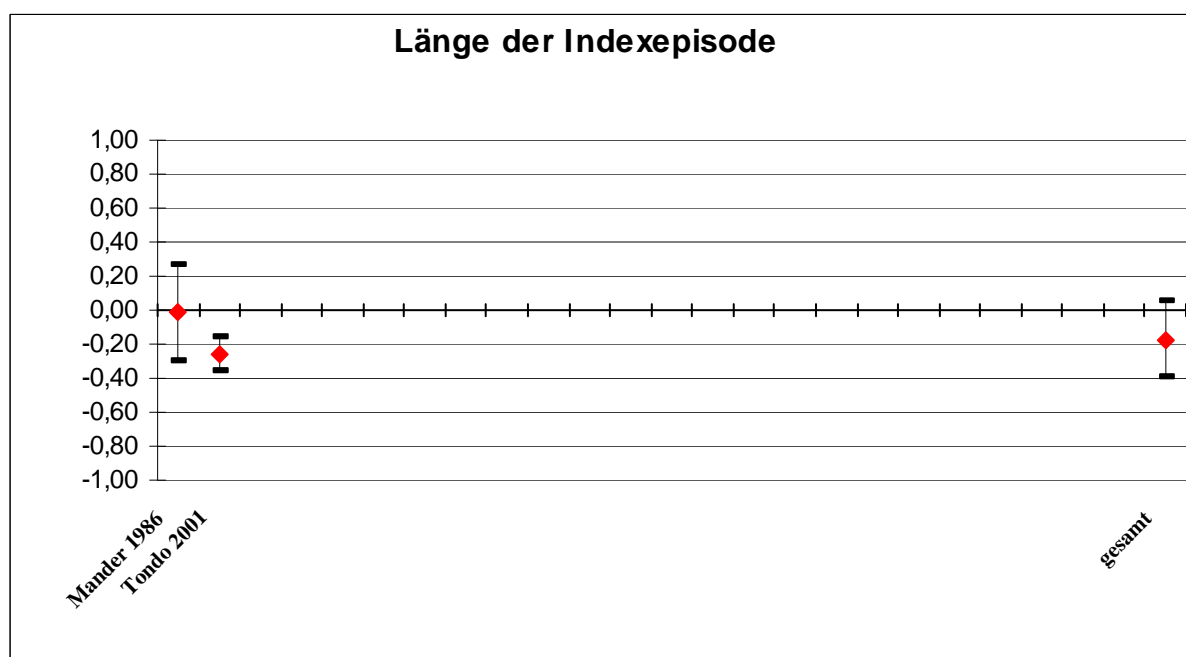


Abbildung 20. Korrelation  $r$  zwischen der Länge der Indexepisode und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.1.18 Vorliegen einer sekundären affektiven Erkrankung

Offenbar die einzige den Einschlusskriterien genügende Untersuchung zur Prädiktion der Lithiumresponse aufgrund einer sekundären (etwa organisch bedingten) affektiven Erkrankung, ist die Studie von Kocsis et al. (1979). In dieser Studie zeigte sich kein Zusammenhang zwischen der Diagnose einer sekundären affektiven Erkrankung und dem Ansprechen auf eine Lithiumprophylaxe ( $r = -0.007$ , KFI = [-0.259, 0.245],  $p = 0.96$ ).

## 4.4.2 Krankheitsanamnese

### 4.4.2.1 Anzahl bisheriger Krankheitsphasen

Insgesamt wurden 15 Primärstudien mit insgesamt 1204 Patienten zum Zusammenhang zwischen der Anzahl bisheriger affektiver Krankheitsphasen und dem darauffolgenden Ansprechen auf eine Lithiumprophylaxe lokalisiert. Der Test auf Heterogenität der Effektstärken zeigt eine deutliche Verletzung der Homogenitätsannahme an ( $\chi^2 = 47.5$ ,  $p = 0.00002$ ). Aus Abb. 21 wird deutlich, daß diese Heterogenität zu einem Großteil durch die Studie von Yaziki et al (1999) bedingt ist. Ein Vergleich dieser Studie mit den anderen, in sich relativ homogenen Studien ergab allerdings keinen eindeutig ersichtlichen Grund, wie es zu dieser starken Abweichung kam. Auf (die sicher zu erwägende) Entfernung dieser Studie aus der Metaanalyse wurde daher verzichtet. Bei einer quantitativen Zusammenfassung aller Studien (random effects model) zeigt sich kein signifikanter Zusammenhang zwischen der Anzahl bisheriger Krankheitsphasen und der Response ( $\bar{r} = -0.074$ ,  $KFI = [-0.189, 0.043]$ ,  $p = 0.22$ ). Der Funnelplot (Abb. 22) zeigt wieder den Ausreißer an (die Studie von Yaziki et al. 1999), gibt aber keinen Anhalt für einen Bias aufgrund nicht publizierter Studien („Schubladenbias“).

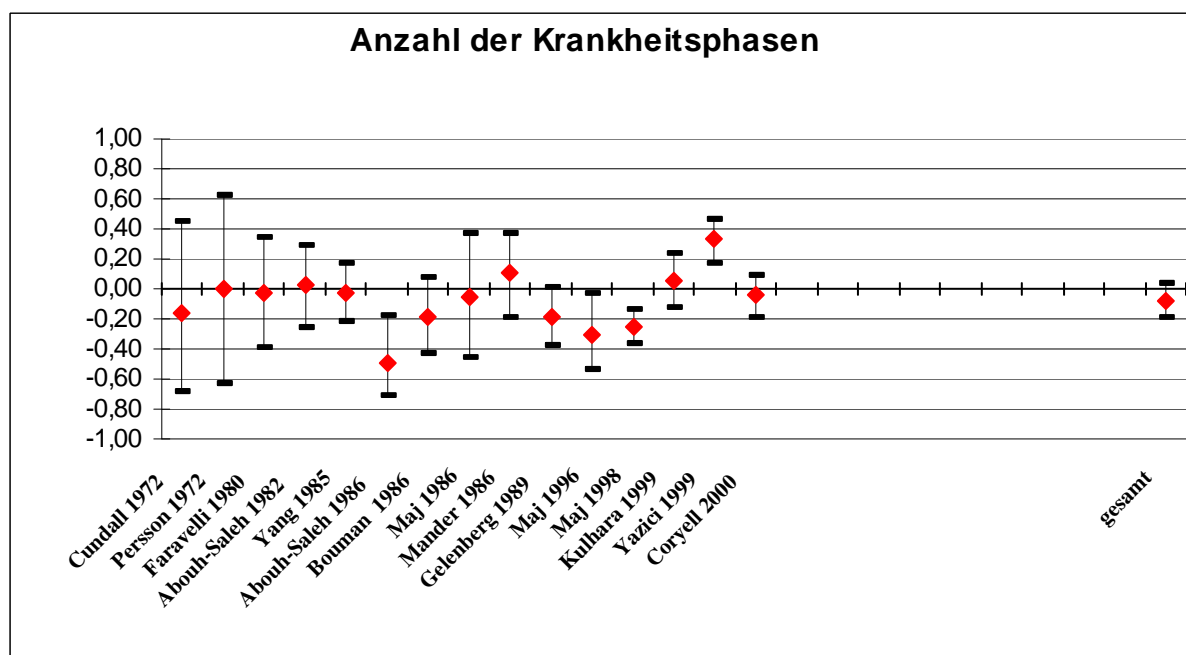


Abbildung 21. Korrelation  $r$  zwischen der Anzahl der Krankheitsphasen und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).



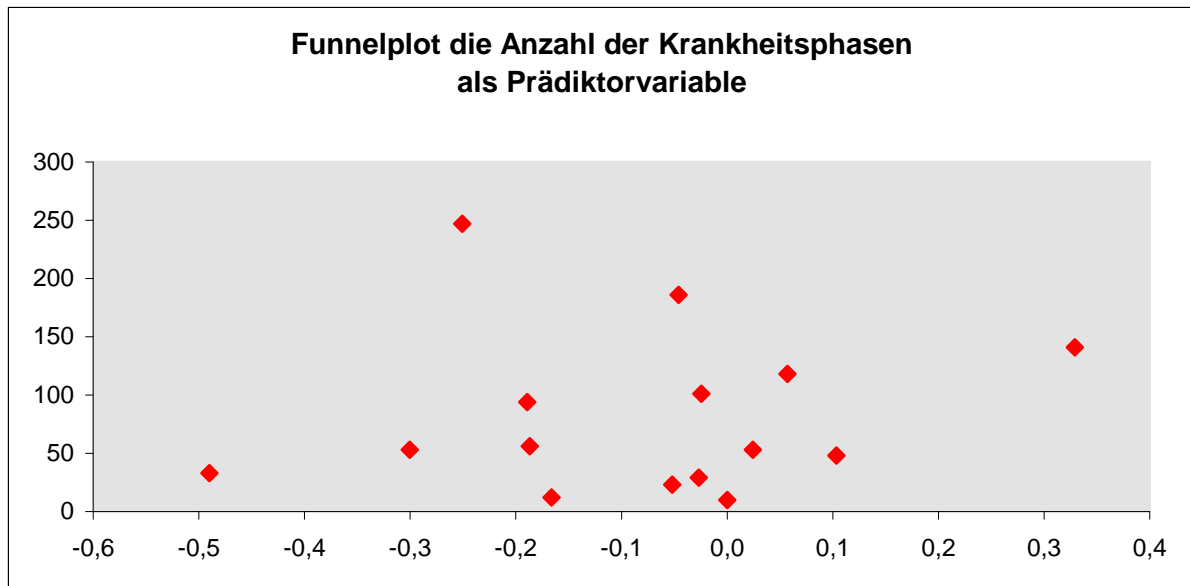


Abbildung 22. Funnelplot für die Anzahl der Krankheitsphasen: Effektstärke  $r$  in Abhängigkeit von der Fallzahl  $n$  in den einzelnen Stichproben.

#### 4.4.2.2 Bisherige Anzahl stationärer Krankheitsphasen

Insgesamt gingen die Daten von 677 Patienten aus 4 Studien in die Metaanalyse zum Zusammenhang zwischen der bisherigen Anzahl stationärer Krankheitsphasen und dem Ansprechen auf eine Lithiumprophylaxe ein. In jeder einzelnen dieser Studien, und noch mehr bei einer metaanalytischen Zusammenfassung der Primärergebnisse erwies sich die bisherige Anzahl stationärer Episoden als Risikofaktor für den weiteren Verlauf unter Lithium ( $\bar{r} = -0.347$ , KFI = [-0.417, -0.272],  $p < 0.0000000001$ , vgl. Abb. 23). Die empirische Heterogenität der Effektstärken zwischen den Studien ist statistisch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 3.39$ ,  $p = 0.34$ ). Das Ergebnis der Metaanalyse spricht eindeutig dafür, daß die Anzahl bisheriger stationärer Episoden ein klinisch und statistisch bedeutsamer Risikofaktor für ein schlechtes Ansprechen auf eine Lithiumprophylaxe ist.

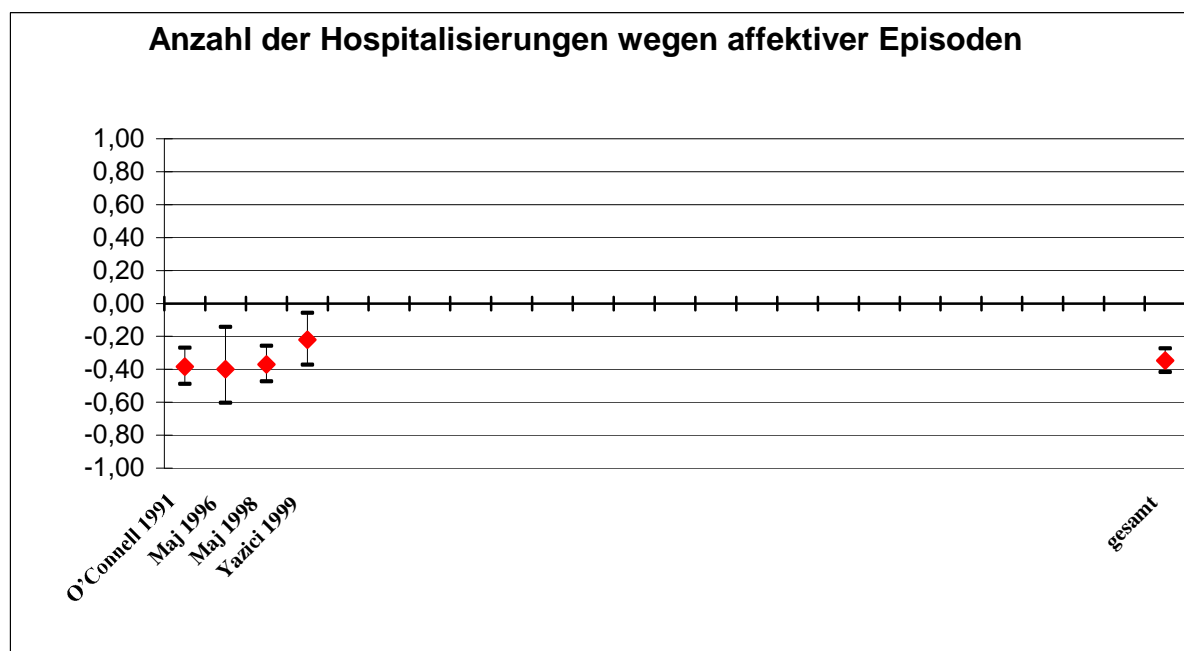


Abbildung 23. Korrelation  $r$  zwischen der Anzahl bisheriger Hospitalisierungen aufgrund affektiver Episoden und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.2.3 Anzahl depressiver Krankheitsepisoden

Zum Zusammenhang zwischen der Anzahl depressiver Krankheitsepisoden und dem Verlauf unter einer Lithiumprophylaxe konnten zwei Studien mit insgesamt  $n = 147$  Patienten lokalisiert werden (Faravelli et al. 1980, Kulhara et al. 1999). Bei einer quantitativen Zusammenfassung der beiden Studien ergab sich kein klarer Zusammenhang zwischen der Anzahl depressiver Episoden in der Krankengeschichte und einem guten Ansprechen auf Lithium ( $\bar{r} = -0.050$ ,  $KFI = [-0.395, 0.307]$ ,  $p = 0.79$ , vgl. Abb. 24). Die empirische Heterogenität der Effektstärken zwischen den Studien ist statistisch noch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 3.11$ ,  $p = 0.08$ ).

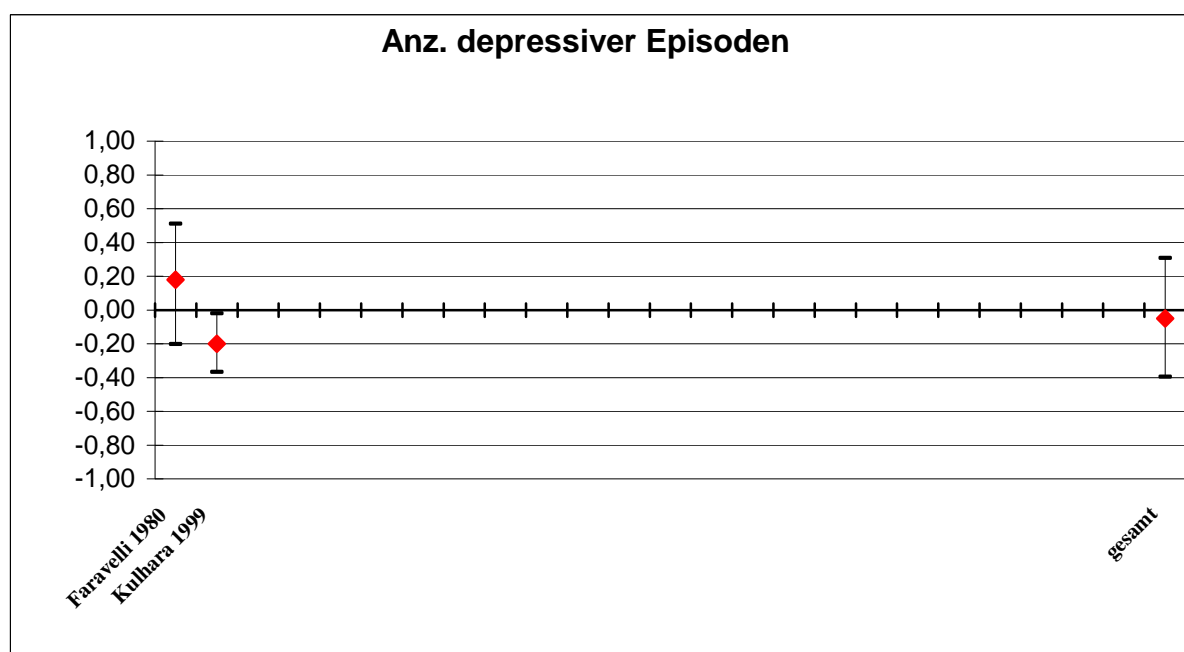


Abbildung 24. Korrelation  $r$  zwischen der Anzahl depressiver Episoden in der Vorgeschichte und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.2.4 Anzahl manischer Krankheitsepisoden

Auch zum Zusammenhang zwischen der Anzahl manischer Krankheitsepisoden und dem Verlauf unter einer Lithiumprophylaxe liegen die beiden selben Studien vor wie zur Prädiktion aufgrund der Anzahl depressiver Episoden – Faravelli et al. (1980) und Kulhara et al. (1999). Im Ergebnis ergab eine Zusammenfassung der beiden Studien mit insgesamt  $n = 147$  Patienten keinen Zusammenhang zwischen der Anzahl manischer Episoden und der Lithiumresponse ( $\bar{r} = -0.022$ , KFI =  $[-0.185, 0.140]$ ,  $p = 0.80$ , vgl. Abb. 25). Der Test auf Heterogenität der Effektstärken gibt keinen Anhalt für die Verletzung der Homogenitätsannahme ( $\chi^2 = 0.673$ ,  $p = 0.41$ ).

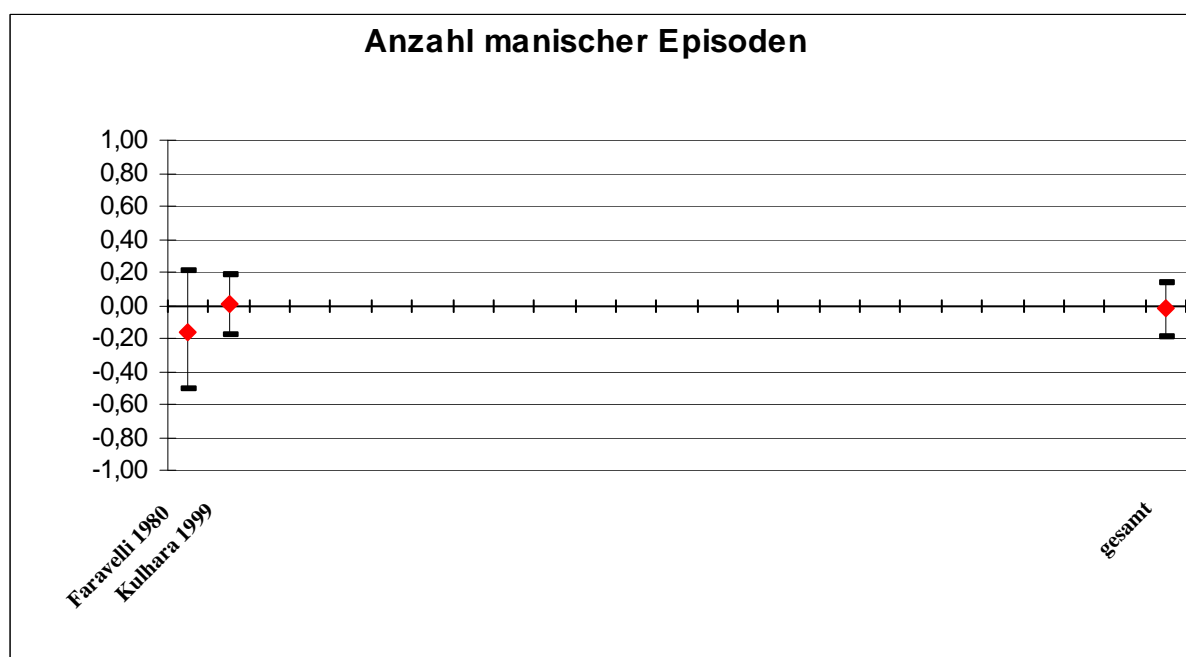


Abbildung 25. Korrelation  $r$  zwischen der Anzahl manischer Episoden in der Vorgeschichte und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.2.5 Hohe Phasenfrequenz und rapid cycling

Zum Zusammenhang zwischen der bisherigen Frequenz affektiver Episoden und dem Ansprechen auf eine Lithiumprophylaxe wurden 8 Primärstudien mit insgesamt 908 Patienten lokalisiert. Der Test auf Heterogenität der Effektstärken zeigt eine deutliche Verletzung der Homogenitätsannahme an ( $\chi^2 = 40.9$ ,  $p = 0.0000003$ ). Dies dürfte damit zusammenhängen, daß die Evaluation eines guten vs. schlechten Ansprechens auf Lithium in 3 der 8 Studien (Kocsis et al. 1979, Yaziki et al. 1999 und Tondo et al. 2001) auf prä-post Vergleichen beruht. Numerisch wäre damit für Patienten mit einer hohen Phasenfrequenz eine größere Chance für eine Verbesserung des schlechten Ausgangsniveaus (der hohen Phasenfrequenz vor Studienbeginn) gegeben. Um diese Hypothese zu prüfen, wurde der ursprünglich einheitliche Forestplot (Abb. 26) für die beiden Schichten (outcome-Messung allein auf Grundlage der Beobachtungsperiode vs. outcome-Messung auf Grundlage von prä-post Vergleichen) getrennt dargestellt (Abb. 27 und 28). Das Ergebnis dieser Aufteilung der Studien entspricht exakt der Hypothesen, daß die Art der Therapieevaluation für die Prädiktion aufgrund der Phasenfrequenz von großer Bedeutung ist. Demnach kann es bei prä-post Vergleichen zu einem kontraintuitiven Effekt kommen – eine hohe Phasenfrequenz in der Vorgeschichte begünstigt möglicherweise sogar eine relative Reduktion dieser hohen Phasenfrequenz unter Lithium.

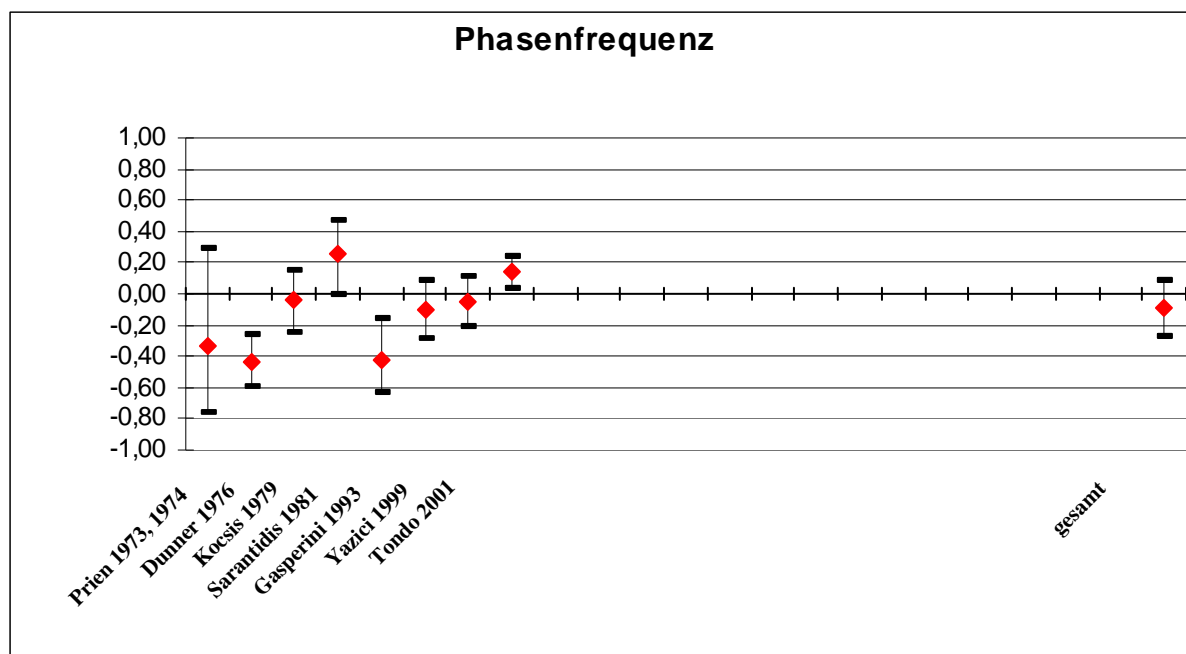


Abbildung 26. Korrelation  $r$  zwischen der bisherigen Phasenfrequenz und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

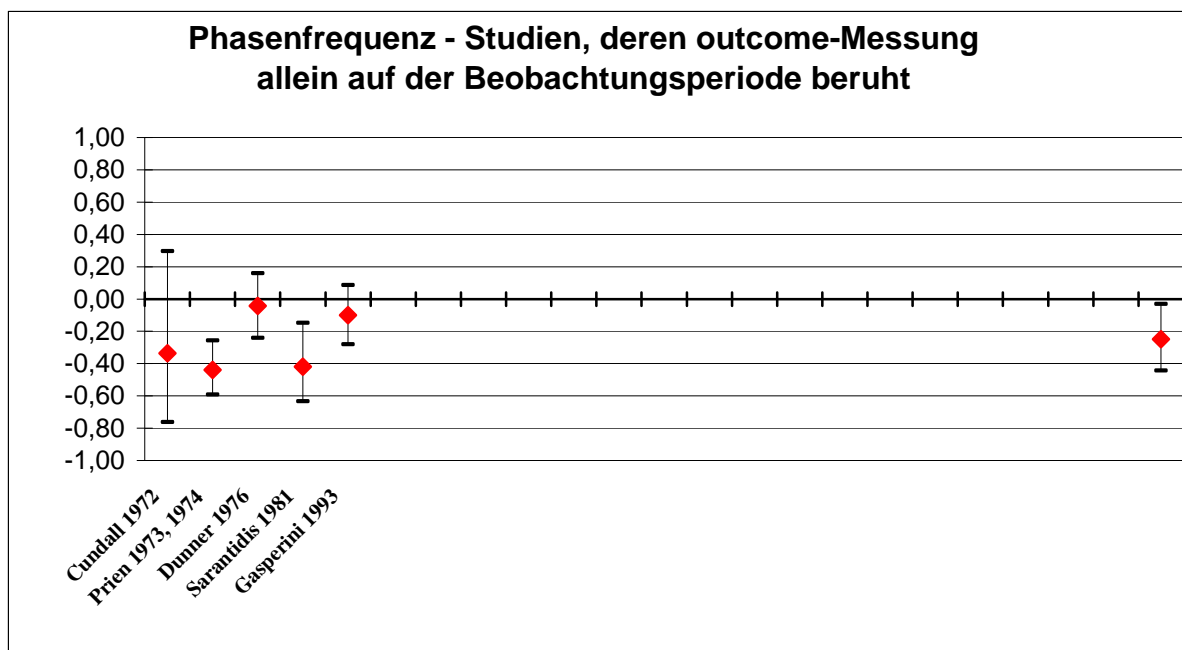


Abbildung 27. Korrelation  $r$  zwischen der bisherigen Phasenfrequenz und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz). Die Graphik bezieht sich ausschließlich auf Studien, deren outcome-Messung auf der Beobachtungszeit beruht.

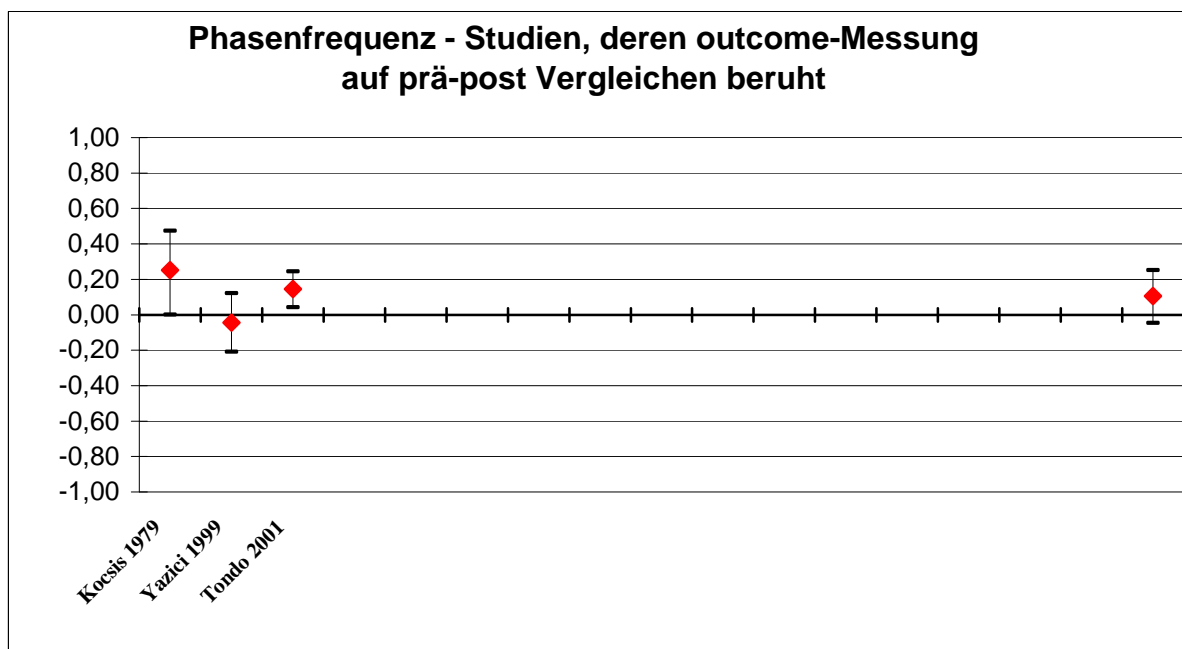


Abbildung 28. Korrelation  $r$  zwischen der bisherigen Phasenfrequenz und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz). Die Graphik bezieht sich ausschließlich auf Studien, deren outcome-Messung auf prä-post Vergleichen beruht.

Da es in der vorliegenden Arbeit primär um die Vorhersage des absoluten Verlaufs unter Lithium (ohne Einbeziehung des bisherigen Verlaufs) geht, fokussiert die metaanalytische Auswertung auf die 5 Studien, deren outcome-Messung allein auf der Beobachtungsperiode beruht. Bei einer quantitativen Zusammenfassung dieser Studien nach dem random effects model zeigt sich ein signifikanter Zusammenhang zwischen einer hohen Phasenfrequenz in der Vorgeschichte und einem ungünstigen Verlauf unter einer Lithiumprophylaxe ( $\bar{r} = -0.284$ , KFI = [-0.442, -0.032],  $p = 0.03$ ). Da keine empirisch belegbaren Anhaltspunkte für einen Bias vorliegen, wird eine hohe Phasenfrequenz im folgenden als Risikofaktor für ein affektives Rezidiv unter einer Lithiumprophylaxe betrachtet.

Rapid cycling stellt eine Sonderform einer hohen Phasenfrequenz dar (von rapid cycling spricht man, falls mindestens 4 Phasen pro Jahr vorliegen). Da rapid cycling eine in der Prädiktionsforschung viel diskutierte Variable darstellt, werden die Ergebnisse für diese Variable gesondert aufgeführt. Insgesamt konnten 7 Primärstudien lokalisiert werden, die eine Verbindung zwischen rapid cycling und Lithiumresponse herstellen. Die Studien von Baldessarini et al. (2000) und von Tondo et al. (2001) beziehen sich offensichtlich auf dasselbe Patientenkollektiv und deshalb wurde nur eine dieser beiden Studien in die Metaanalyse eingeschlossen, so daß insgesamt  $n = 699$  Patienten aus 6 Studien in die metaanalytische Auswertung eingehen.

Der Test auf Heterogenität der Effektstärken weist auf eine Verletzung der Homogenitätsannahme hin ( $\chi^2 = 16.3$ ,  $p = 0.003$ ). Dies könnte wieder mit der unterschiedlichen Evaluation des Behandlungserfolges (prä-post Vergleiche vs. Evaluation allein aufgrund des Beobachtungszeitraums, s.o.) zusammenhängen. Allerdings sind die Ergebnisse insgesamt einheitlicher als im letzten Abschnitt – dies hängt damit zusammen, daß rapid cycling unabhängig von der Art der outcome-Evaluation ein Negativprädiktor in den meisten Primärstudien war. Bei einer quantitativen Zusammenfassung der Effektstärken nach dem random effects model ergab sich ein klinisch wie statistisch signifikanter Zusammenhang zwischen dem Vorliegen eines rapid cycling und einem ungünstigen Verlauf unter Lithium ( $\bar{r} = -0.204$  (KFI = [-0.266, -0.141],  $p = 0.0000000005$ , vgl. Abb. 29). Hierdurch wird das Hauptergebnis dieses Kapitels, wonach eine hohe Phasenfrequenz prognostisch ungünstig für das Ansprechen auf eine Lithiumprophylaxe ist, noch einmal bekräftigt.

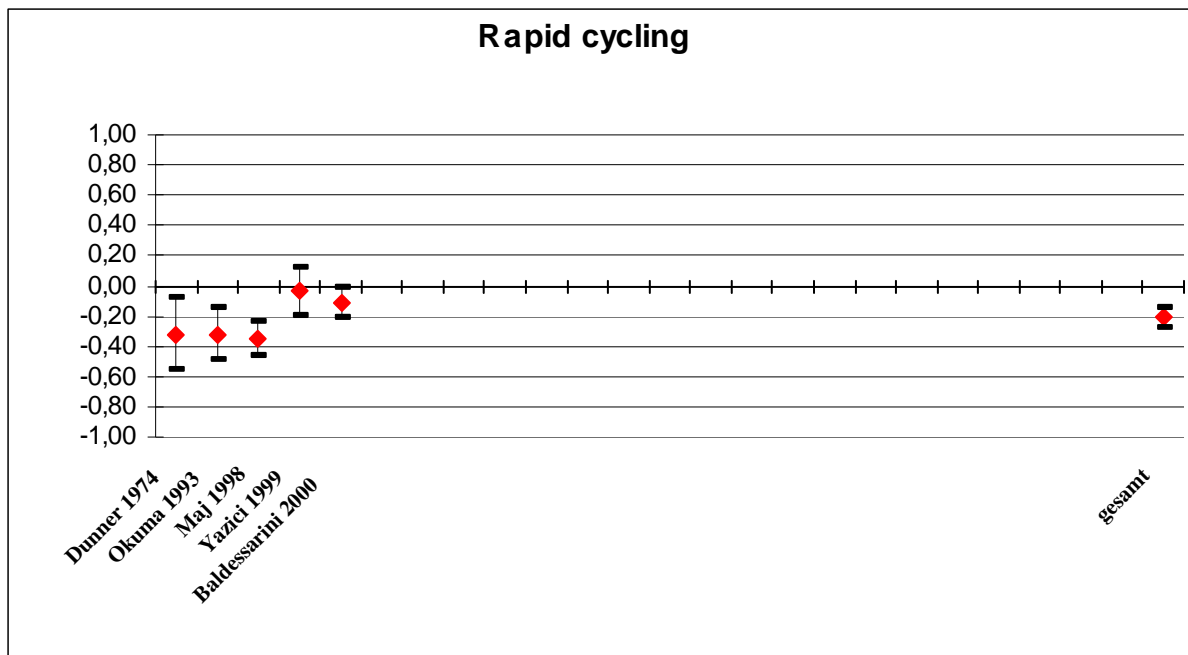


Abbildung 29. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen eines rapid-cyclings und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).



#### 4.4.2.6 Krankheitsdauer

Insgesamt wurden 8 Studien mit insgesamt 905 Patienten lokalisiert, für die sich der Zusammenhang zwischen der Zeit seit der Erstdiagnose oder der Zeit seit dem ersten psychiatrischen Kontakt („Krankheitsdauer“) und dem Ansprechen auf eine Lithiumprophylaxe quantifizieren lässt. Eine neunte Studie (Tondo et al. 2001) wurde nicht eingeschlossen, da sie sich offensichtlich weitgehend auf dasselbe Patientenkollektiv bezieht wie die Studie von Tondo et al. (1998).

Der Test auf Heterogenität der Effektstärken für die eingeschlossenen Studien zeigt eine deutliche Heterogenität der Primärstudien an ( $\chi^2 = 46.3$ ,  $p = 0.0000005$ ). Hierfür waren keine klar ersichtlichen inhaltlichen Gründe eruierbar. Bei einer Aggregation der Effektstärken nach dem random-effects-model (DerSimonian-Laird) ergibt sich eine gemittelte Effektstärke von  $\bar{r} = -0.160$  (KFI =  $[-0.298, -0.015]$ ,  $p = 0.03$ ). Eine lange Krankheitsdauer erweist sich demnach als prognostisch ungünstig für den Verlauf unter Lithium (vgl. Abb. 30). Auch im Funnelplot (Abb. 31) wird die Heterogenität der Primärstudien sichtbar. Ein Hinweis auf nicht publizierte Studien mit geringer absoluter Effektstärke ergibt sich allerdings nicht. Aufgrund der fehlenden Evidenz für einen Bias wird daher eine lange Krankheitsdauer im folgenden als Risikofaktor für einen ungünstigen Verlauf unter Lithium betrachtet.

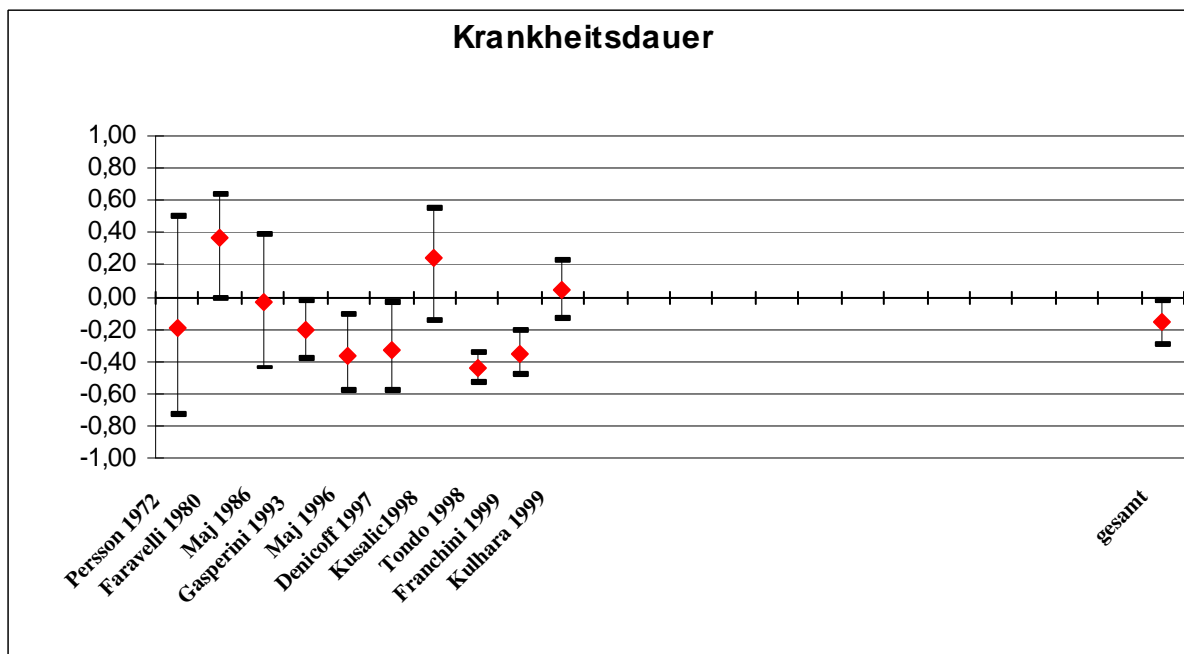


Abbildung 30. Korrelation  $r$  zwischen der bisherigen Krankheitsdauer und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

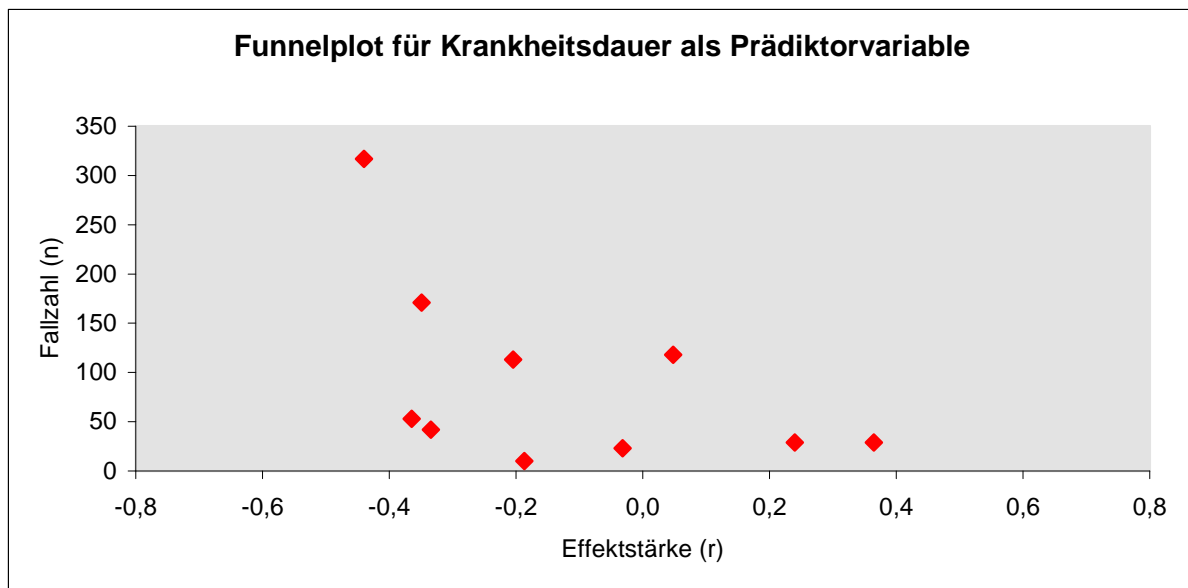


Abbildung 31. Funnelplot für die Krankheitsdauer als Prädiktorvariable: Effektstärke  $r$  in Abhängigkeit von der Fallzahl  $n$  in den einzelnen Stichproben.

#### 4.4.2.7 Alter bei Ersterkrankung

Das Alter bei Ersterkrankung wurde in 10 Primärstudien mit insgesamt 1138 Patienten in einen quantifizierbaren Zusammenhang mit der Lithiumresponse gebracht. Die empirische Heterogenität der Effektstärken aus diesen 10 Studien ist statistisch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 12.6$ ,  $p = 0.18$ ). Bei einer quantitativen Zusammenfassung der Effektstärken nach dem random-effects-model erweist sich ein hohes Ersterkrankungsalter als prognostisch günstig für den Verlauf unter Lithium ( $\bar{r} = 0.114$  (KFI = [0.039, 0.188],  $p = 0.003$ , vgl. Abb. 32). Aufgrund des Funnelplots (Abb. 33) ergeben sich keine Hinweise für einen Publikationsbias. In 4 weiteren Studien wird von einem nicht signifikanten, allerdings nicht quantifizierbaren Einfluss des Ersterkrankungsalters auf die Lithiumresponse berichtet (Prien et al. 1973 und 1974, Abouh-Saleh & Coppen 1986, O'Connell et al. 1991, Denicoff et al. 1997). Die fehlende Möglichkeit, die Ergebnisse aus diesen 4 Studien zu quantifizieren könnte tendenziell zu einer Überschätzung der gemittelten Effektstärke im Rahmen dieser Metaanalyse führen. Aufgrund des klaren Ergebnisses bei der metaanalytischen Zusammenfassung der 10 Studien, für die sich die Effektstärke quantifizieren lässt, wird im folgenden jedoch davon ausgegangen, daß ein hohes Ersterkrankungsalter positiv mit einem günstigen Verlauf unter Lithium korreliert ist.

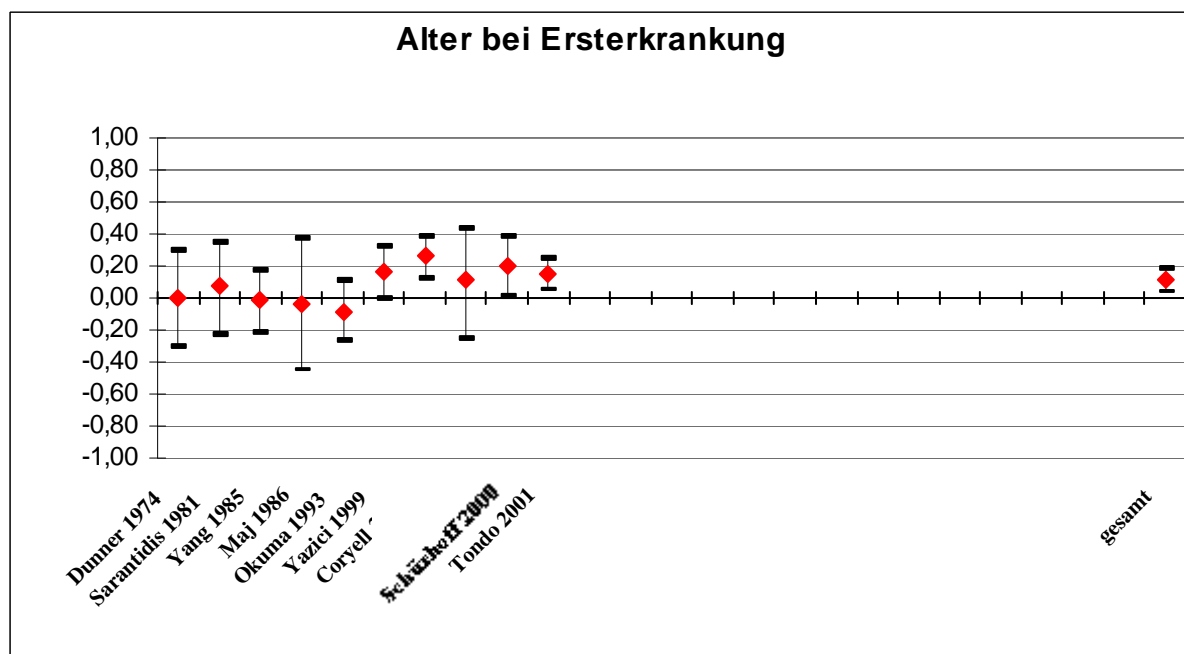


Abbildung 32. Korrelation  $r$  zwischen dem Alter bei Ersterkrankung und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

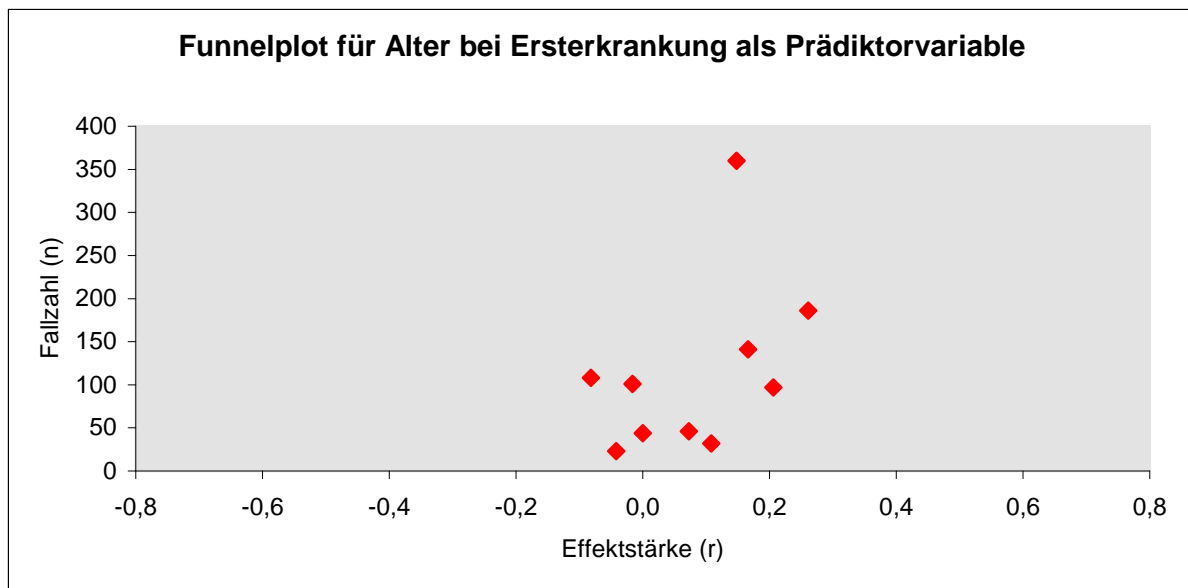


Abbildung 33. Funnelplot für das Alter bei Ersterkrankung als Prädiktorvariable: Effektstärke  $r$  in Abhängigkeit von der Fallzahl  $n$  in den einzelnen Stichproben.

#### 4.4.2.8 Alter bei Erstbehandlung

Zum Zusammenhang zwischen dem Alter bei Erstbehandlung und dem Ansprechen auf eine Lithiumprophylaxe wurden 3 Studien mit insgesamt 342 Patienten lokalisiert. In zwei dieser Studien war dieser Zusammenhang direkt als Effektstärke quantifizierbar (Maj et al. 1996, Maj et al. 1998), in der dritten Studie (Denicoff et al. 1997) wurde nur berichtet, daß eine signifikant negative Beziehung zwischen einem hohen Ersterkrankungsalter und einer guten Lithiumresponse bestand ( $p < 0.01$ ). Da diese Aussage Information über Ausmaß und Richtung des Zusammenhangs zwischen Alter bei Erstbehandlung und Lithiumresponse enthält, wäre ein Weglassen dieser Information mit einem vermutlich größeren Fehler verbunden gewesen als eine Schätzung des Zusammenhangs aus den angegebenen Informationen. Diese Schätzung der Effektstärke  $r$  wurde konservativ vorgenommen, d.h. für  $p$  wurde die Signifikanzschwelle 0.01 angesetzt, hieraus wurde dann über die der Stichprobengröße entsprechende  $t$ -Verteilung ein  $t$ -Wert bestimmt und schließlich wurde dieser  $t$ -Wert in eine Effektstärkeschätzung  $r$  übersetzt (zur Berechnung von  $r$  auf Grundlage eines  $p$ -Wertes vgl. Kap. 4.1.3).

Der Test auf Heterogenität der Effektstärken weist auf eine Verletzung der Homogenitätsannahme hin ( $\chi^2 = 7.05$ ,  $p = 0.03$ ), ohne daß hierfür inhaltliche Gründe offensichtlich gewesen wären. Bei einer quantitativen Zusammenfassung der Effektstärken nach dem random effects model ergab sich kein signifikanter Zusammenhang zwischen dem Alter bei Erstbehandlung und dem Verlauf unter Lithium ( $\bar{r} = -0.061$  (KFI = [-0.305, 0.190],  $p = 0.63$ , vgl. Abb. 34).

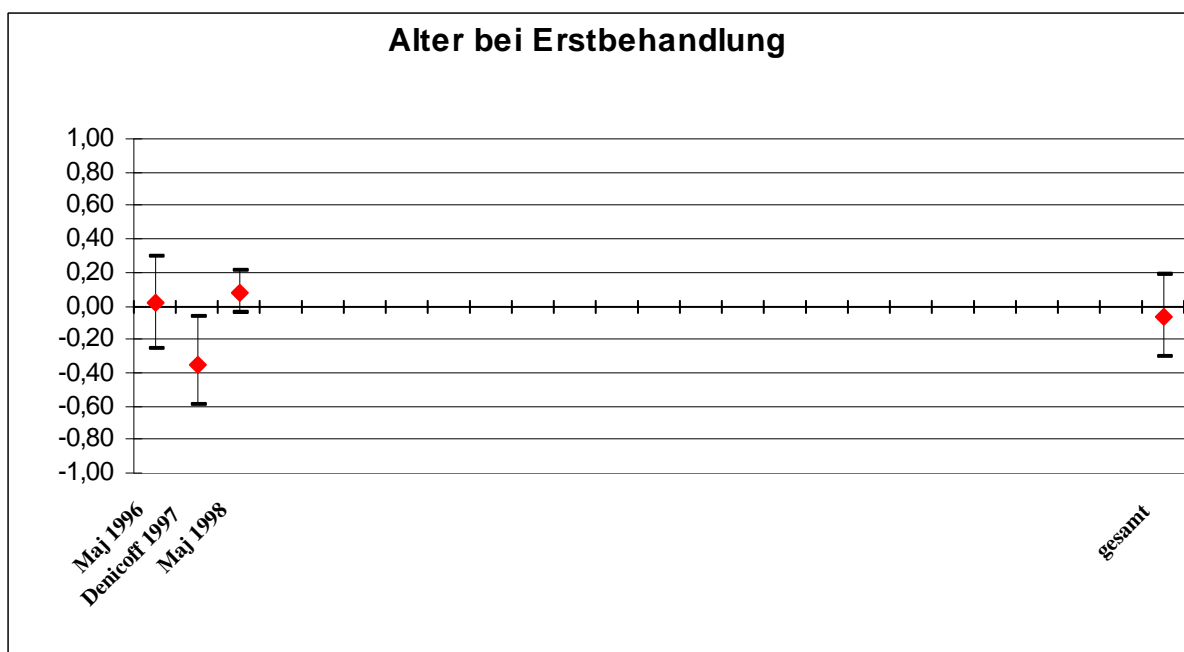


Abbildung 34. Korrelation  $r$  zwischen dem Alter bei Erstbehandlung und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.2.9 Verlaufsmuster MDI

Insbesondere seit den Arbeiten von Kukopulos (1980a,b) und Angst (1978) findet das vorherrschende Verlaufsmuster in der Krankengeschichte große Beachtung in der Prädiktionsforschung bipolarer Störungen. Die wichtigsten Komponenten, auf deren Grundlage das Verlaufsmuster beschrieben wird, sind die Manie (M), die Depression (D), das freie Intervall zwischen den Episoden (I), sowie symptomatisch leichte Manien oder Hypomanien (m) und symptomatisch leichte depressive Zustände (d). Diese Komponenten kommen in einer Vielzahl von Verlaufsmustern vor und dementsprechend sind auch viele verschiedene Verlaufsmuster beschrieben und als mögliche Prädiktoren diskutiert worden.

MDI bedeutet, daß der bisherige Verlauf vor allem durch die Episodenfolge Manie – Depression – freies Intervall gekennzeichnet ist. Zum Zusammenhang zwischen dem Verlaufsmuster MDI und dem Ansprechen auf eine Lithiumprophylaxe konnten 7 Studien mit insgesamt 904 Patienten lokalisiert werden. Drei dieser Studien waren ihrer Substanz noch ein weiteres mal publiziert, wurden aber, wie stets in der vorliegenden Arbeit, nur einmal in die Metaanalyse eingeschlossen. Der Test auf Heterogenität der Effektstärken zeigt eine gewisse Heterogenität der 7 eingeschlossenen Primärstudien an ( $\chi^2 = 14.0$ ,  $p = 0.03$ ). Bei der Aggregation der Effektstärken nach dem random effects model ergab sich ein signifikanter Zusammenhang zwischen dem Verlaufsmuster MDI vor Beginn der Behandlung und dem Verlauf unter Lithium ( $\bar{r} = 0.203$  (KFI = [0.092, 0.310],  $p = 0.0004$ , vgl. Abb. 35). Da keine empirisch belegbaren Anhaltspunkte für einen Bias vorliegen, wird das Verlaufsmuster MDI im folgenden als Positivprädiktor für einen günstigen Verlauf unter Lithium betrachtet.

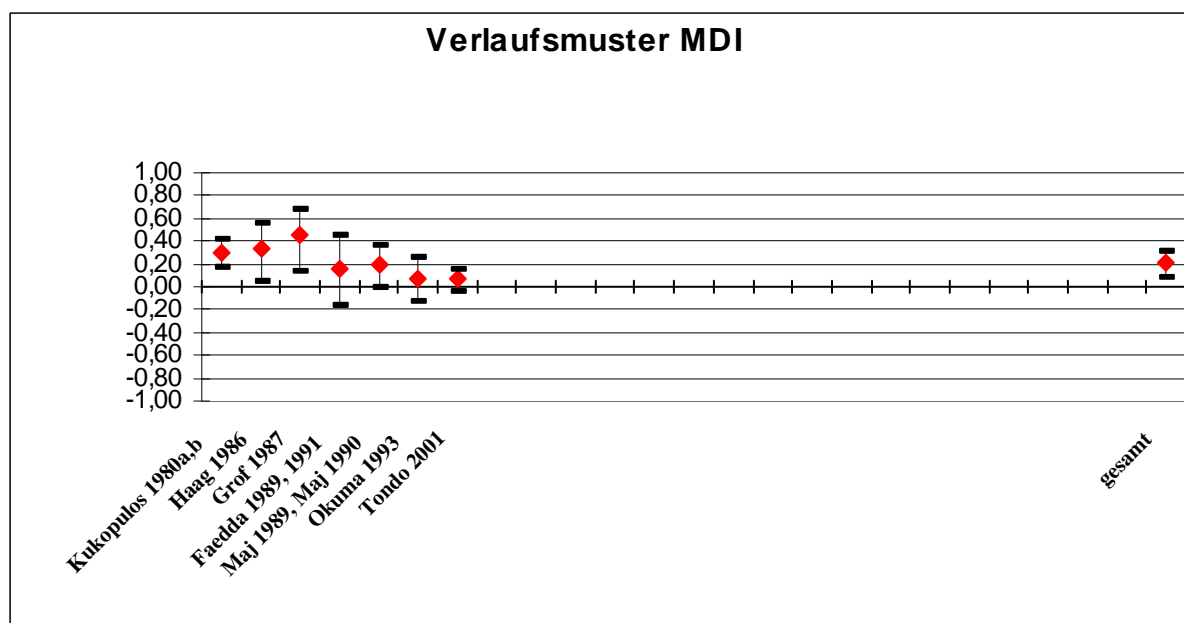


Abbildung 35. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen des Verlaufsmusters MDI und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.2.10 Verlaufsmuster DMI

DMI bedeutet, daß der bisherige Verlauf vor allem durch die Episodenfolge Depression – Manie – freies Intervall gekennzeichnet ist (vgl. Kap. 4.4.2.9). Hierbei ist zu beachten, daß DMI nicht das exakte Gegenteil von MDI ist. Vielmehr gibt es eine Vielzahl weiterer Verlaufsvarianten, die zum Teil in den folgenden Kapiteln besprochen werden.

Das Verlaufsmuster DMI wurde in 8 Studien mit insgesamt 1151 Patienten mit dem Ansprechen auf Lithium in Verbindung gebracht. Auch hier zeigte sich eine gewisse Heterogenität der Primärstudien hinsichtlich der Effektstärken ( $\chi^2 = 15.3$ ,  $p = 0.03$ ). Bei der Aggregation der Effektstärken nach dem random effects model (DerSimonian-Laird) erwies sich das Verlaufsmuster DMI als ein Risikofaktor für einen ungünstigen Verlauf unter einer Lithiumprophylaxe ( $\bar{r} = -0.122$  (KFI = [-0.206, -0.017],  $p = 0.02$ , vgl. Abb. 36). Aufgrund des Funnelplots (Abb. 37) ergibt sich keinerlei Hinweis für einen Publikationsbias. Da auch keine weiteren empirisch belegbaren Anhaltspunkte für einen Bias vorliegen, wird das Verlaufsmuster DMI im folgenden als Risikofaktor betrachtet.

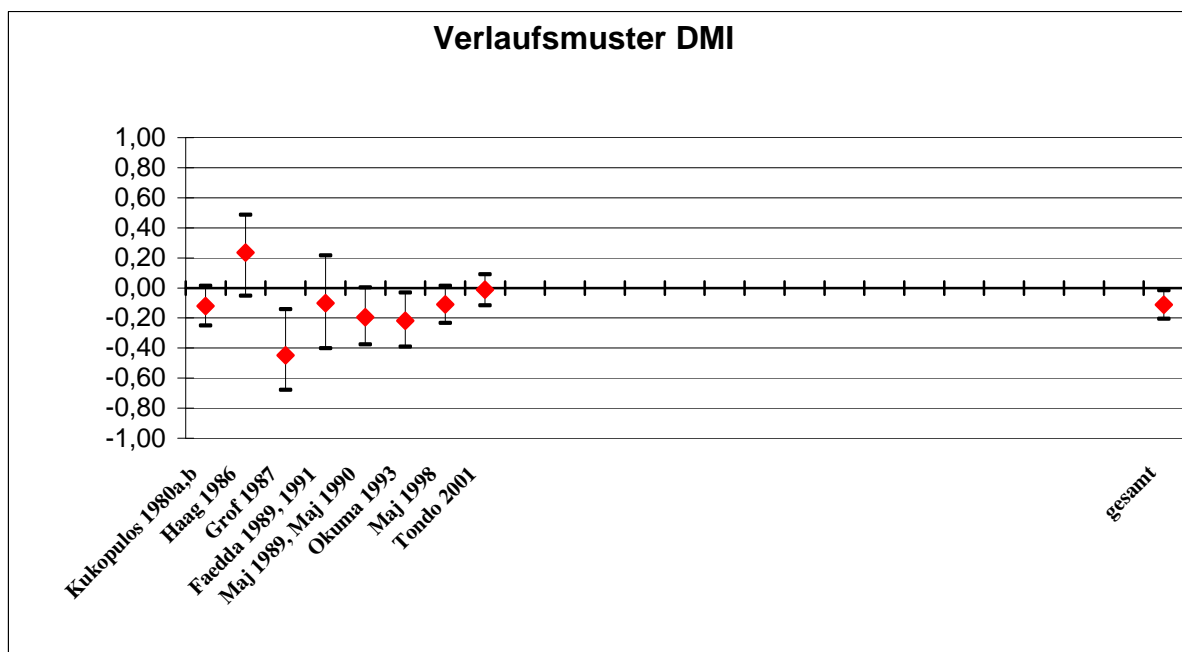


Abbildung 36. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen des Verlaufsmusters DMI und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).



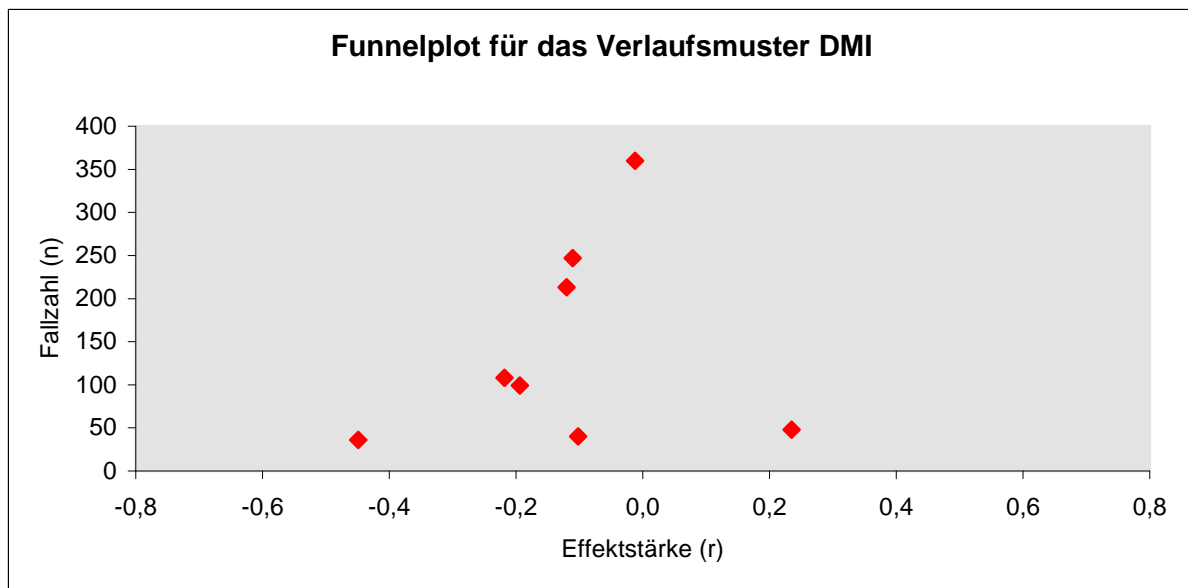


Abbildung 37. Funnelplot für das Verlaufsmuster DMI als Prädiktorvariable: Effektstärke  $r$  in Abhängigkeit von der Fallzahl  $n$  in den einzelnen Stichproben.

#### 4.4.2.11 Verlaufsmuster CC

CC bedeutet, daß der bisherige Verlauf vor allem durch eine Episodenfolge dominiert ist, bei der einzelne Episoden ohne zwischenliegendes freies Intervall ineinander übergehen (CC steht für "continuous cycling", vgl. auch Kap. 4.4.2.9).

Das Verlaufsmuster CC wurde an insgesamt 400 Studienpatienten untersucht (7 Studien, davon 4 Studien, deren Studienpopulationen sich offensichtlich nicht überschneiden). Die empirische Heterogenität der Effektstärken zwischen den eingeschlossenen Studien ist gering und statistisch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 2.53$ ,  $p = 0.47$ ). Im Ergebnis erwies sich ein continuous cycling bei einer quantitativen Zusammenfassung der Studien als ein Risikofaktor für einen ungünstigen Verlauf unter einer Lithiumprophylaxe ( $\bar{r} = -0.213$  (KFI = [-0.306, -0.116],  $p = 0.00002$ , vgl. Abb. 38). Da auch hier keine Anhaltspunkte für einen Bias vorliegen, wird das Verlaufsmuster CC im folgenden als Risikofaktor für eine schlechte Response auf eine Lithiumprophylaxe betrachtet.

In einer der eingeschlossenen Studien (Maj et al. 1989, 1990) werden zwei Typen von continuous cycling unterschieden: CC mit kurzen Zyklen und CC mit langen Zyklen. Aufgrund der starken inhaltlichen Überschneidung mit Phasenfrequenz und rapid cycling wurde diese Unterscheidung in der Metaanalyse nicht übernommen.

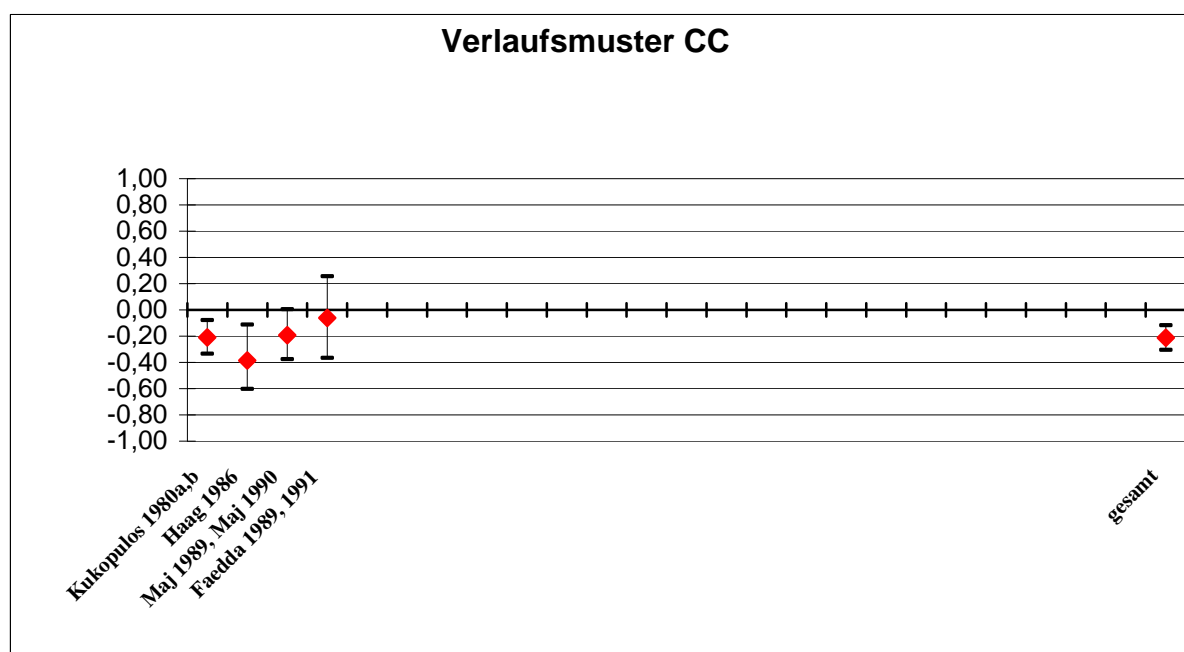


Abbildung 38. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen des Verlaufsmusters CC und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.2.12 Isolierte Krankheitsepisoden

Isolierte Krankheitsepisoden sind Episoden, die an Anfang und Ende durch ein freies Intervall begrenzt werden (z.B. freies Intervall – Manie – freies Intervall). Zum Zusammenhang zwischen dem Vorherrschen dieses Verlauftyps und der Response auf eine Lithiumprophylaxe lag nur eine einzige Primärstudie vor – die Studie von Okuma (1993). Gemäß dieser Studie ist das isolierte Auftreten der Krankheitsepisoden mit einem günstigen Verlauf unter einer Lithiumprophylaxe assoziiert ( $\bar{r} = 0.324$ , KFI = [0.144, 0.483],  $p = 0.0006$ ). Zwar ist – wie immer in solchen Fällen – mit einem gewissen Publikationsbias zu rechnen (vgl. etwa Kap. 4.4.1.15). Da es aber keine empirisch belegbaren Anhaltspunkte für einen Bias gibt, wird das isolierte Auftreten von Krankheitsepisoden im folgenden als Prädiktor für einen günstigen Verlauf unter Lithium betrachtet.

#### 4.4.2.13 Irreguläres Verlaufsmuster

Ein irreguläres Verlaufsmuster bedeutet, daß der bisherige Verlauf keinem eindeutigen Muster folgt, sondern Episoden und freie Intervalle in irregulärer Folge wechseln (vgl. auch Kap. 4.4.2.9).

Ein irreguläres Verlaufsmuster wurde für insgesamt 360 Studienpatienten mit dem Ansprechen auf eine Lithiumprophylaxe in Zusammenhang gebracht. Die empirische Heterogenität der Effektstärken zwischen den insgesamt 3 Studien, die sich auf verschiedene Patientenpopulationen beziehen, ist statistisch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 3.61$ ,  $p = 0.16$ ). Wie in Abb. 39 dargestellt ergab sich bei einer Zusammenfassung der Effektstärken kein statistisch signifikanter Zusammenhang zwischen einem irregulären Verlaufsmuster und einer Lithiumresponse ( $\bar{r} = -0.118$ ,  $KFI = [-0.266, 0.035]$ ,  $p = 0.13$ ).

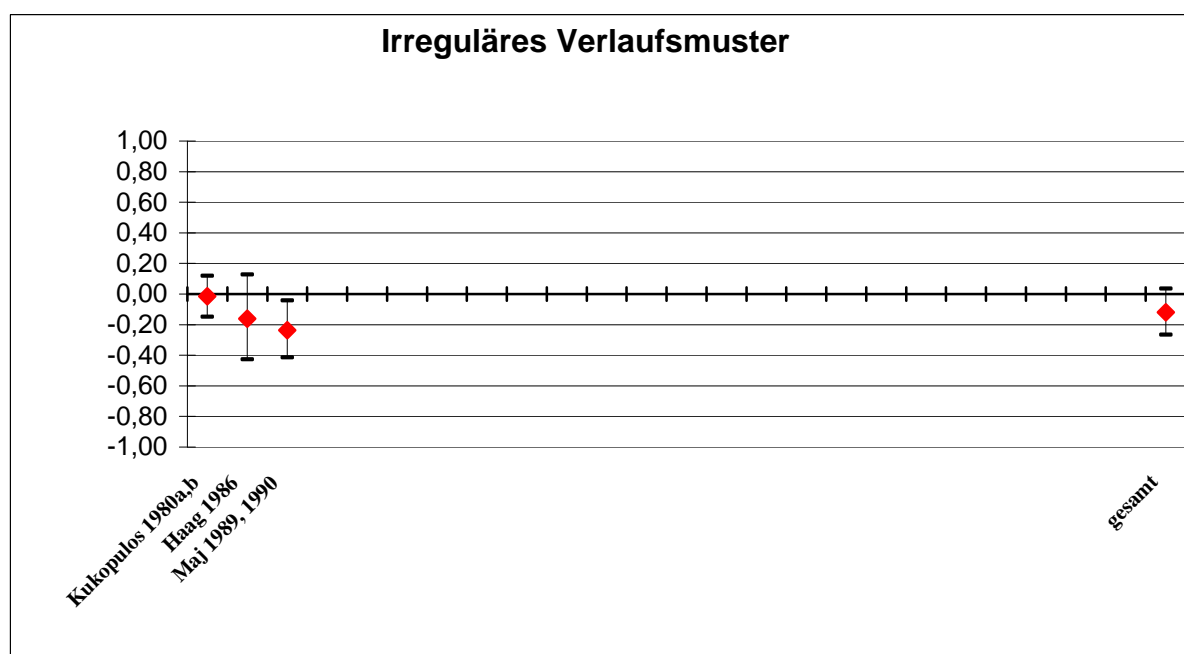


Abbildung 39. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen eines irregulären Verlaufsmusters und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.2.14 Freies Intervall

Die prädiktive Bedeutung eines freien, euthymen Intervall zwischen den Krankheitsepisoden wird in der Literatur meist im Zusammenhang mit den verschiedenen Verlaufsmustern diskutiert – so finden sich beim Verlaufsmuster MDI per definitionem freie Intervalle (vgl. Kap. 4.4.2.9), während beim Verlaufsmuster CC freie Intervalle allenfalls sporadisch auftreten. Daher ist in den Studien zum vorherrschenden Verlaufsmuster die Prädiktion aufgrund des freien Intervalls mit anderen Charakteristika der Verlaufsmuster amalgamiert. Die isolierte Bedeutung des freien Intervalls wurde in einer Studie (Yaziki et al. 1999) an  $n = 141$  Patienten untersucht. Im Ergebnis zeigte sich kein Zusammenhang zwischen dem Vorliegen freier Intervalle und der Lithiumresponse ( $r = -0.002$ , KFI =  $[-0.165, 0.169]$ ,  $p = 0.98$ ).

#### 4.4.2.15 Verlaufsscharakteristika nach J. Angst

Die meisten Untersuchungen zur prädiktiven Bedeutung von Charakteristika in der Abfolge der Krankheitsepisoden beziehen sich auf die von Kukopulos (1980a,b) vertretene Klassifikation (vgl. Kap. 4.4.2.9 – 4.4.2.14). Angst (1978, 1980) unterteilt das vorwiegende Verlaufsmuster nicht nur aufgrund der Polarität der Episoden (manisch vs. depressiv), sondern auch nach der Stärke deren Ausprägung (vgl. auch Kap. 4.4.2.9). Innerhalb der bipolaren Erkrankungen unterscheidet er drei Untergruppen: Dm, MD und Md. Dm steht für schwere Depressionen mit Hospitalisierung + Hypomanien, die keine Hospitalisierung erfordern (Angst 1978, 1980). MD steht für schwere Manien + Depressionen, die beide eine Hospitalisierung nötig machen. Md bedeutet schwere Manien mit Hospitalisierung + leichte Depressionen, die keinen stationären Aufenthalt erfordern. Insgesamt wurden 2 Studien lokalisiert, für die das Ansprechen auf Lithium mit Verlaufsmustern nach Angst in Verbindung gebracht wurden (Abouh-Saleh & Coppen 1986, Kato et al. 2000), wobei nur in der Studie von Kato et al. (2000) quantifizierbare Ergebnisse berichtet werden. In dieser Studie vom Stichprobenumfang  $n = 32$  fand sich kein Zusammenhang zwischen der Abfolge der Krankheitsepisoden und der Lithiumresponse. Für die Episodenfolge MD fand sich eine Effektstärke von  $r = -0.073$  (KFI =  $[-0.411, 0.283]$ ,  $p = 0.69$ ), für die Episodenfolge Md eine Effektstärke von  $r = -0.000$  (KFI =  $[-0.349, 0.349]$ ,  $p = 1.00$ ), und für die Abfolge mD lag die Effektstärke bei  $r = 0.093$  (KFI =  $[-0.265, 0.427]$ ,  $p = 0.62$ ). Diese Resultate sind in Einklang mit dem von Abouh-Saleh & Coppen (1986) berichteten Ergebnis, wonach das Muster der Episodenabfolge bei einer Stichprobe vom Umfang  $n = 33$  „ohne Einfluss“ auf den Behandlungserfolg unter Lithium war.

#### 4.4.2.16 Abruptes vs. schleichendes Auftreten der Manie

Die wohl einzige Studie zum Zusammenhang zwischen einem abrupten (vs. schleichenden) Auftreten der Manie und dem Ansprechen auf Lithium ist die Studie von Dunner et al. (1976). Leider geht aus der Arbeit nicht klar hervor, auf welche manische Episode sich das abrupte Auftreten der Manie bezieht (etwa auf die Indexepisode oder auf die erste Episode überhaupt). Im Ergebnis fand sich kein Zusammenhang zwischen abruptem Auftreten und dem Verlauf unter Lithium ( $r = -0.075$ , KFI = [-0.314, 0.172],  $p = 0.56$ ).

#### 4.4.2.17 Anteil manischer Episoden

Insgesamt gingen die Daten von 302 Patienten aus 3 Studien in die Metaanalyse zum Zusammenhang zwischen dem Anteil manischer (vs. depressiver) Episoden in der Vorgeschichte und der Lithiumresponse ein. Die empirische Heterogenität der Effektstärken zwischen den Studien ist sehr hoch und statistisch nicht mit der Homogenitätsannahme vereinbar ( $\chi^2 = 17.1$ ,  $p = 0.0002$ ). Eine plausible Erklärung für diese Heterogenität war nicht ersichtlich. Wie in Abb. 40 dargestellt ergab sich bei einer quantitativen Zusammenfassung dieser Studien kein statistisch signifikanter Zusammenhang zwischen dem Anteil manischer Episoden und einem günstigen Verlauf unter Lithium ( $\bar{r} = 0.091$ , KFI = [-0.199, 0.367],  $p = 0.54$ ).

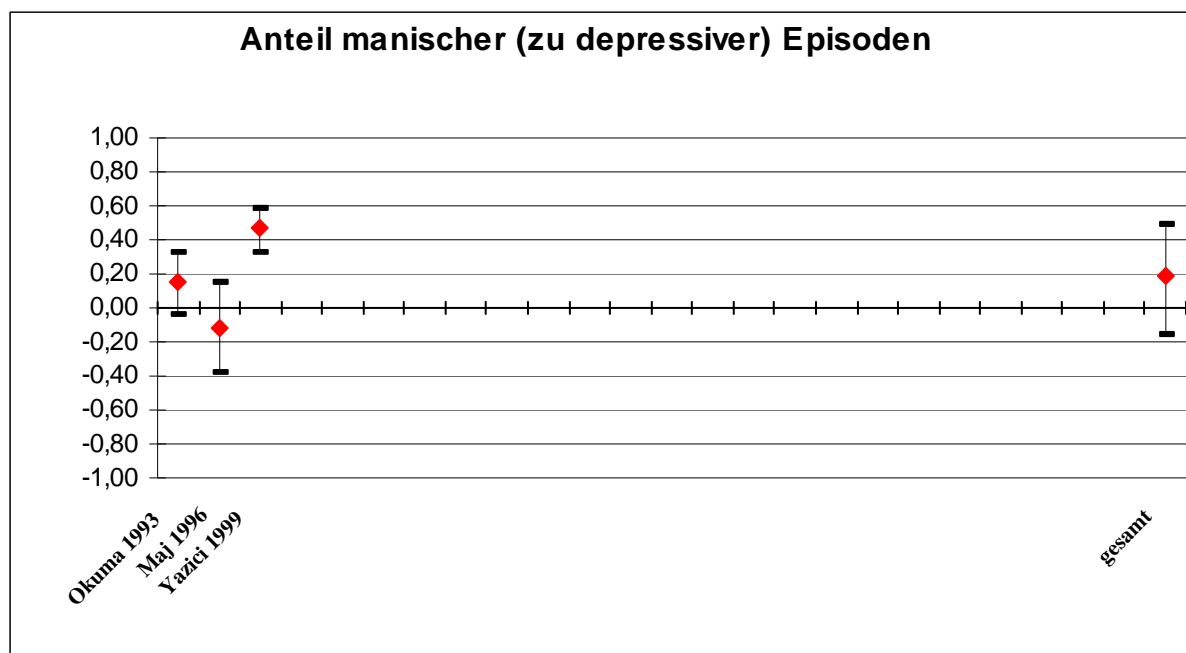


Abbildung 40. Korrelation  $r$  zwischen dem Anteil manischer Phasen und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.2.18 Hoher Anteil gemischter Episoden

Offenbar die einzige Studie, in der untersucht wurde, ob ein hoher Anteil gemischter Episoden in der Vorgeschichte mit der Response auf eine Lithiumprophylaxe korreliert, ist die Studie von Tondo et al. (2001). Im Ergebnis fand sich kein Zusammenhang zwischen dem Überwiegen gemischter Episoden in der Krankenanamnese und dem Verlauf unter Lithium ( $r = 0.032$ , KFI = [-0.135, 0.072],  $p = 0.55$ ).

#### 4.4.2.19 Typ der ersten Episode

Insgesamt lagen 5 Studien vor, in denen der Typ der ersten affektiven Episode mit dem Ansprechen auf eine Lithiumprophylaxe in Verbindung gebracht wurden. In drei Studien war dieser Zusammenhang quantifizierbar (Ananth et al. 1979, Denicoff et al. 1997, Kato 2000). In der Studie von Prien et al. (1973, 1974) wurde lediglich berichtet, daß kein signifikanter Zusammenhang bestand, in der Arbeit von Gasperini et al. 1993 war zwar der Betrag der Effektstärke ermittelbar ( $|r| = 0.1511$ ), aber nicht das Vorzeichen. Da diese Studie somit keine metaanalytisch verwertbare Information enthielt, wurde sie von der Analyse ausgeschlossen. (Zur Sicherheit wurden zusätzlich die beiden Varianten mit positivem und negativem Vorzeichen berechnet. Im Ergebnis ergab sich kein bedeutsamer Unterschied.) Auf Grundlage der 3 Primärstudien mit auswertbaren Resultaten (insgesamt  $n = 133$  Patienten) ergaben sich die im folgenden dargestellten Ergebnisse.

Die empirische Heterogenität der Effektstärken war statistisch noch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 5.83$ ,  $p = 0.054$ ). Wie in Abb. 41 dargestellt ergab sich bei einer quantitativen Zusammenfassung dieser 3 Studien (random effects model) kein statistisch signifikanter Zusammenhang zwischen dem Typ der ersten Indexepisode und einem günstigen Verlauf unter Lithium ( $\bar{r} = 0.091$ ,  $KFI = [-0.351, 0.242]$ ,  $p = 0.70$ ).

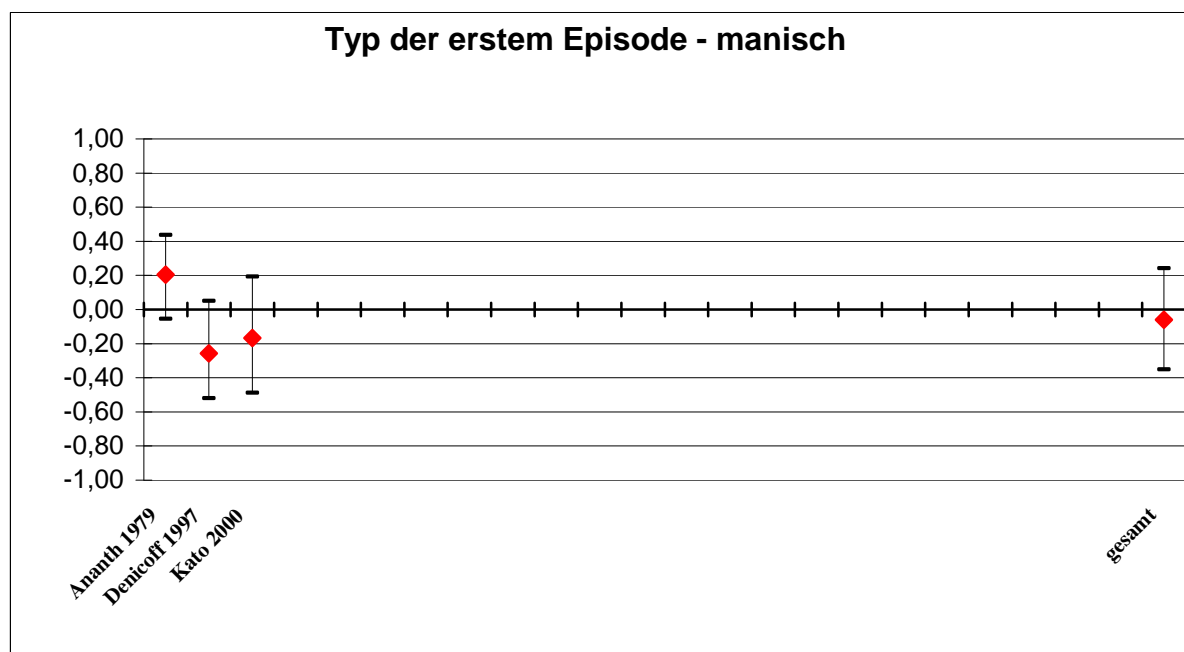


Abbildung 41. Punkt-biseriale Korrelation  $r$  zwischen dem Typ der ersten Phase (manisch) und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).



#### 4.4.2.20 Länge des ersten freien Intervalls

Zum Zusammenhang zwischen der Länge des ersten freien Intervalls und der Response auf eine Lithiumprophylaxe lag nur eine Primärstudie vor (Tondo et al. 2001). In dieser Studie mit  $n = 360$  Patienten war ein langes freies Intervall zwischen der ersten und zweiten Krankheitsphase mit einer schlechten Lithiumresponse assoziiert ( $r = -0.162$ ,  $KFI = [-0.261, -0.060]$ ,  $p = 0.002$ ). Falls – wie in diesem Fall – nur eine Studie zu einem Zusammenhang publiziert ist und in dieser Studie ein signifikantes Ergebnis berichtet wird, ist grundsätzlich mit einem Publikationsbias zu rechnen – etwa, weil signifikante Ergebnisse bevorzugt veröffentlicht werden. Diese Tendenz ist zwar allgemein bekannt, lässt sich aber im Einzelfall kaum nachweisen. Da im konkreten Fall keine klaren Anhaltspunkte für einen Bias vorliegen, wird ein langes Intervall zwischen der ersten und zweiten Krankheitsphase im folgenden als Risikofaktor für einen ungünstigen Verlauf unter Lithium betrachtet. Dieses Ergebnis steht inhaltlich in einem gewissen Widerspruch zu den Ergebnissen der Kapitel 4.4.2.2 und 4.4.2.6, wonach eine hohe Anzahl bzw. Frequenz bisheriger Krankheitsepisoden prognostisch ungünstig ist. Eine mögliche Erklärung hierfür ist, daß die Primärstudie von Tondo et al. (2001) auf prä-post Vergleichen beruht. Demnach hätten Patienten mit einem langen freien Intervall zwischen der ersten und zweiten (und vermutlich auch zwischen den weiteren) Krankheitsepisoden eine numerisch geringere Chance sich unter einer Therapie im prä-post Vergleich zu verbessern. Dies könnte ein Grund für den von Tondo et al. (2001) berichteten Befund sein, wonach ein langes freies Intervall prognostisch ungünstig ist.

#### 4.4.2.21 Psychotische Merkmale während mindestens einer Episode

####Mendlewicz muß raus#### Zum Einfluss von psychotischen Merkmalen während mindestens einer Episode auf den Verlauf unter einer Lithiumprophylaxe konnten vier Studien lokalisiert werden. In diesem Fall liegt eine erhebliche Heterogenität zwischen den Effektstärken vor ( $\chi^2 = 26.0$ ,  $p = 0.00003$ ). Möglicherweise hängt diese Heterogenität wieder damit zusammen, daß unter psychotischen Symptomen in den verschiedenen Studien unterschiedliches gemeint war (vgl. hierzu Kap. 4.4.1.6). Bei einer metaanalytischen Zusammenfassung der Effektstärken (random effects) ergab sich für die insgesamt 863 Patienten aus allen 4 Studien eine gemittelte Effektstärke von  $\bar{r} = -0.149$  (KFI = [-0.325, 0.038],  $p = 0.12$ , vgl. Abb. 42).

In der Studie von Yaziki et al. (1999) wurde neben der Variablen „psychotische Merkmale während mindestens einer Episode“ die sehr ähnliche Variable „psychotische Merkmale während der ersten Episode“ als möglicher Prädiktor untersucht. Die Korrelation zwischen der letztgenannten Variablen und der Lithiumresponse war numerisch nur geringfügig anders ( $r = -0.358$  für die Variable „psychotische Merkmale während mindestens einer Episode“ vs.  $r = -0.307$  für die Variable „psychotische Merkmale während mindestens einer Episode“).

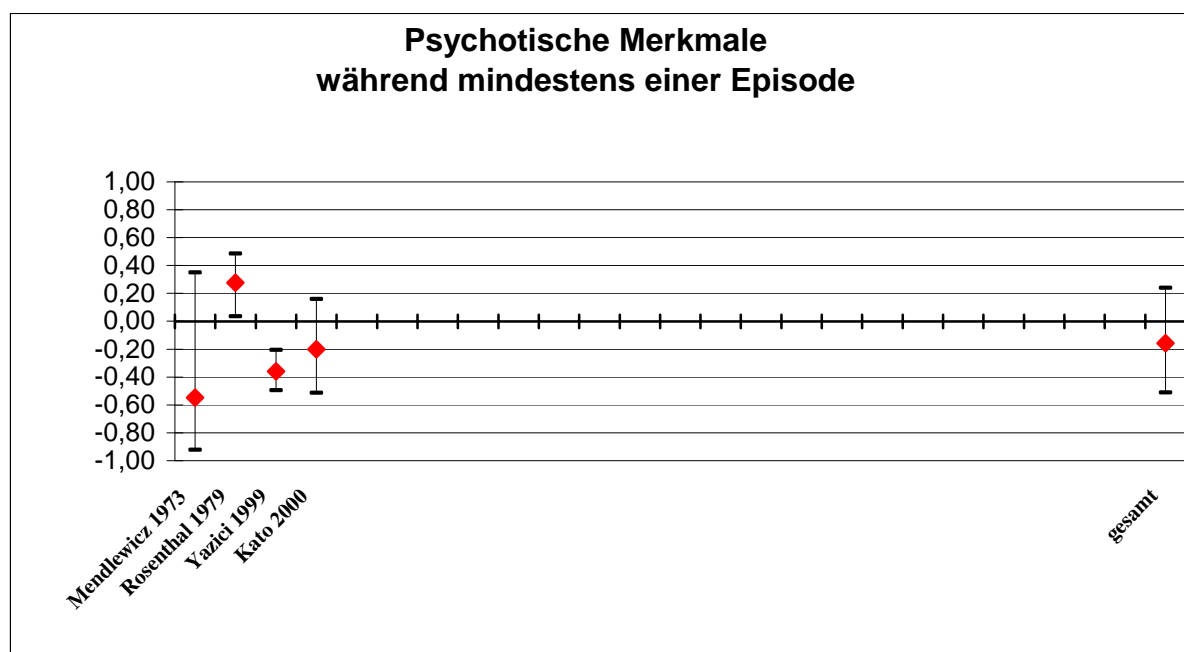


Abbildung 42. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen psychotischer Merkmale während mindestens einer Episode und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.2.22 Stimmungsinkongruente psychotische Merkmale während mindestens einer Episode

Zum Einfluss stimmungsinkongruenter psychotische Merkmale während mindestens einer Episode auf den Verlauf unter Lithium wurde nur die Primärstudie von Yaziki et al. (1999) lokalisiert. In dieser Studie mit  $n = 141$  Patienten war das Auftreten stimmungsinkongruenter psychotischer Merkmale in irgend einer der bisherigen Episoden signifikant mit einer schlechten Lithiumresponse assoziiert ( $r = -0.409$ , KFI =  $[-0.538, -0.262]$ ,  $p = 0.0000003$ ). Auch hier besteht die Problematik, daß bei nur einer publizierten Studie generell mit einem Publikationsbias gerechnet werden muss (vgl. Kap. 4.4.1.15). Aufgrund fehlender klarer Anhaltspunkte für das konkrete Vorliegen eines derartigen Bias wird das Vorliegen stimmungsinkongruenter psychotischer Merkmale während mindestens einer Episode im folgenden als Risikofaktor für einen ungünstigen Verlauf unter Lithium betrachtet.

#### 4.4.2.23 Saisonales Muster

Ob das Vorliegen eines saisonalen Musters für das Ansprechen auf eine Lithiumprophylaxe prädiktiv ist, wurde wohl nur in der Studie von Yaziki et al. (1999) untersucht. In dieser Studie fand sich kein Zusammenhang zwischen dem Vorliegen eines saisonalen Musters und dem Verlauf unter Lithium ( $r = 0.103$ , KFI =  $[-0.063, 0.264]$ ,  $p = 0.22$ ).

#### 4.4.2.24 Suizidversuch

Auch zum möglichen Zusammenhang zwischen einem Suizidversuch in der Vorgeschichte und der Lithiumresponse liegt offenbar nur die Studie von Yaziki et al. (1999) vor. Im Ergebnis fand sich kein Zusammenhang zwischen einem Suizidversuch und der Lithiumresponse ( $r = -0.079$ , KFI =  $[-0.241, 0.087]$ ,  $p = 0.35$ ).

#### 4.4.2.25 Durchschnittliche Morbidität in den Jahren vor der Prophylaxe

Zum Einfluss der durchschnittlichen Morbidität in den Jahren vor Beginn einer Lithiumprophylaxe auf das Rezidivrisiko wurden 2 Studie mit insgesamt  $n = 413$  Patienten lokalisiert (Maj et al. 1996, Tondo et al. 2001). Der Test auf Heterogenität der Effektstärken weist auf eine Verletzung der Homogenitätsannahme hin ( $\chi^2 = 5.02$ ,  $p = 0.03$ ). Dies dürfte damit zusammenhängen, daß in einer der beiden Studien (Tondo et al. 2001) die outcome-Messung auf prä-post Vergleichen beruht. Numerisch war damit für Patienten mit einer hohen Morbidität eine größere Chance für eine Verbesserung des schlechten Ausgangsniveaus gegeben. Hieraus würde sich auch das kontraintuitive Ergebnis der Studie von Tondo et al. (2001) erklären. Bei einer quantitativen Zusammenfassung der beiden Studien nach dem random effects model zeigt sich kein signifikanter Einfluss eines der durchschnittlichen Morbidität vor Prophylaxe auf die Response ( $\bar{r} = 0.106$ , KFI =  $[-0.218, 0.408]$ ,  $p = 0.53$ , vgl. Abb. 43).

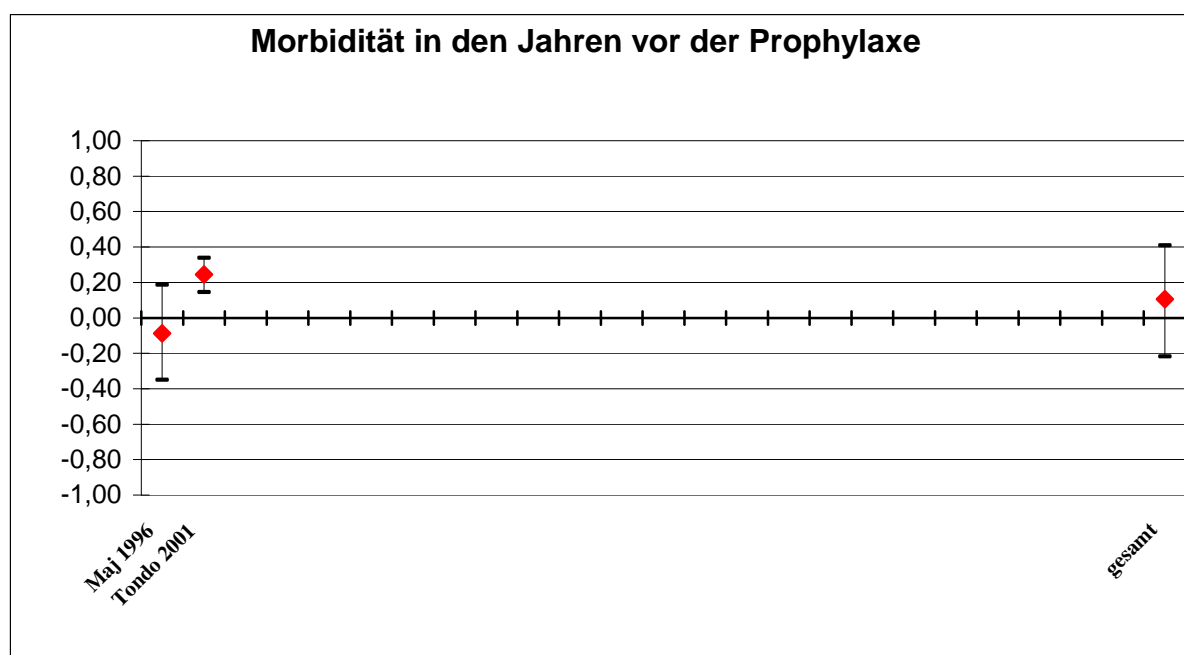


Abbildung 43. Korrelation  $r$  zwischen der durchschnittlichen Morbidität in der Zeit vor Beginn der Prophylaxe und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

### 4.4.3 Psychiatrische Komorbidität

#### 4.4.3.1 Substanzmissbrauch

Der Zusammenhang zwischen komorbidem Substanzmissbrauch und Lithiumresponse war für zwei 2 Primärstudien (zusammen 101 Patienten) quantifizierbar (Mander 1986, Maj et al. 1996). In einer weiteren Studie (O’Connell et al. 1991) wurde berichtet, daß die Inzidenz des Substanzmissbrauchs in der Gruppe mit schlechtem outcome höher war. Leider ist diese Aussage nicht quantifizierbar und muss im Rahmen dieser Metaanalyse daher unberücksichtigt bleiben.

Für die beiden eingeschlossenen Studien ergab sich kein Anhalt auf Heterogenität ( $\chi^2 = 0.066$ ,  $p = 0.80$ ). Bei einer quantitativen Zusammenfassung der Effektstärken zeigt sich kein signifikanter Zusammenhang zwischen Substanzmissbrauch und dem Verlauf unter Lithium ( $\bar{r} = -0.100$  (KFI = [-0.293, 0.100],  $p = 0.33$ , vgl. Abb. 44).

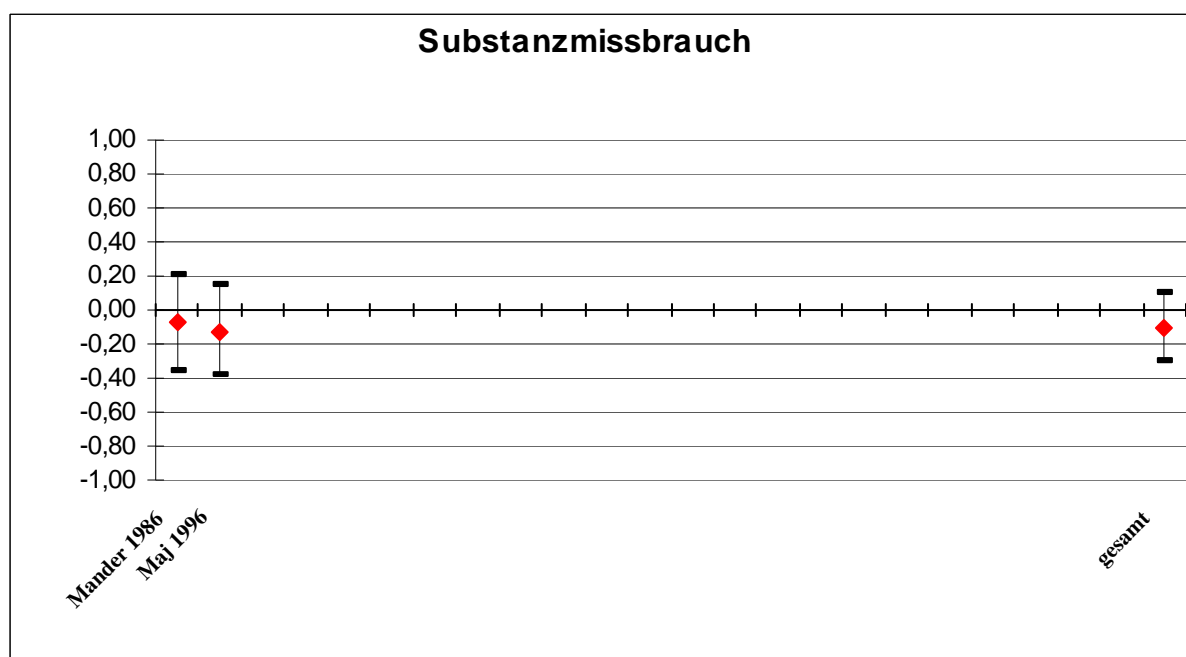


Abbildung 44. Punkt-biseriale Korrelation  $r$  zwischen der Diagnose eines Substanzmissbrauchs und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.4.3.2 Alkoholmissbrauch

Zum Einfluss von Alkoholmissbrauch auf den Verlauf unter einer Lithiumprophylaxe lag nur die Primärstudie von Coryell et al. (2000) vor. In dieser Studie zeigte sich keine signifikante Assoziation zwischen Alkoholmissbrauch und Lithiumresponse ( $r = -0.082$ , KFI =  $[-0.223, 0.063]$ ,  $p = 0.27$ ).

#### 4.4.3.3 Angst- oder Panikstörung

Die einzige Studie, die den Einschlusskriterien entsprach und den Einfluss einer komorbiden Angst- oder Panikstörung auf die Lithiumresponse untersuchte, war die Studie von Young et al. (1993). Im Ergebnis ergab sich keine signifikante Assoziation zwischen einer Angst- oder Panikstörung und dem Ansprechen auf eine Lithiumprophylaxe ( $r = -0.167$ , KFI =  $[-0.391, 0.077]$ ,  $p = 0.18$ ).

#### 4.4.3.4 Persönlichkeitsstörung

Der Zusammenhang zwischen einer komorbiden Persönlichkeitsstörung auf dem Erfolg einer Lithiumprophylaxe wurde in 2 Studien (Gasperini et al. 1993, Yazici et al. 1999) an insgesamt 138 Studienpatienten untersucht. Die empirische Heterogenität der Effektstärken zwischen den eingeschlossenen Studien ist sehr gering und statistisch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 0.044$ ,  $p = 0.83$ ). Bei einer quantitativen Zusammenfassung der beiden Studien erwies sich das Vorliegen einer komorbiden Persönlichkeitsstörung als Risikofaktor für einen ungünstigen Verlauf unter einer Lithiumprophylaxe ( $\bar{r} = -0.272$  (KFI = [-0.421, -0.108],  $p = 0.001$ , vgl. Abb. 45). Da keine Anhaltspunkte für einen Bias vorliegen, wird das Vorliegen einer Persönlichkeitsstörung im folgenden als Risikofaktor für eine schlechte Response auf eine Lithiumprophylaxe betrachtet.

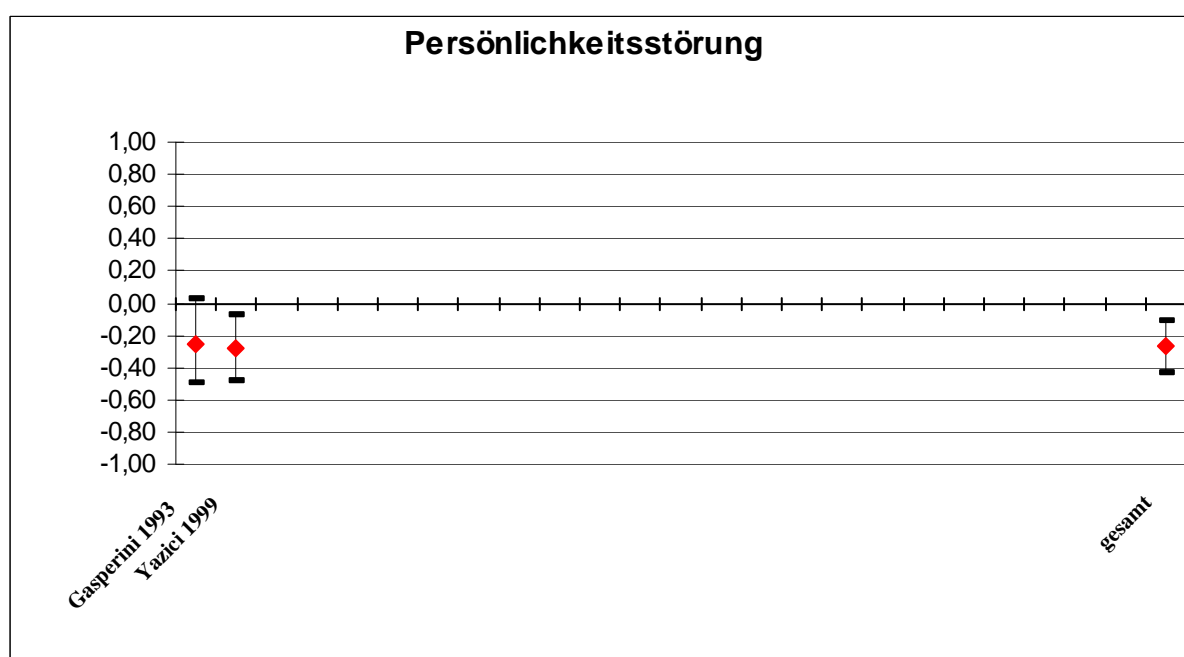


Abbildung 45. Punkt-biseriale Korrelation  $r$  zwischen der Diagnose einer Persönlichkeitsstörung und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

## 4.5 Ergebnisse der Metaanalyse zu psychosozialen Variablen

### 4.5.1 Soziale Schicht

Zum Zusammenhang zwischen sozialer Schicht und Verlauf unter einer Lithiumprophylaxe konnten zwei Studien lokalisiert werden (Mander 1986, O'Connell et al. 1991). In beiden Studien war eine hohe soziale Schicht tendenziell bzw. signifikant mit einem positiven Verlauf assoziiert (vgl. Abb. 46). Insgesamt ergab sich für die insgesamt 295 Patienten aus beiden Studien eine gemittelte Effektstärke von  $\bar{r} = 0.280$  (KFI = [0.171, 0.383],  $p = 0.000001$ ). Die empirische Heterogenität der Effektstärken zwischen den Studien ist gut mit der Homogenitätsannahme vereinbar ( $\chi^2 = 0.015$ ,  $p = 0.90$ ). Da hier keine empirisch belegbaren Anhaltspunkte für einen Bias vorliegen, wird eine hohe soziale Schicht im folgenden als Positivprädiktor für einen günstigen Verlauf unter Lithium betrachtet.

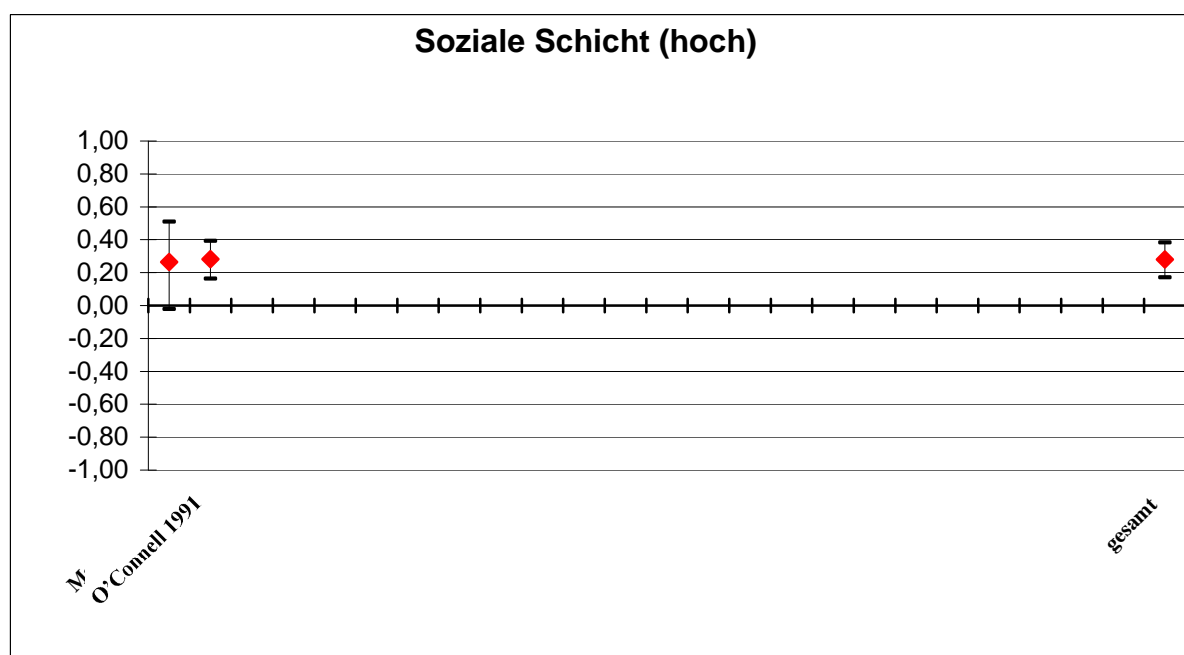


Abbildung 46. Punkt-biseriale Korrelation  $r$  zwischen hoher sozialer Schicht und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).



### 4.5.2 Arbeitslosigkeit

Zum Einfluss von Arbeitslosigkeit auf den Verlauf unter einer Lithiumprophylaxe lag nur die Studie von Tondo et al. (2001) vor, wonach Arbeitslosigkeit zu Therapiebeginn mit einem ungünstigen Verlauf unter einer Lithiumprophylaxe assoziiert ist ( $\bar{r} = -0.118$ , KFI = [-0.219, -0.015],  $p = 0.03$ ). Zwar ist – wie immer, wenn nur eine Arbeit vorliegt und einen signifikanten Zusammenhang berichtet – mit einem gewissen Publikationsbias zu rechnen (vgl. Kap. 4.4.1.15). Da es aber keine empirisch belegbaren Anhaltspunkte für einen Bias gibt, wird Arbeitslosigkeit im folgenden als Risikofaktor für einen ungünstigen Verlauf unter Lithium betrachtet.

### 4.5.3 Soziale Unterstützung (social support)

Insgesamt gingen die Daten von 375 Patienten aus 3 Studien in die Metaanalyse zum Zusammenhang zwischen sozialer Unterstützung und Lithiumresponse ein. Wie in Abb. 47 dargestellt, erwies sich bei einer quantitativen Zusammenfassung dieser Studien eine gute soziale Unterstützung als prognostisch günstig für eine gute Lithiumresponse ( $\bar{r} = 0.171$ ,  $KFI = [0.070, 0.268]$ ,  $p = 0.001$ ). Der Test auf Heterogenität der Effektstärken gibt keinen Anhalt für die Verletzung der Homogenitätsannahme ( $\chi^2 = 1.62$ ,  $p = 0.44$ ). Da auch hier keine Anhaltspunkte für einen Bias vorliegen, wird eine gute soziale Unterstützung im folgenden als Positivprädiktor betrachtet.

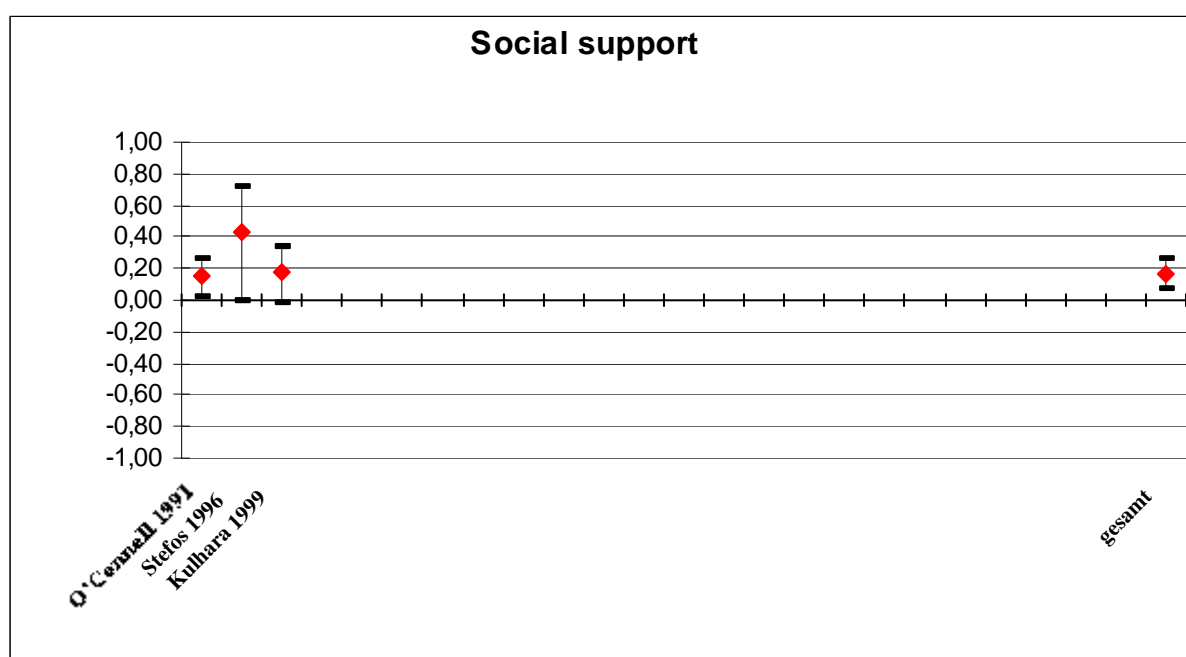


Abbildung 47. Punkt-biseriale Korrelation  $r$  zwischen guter sozialer Unterstützung und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### 4.5.4 Familienatmosphäre (expressed emotions und affective style)

Zum Zusammenhang zwischen der Familienatmosphäre und dem Verlauf unter einer Lithiumprophylaxe lag nur die Primärstudie von O'Connell et al. (1991) vor. In dieser Studie wurden zwei Variablen mit der Lithiumresponse in Zusammenhang gebracht – expressed emotions (EE) und affective style (AS). Das Konstrukt der expressed emotions setzt sich aus zwei Dimensionen zusammen: "Kritische Bemerkungen" und "Emotionales Überengagement" (jeweils von Seiten der Angehörigen). Von "high expressed emotions" spricht man, falls in mindestens einer dieser beiden Dimensionen ein empirisch definierter Schwellenwert überschritten wird. Mit einem "negative affective style" ist insbesondere ein herabwürdigender, zudringlicher oder überkritisch-respektloser Kommunikationsstil der Angehörigen gemeint (vgl. Miklowitz 1988, O'Connell et al. 1991).

Im Ergebnis zeigte sich ein signifikant negativer Einfluss von high expressed emotions auf den Verlauf unter Lithium ( $\bar{r} = -0.200$ , KFI = [-0.341, -0.51],  $p = 0.009$ ) und ein negativer Trend für die Variable "negative affective style" ( $\bar{r} = -0.115$ , KFI = [-0.292, 0.003],  $p = 0.055$ ). Zwar ist – wie stets bei isoliert publizierten signifikanten Ergebnissen – mit einem gewissen Publikationsbias zu rechnen (vgl. Kap. 4.4.1.15). Allerdings liegen keine empirisch belegbaren Anhaltspunkte für einen Bias vor. Daher wird die Variable "high expressed emotions" im folgenden als Risikofaktor für ein affektives Rezidiv unter einer Lithiumprophylaxe betrachtet.

Diese Ergebnisse werden unterstützt durch eine verwandte Arbeit von Stefos et al. (1996). In dieser Arbeit fand sich trotz der kleinen Fallzahl von  $n = 21$  ein signifikant positiver Einfluss einer hohen Beziehungsqualität zu Familie und familiärem Umfeld auf den Verlauf unter Lithium ( $\bar{r} = -0.200$ , KFI = [0.026, 0.740],  $p = 0.04$ ).

#### 4.5.5 Adäquate Freizeitaktivitäten

Zum Zusammenhang zwischen Freizeitaktivitäten und Lithiumresponse lag nur eine und dazu sehr kleine ( $n = 21$ ) Primärstudie vor (Stefos et al. 1996). In dieser Studie waren adäquate Freizeitaktivitäten tendenziell protektiv gegen einen ungünstigen Verlauf unter einer Lithiumprophylaxe ( $\bar{r} = 0.4301$ , KFI = [-0.002, 0.727]), aufgrund der geringen statistischen power erreichte das Ergebnis allerdings nicht ganz die Signifikanzschwelle ( $p = 0.051$ ).

#### 4.5.6 Stress

Auch zum Einfluss von Stress auf den Verlauf unter Lithium konnte nur eine Primärstudie lokalisiert werden (Kulhara et al. 1999). Hierbei zeigte sich eine deutlich negative Assoziation zwischen Stress und Rezidivfreiheit unter Lithium ( $\bar{r} = -0.254$ , KFI =  $[-0.416, -0.077]$ ,  $p = 0.005$ ). Falls nur wenige oder gar nur eine Studie zu einem Zusammenhang publiziert ist und in dieser Studie ein signifikantes Ergebnis berichtet wird, ist immer damit zu rechnen, daß bei der Publikation eine gewisse Auslese stattgefunden hat – etwa, weil sogenannte negative Ergebnisse schwerer publizierbar sind. Diese Tendenz ist zwar allgemein bekannt, lässt sich aber im Einzelfall nur selten nachprüfen. Da im konkreten Fall keine empirisch belegbaren Anhaltspunkte für einen Bias vorliegen, wird Stress im folgenden als Risikofaktor für einen ungünstigen Verlauf unter Lithium betrachtet.

#### 4.5.7 Kritische Lebensereignisse (life events)

Insgesamt konnte für  $n = 699$  Patienten aus 6 Primärstudien ein Zusammenhang zwischen kritischen Lebensereignissen (sog. "life events") und dem Ansprechen auf eine Lithiumprophylaxe hergestellt werden. Wie aus Abb. 48 ersichtlich ergab sich bei einer quantitativen Zusammenfassung dieser Studien ein signifikant negativer Zusammenhang zwischen der Anzahl von life events und einem günstigen Verlauf unter Lithium ( $\bar{r} = -0.085$ , KFI = [-0.166, -0.002],  $p = 0.045$ ). Die empirische Heterogenität der Effektstärken zwischen den Studien ist statistisch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 5.78$ ,  $p = 0.33$ ). Da auch in diesem Falle kein Anhaltspunkt für einen Bias vorliegt, wird eine große Zahl kritischer Lebensereignisse in der Vorgeschichte im folgenden als Risikofaktor für einen ungünstigen Verlauf unter Lithium betrachtet.

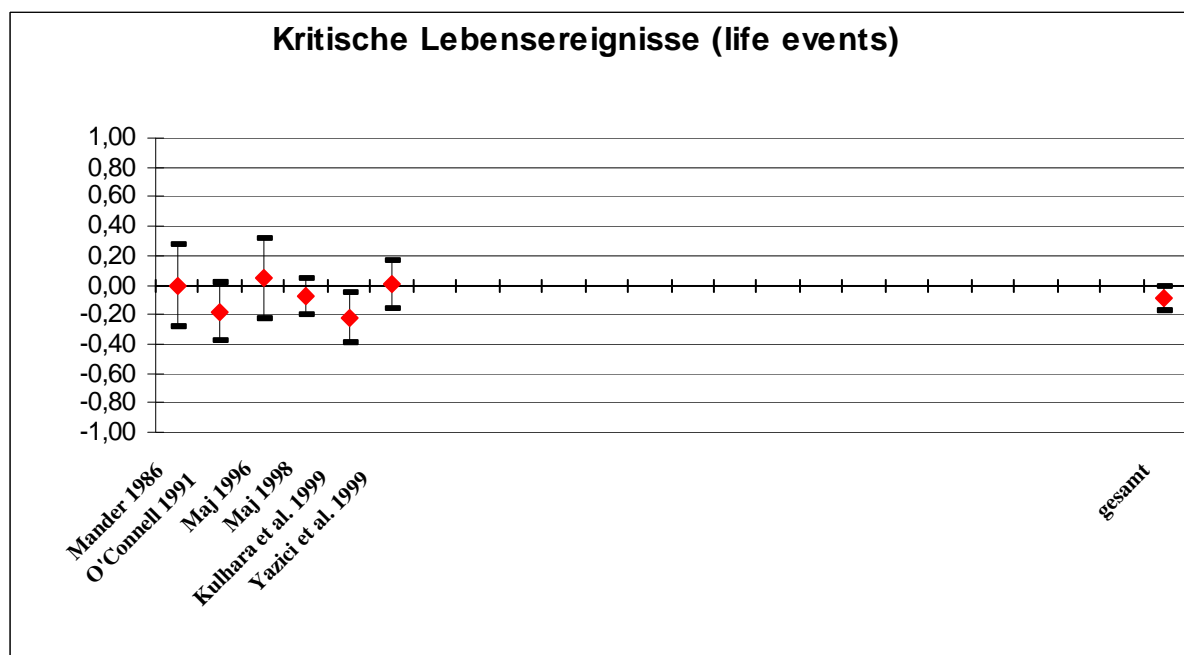


Abbildung 48. Punkt-biseriale Korrelation  $r$  zwischen dem Vorliegen kritischer Lebensereignisse (life events) und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

### 4.5.8 Ängstlichkeit

Der Zusammenhang zwischen Ängstlichkeit und dem Verlauf unter einer Lithiumprophylaxe wurde in drei der lokalisierten Primärstudien beschrieben. In zwei dieser Studien war der Zusammenhang in direkt quantifizierbarer Weise angegeben, in der dritten Studie (Abou-Saleh & Coppen 1986) wurde nur berichtet, daß eine signifikant negative Beziehung zwischen einem hohen Ängstlichkeits-Score und einer guten Lithiumresponse bestand. Diese Aussage enthält Information über Ausmaß und Richtung des Zusammenhangs zwischen Ängstlichkeit und Lithiumresponse, und daher wäre ein Weglassen dieser Information mit einem vermutlich größeren Fehler verbunden gewesen als eine Schätzung des Zusammenhangs aus den angegebenen Informationen. Diese Schätzung der Effektstärke  $r$  wurde konservativ vorgenommen, d.h. für  $p$  wurde die Signifikanzschwelle 0.05 angesetzt. Hieraus wurde über die der Stichprobengröße entsprechende  $t$ -Verteilung ein  $t$ -Wert bestimmt und dieser  $t$ -Wert wurde schließlich in eine Effektstärkeschätzung übersetzt (vgl. Kap. 4.1.3).

Im Ergebnis fand sich für die insgesamt 178 Patienten aus den drei Studien kein statistisch signifikanter Zusammenhang zwischen Ängstlichkeit und Lithiumresponse ( $\bar{r} = -0.115$ , KFI = [-0.260, 0.036],  $p = 0.13$ , vgl. Abb. 49). Die empirische Heterogenität der Effektstärken ist mit der Homogenitätsannahme vereinbar ( $\chi^2 = 1.848$ ,  $p = 0.40$ ).

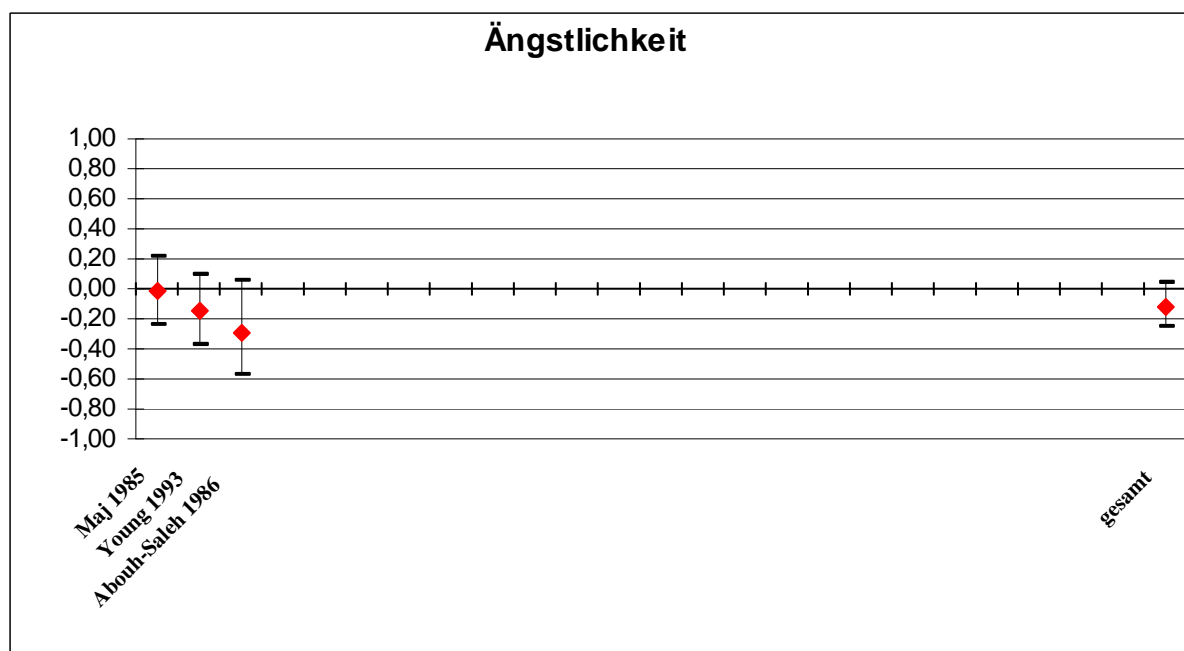


Abbildung 49. Korrelation  $r$  zwischen hoher Ängstlichkeit und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

### 4.5.9 Neurotizismus

Zum Zusammenhang zwischen Neurotizismus und Verlauf unter einer Lithiumprophylaxe konnten drei Studien lokalisiert werden. In zwei dieser Studien war der Zusammenhang direkt als Effektstärke quantifizierbar, in der dritten Studie (Abou-Saleh & Coppen 1986) wurde nur angegeben, daß eine signifikant negative Beziehung zwischen einem hohen Neurotizismus-Score und einer guten Lithiumresponse bestand. Da diese Aussage Information über das Ausmaß und die Richtung des Zusammenhangs zwischen Neurotizismus und Lithiumresponse enthält, wurde sie in eine (konservative) Schätzung der Effektstärke  $r$  übersetzt, für  $p$  wurde also die Signifikanzschwelle 0.05 angesetzt. Hieraus wurde über die der Stichprobengröße entsprechende  $t$ -Verteilung ein  $t$ -Wert bestimmt und dieser  $t$ -Wert wurde nach Formel (F3) in eine Effektstärkeschätzung übersetzt (vgl. Kap. 4.1.3).

Hieraus ergab sich für die insgesamt 143 Patienten aus den drei Studien eine gemittelte Effektstärke (random effects model) von  $\bar{r} = -0.259$  (KFI = [-0.471, -0.019],  $p = 0.03$ , vgl. Abb. 50). Die empirische Heterogenität der Effektstärken zwischen den Studien ist mit der Homogenitätsannahme vereinbar ( $\chi^2 = 3.774$ ,  $p = 0.15$ ). Auch hier liegen keine empirisch belegbaren Anhaltspunkte für einen Bias vor. Daher wird eine hoher Neurotizismus-Score im folgenden als Risikofaktor für einen ungünstigen Verlauf unter Lithium betrachtet.

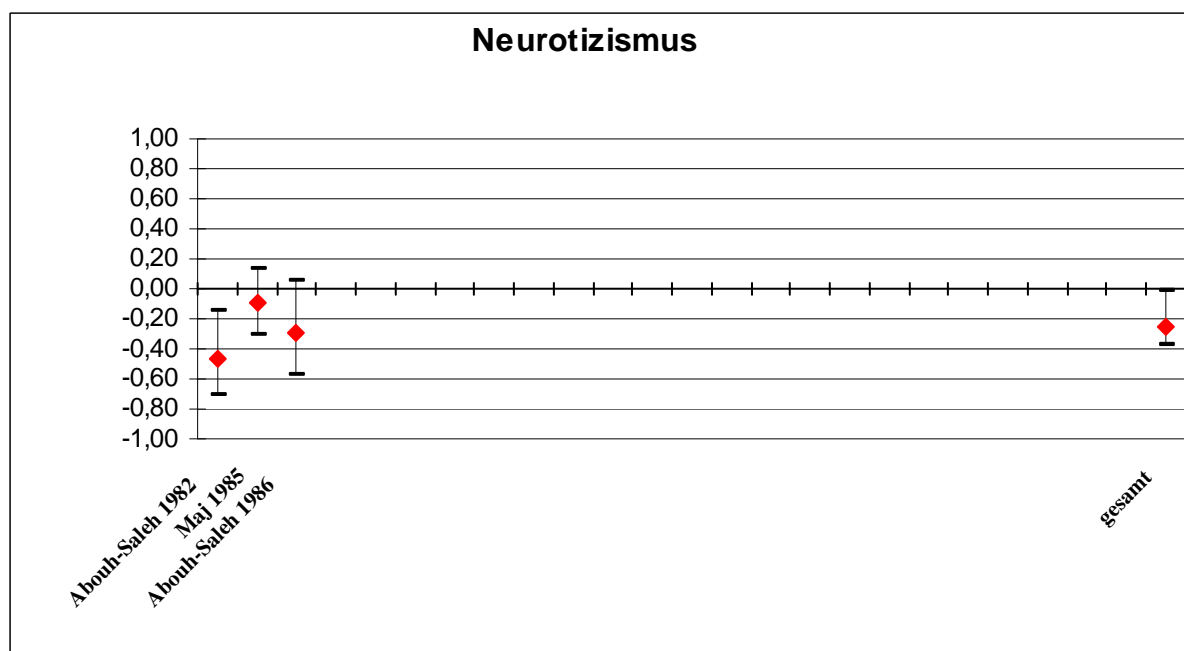


Abbildung 50. Korrelation  $r$  zwischen starkem Neurotizismus und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

### 4.5.10 Extraversion

Aufgrund der drei Primärstudien zur Prädiktion der Lithiumresponse aufgrund von Extraversion-Scores fand sich kein signifikanter Zusammenhang zwischen Extraversion und dem Verlauf unter einer Lithiumprophylaxe ( $\bar{r} = 0.076$  (KFI = [-0.085, 0.232],  $p = 0.35$ , vgl. Abb. 51). Die empirische Heterogenität der Studien hinsichtlich der Effektstärken zum Zusammenhang zwischen Extraversion und Verlauf war mit der Homogenitätsannahme vereinbar ( $\chi^2 = 2.748$ ,  $p = 0.25$ ).

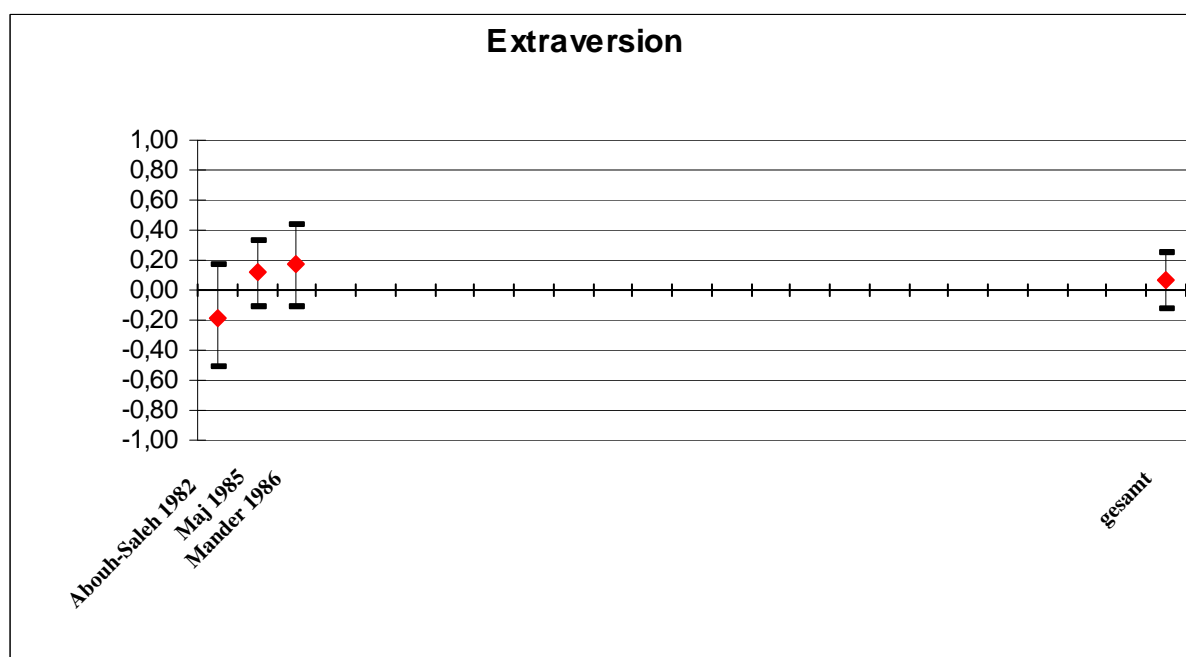


Abbildung 51. Korrelation  $r$  zwischen starker Extraversion und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).



### 4.5.11 Psychotizismus

Bei einer quantitativen Zusammenfassung (random effects model) der beiden Studien zur Prädiktion der Lithiumresponse aufgrund von Psychotizismus-Scores ergab sich kein signifikanter Zusammenhang zwischen Psychotizismus und Response ( $\bar{r} = -0.098$  (KFI = [-0.282, 0.094],  $p = 0.32$ , vgl. Abb. 52). Die empirische Heterogenität der Effektstärken zwischen den Studien ist statistisch mit der Homogenitätsannahme vereinbar ( $\chi^2 = 0.848$ ,  $p = 0.36$ ).

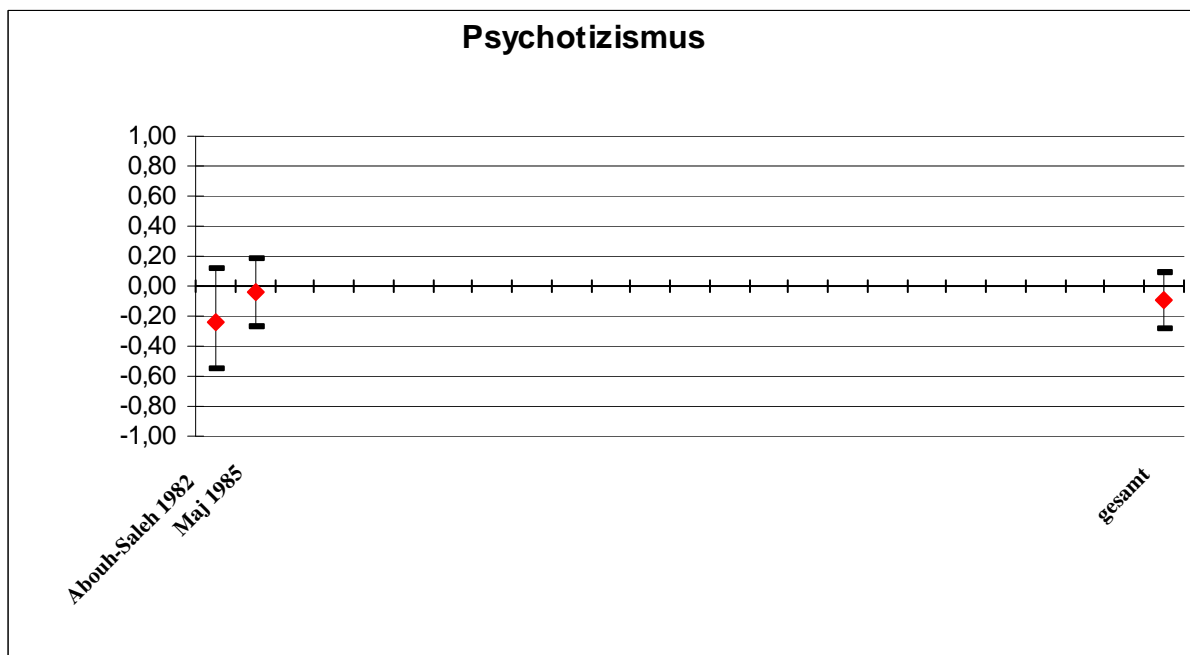


Abbildung 52. Korrelation  $r$  zwischen hohem Psychotizismus und der prophylaktischen Wirksamkeit von Lithium. Gesamtmodell auf Grundlage zufälliger Effekte (DSL-Ansatz).

#### **4.5.12 Zwanghaftigkeit**

Bisher liegt wohl nur eine Prophylaxestudie zum Zusammenhang von Zwanghaftigkeit und Lithiumresponse bei bipolaren Störungen vor (Maj et al. 1985). Für die  $n = 78$  Patienten dieser Studie ergab sich kein Anhalt für einen Zusammenhang zwischen den mit dem Middlesex Hospital Questionnaire erfassten Zwanghaftigkeits-Scores zu Studienbeginn und der Response ( $r = -0.022$  (KFI =  $[-0.244, 0.201]$ ,  $p = 0.85$ ).

#### **4.5.13 Depressive Persönlichkeit**

Zwar liegt eine Reihe von Befunden zur Prädiktion der Lithiumresponse aufgrund von Variablen vor, die dem Inhalt nach der Variablen "depressive Persönlichkeit" verwandt sind. Zur Prädiktion der Response aufgrund eines Scores zur Erfassung depressiver Persönlichkeitszüge liegt für bipolare Patienten aber offenbar nur die Studie von Maj et al. (1985) vor. Im Ergebnis zeigte sich kein signifikanter Zusammenhang zwischen depressiven Persönlichkeitszügen und Response ( $r = -0.060$  (KFI =  $[-0.279, 0.165]$ ,  $p = 0.60$ ,  $n = 78$ ). Depressive Persönlichkeit wurde hierbei als Summenscores der entsprechenden Subskala des Middlesex Hospital Questionnaire erfasst.

#### **4.5.14 Histrionische Persönlichkeit**

Auch zur Response-Prädiktion aufgrund eines Scores zur Erfassung histrionischer Persönlichkeitszüge liegen für bipolare Patienten wohl nur die auf dem Middlesex Hospital Questionnaire basierenden Befunde von Maj et al. (1985) vor. Die Autoren fanden keinen Zusammenhang zwischen histrionischen Persönlichkeitszügen und der Response ( $r = -0.016$  (KFI =  $[-0.237, 0.208]$ ,  $p = 0.89$ ,  $n = 78$ ).

#### **4.5.15 Zykllothyme Persönlichkeit**

Zur Prädiktion der Lithiumresponse bei bipolaren Patienten aufgrund zykllothymen Persönlichkeitszüge wurde nur die an  $n = 59$  Patienten durchgeführte Studie von Ananth et al. (1979) lokalisiert. Im Ergebnis zeigte sich in dieser Studie kein signifikanter Zusammenhang zwischen den Scores für zykllothyme Persönlichkeit und dem Ansprechen auf eine Lithiumprophylaxe ( $r = -0.110$  (KFI =  $[-0.150, 0.356]$ ,  $p = 0.41$ ).

#### 4.5.16 Somatisierungstendenzen

In der bereits mehrfach angeführten Arbeit von Maj et al. (1985) wird unter der Rubrik „Persönlichkeitsvariablen“ der Zusammenhang zwischen Lithiumresponse und der auf Grundlage des Middlesex Hospital Questionnaire erhobenen Somatisierungstendenz berichtet. Im Ergebnis zeigte sich kein signifikanter Einfluss von Somatisierungstendenzen auf den Erfolg einer Lithiumprophylaxe ( $r = -0.044$  (KFI = [-0.264, 0.180],  $p = 0.70$ ,  $n = 78$ ).

#### 4.5.17 Extrapunitivität, Intrapunitivität und Dominanz

Zur Prädiktion der Lithiumresponse aufgrund der drei in der Foulds' Personality Deviance Scale (Foulds 1976) erfassten Dimensionen Extrapunitivität, Intrapunitivität und Dominanz liegt für die Langzeitbehandlung bipolarer Patienten wohl nur die Studie von Abouh-Saleh (1982) vor. Es handelt sich um eine retrospektive Studie bei  $n = 32$  bipolaren Patienten.

Im Ergebnis waren weder Extrapunitivität noch Intrapunitivität mit einem Rezidiv unter Lithium assoziiert ( $r = 0.034$  (KFI = [-0.318, 0.378],  $p = 0.85$  für Intrapunitivität und  $r = 0.073$  (KFI = [-0.283, 0.411],  $p = 0.69$  für Extrapunitivität). Hingegen zeigte sich trotz des relativ geringen Stichprobenumfangs ein signifikant positiver Zusammenhang zwischen dominanten Charakterzügen zu Studienbeginn und dem Ansprechen auf eine Lithiumprophylaxe. ( $r = 0.352$  (KFI = [0.003, 0.624],  $p = 0.048$ ). Da kein empirisch greifbarer Anhaltspunkt für das Vorliegen eines Bias besteht wird im Rahmen dieser Metaanalyse davon ausgegangen, daß Dominanz ein protektiver Faktor gegen das Auftreten von Rezidiven unter einer Lithiumprophylaxe bei bipolaren Störungen ist.

#### 4.5.18 Soziale Erwünschtheit (Normorientiertheit)

Die meisten Persönlichkeitsfragebögen enthalten eine Kontrollskala ("lie scale"). Diese Skala dient zunächst dazu, festzustellen, ob der Patient vermutlich wahrheitsgemäß antwortet, oder ob eine Antworttendenz in Richtung der sozialen Erwünschtheit besteht. Die Tendenz zu einem Norm-orientierten Antwortverhalten kann allerdings auch als eine Persönlichkeitseigenschaft aufgefasst werden, und gerade im Zusammenhang mit affektiven Störungen spielt eine Normorientiertheit eine bedeutende Rolle (vgl. etwa das Konzept des *typus melancholicus* von Tellenbach 1961).

Zum Zusammenhang zwischen Lithiumresponse bei bipolaren Störungen und Normorientiertheit wurde einzig die an  $n = 32$  Patienten durchgeführte Studie von Abou-Saleh

(1983) lokalisiert. Hierbei zeigte sich ein tendenziell positiver Einfluss von Normorientiertheit auf den Erfolg einer Lithiumprophylaxe ( $r = 0.344$  (KFI = [-0.005, 0.618],  $p = 0.054$ ).

#### **4.5.19 Compliance**

Eine gute Patientencompliance gilt als eine der zentralen Variablen für den Erfolg einer Langzeitbehandlung (Leverich et al 1990, Guscott & Taylor 1994, Keck et al. 1997, Schou 1997, Colom et al. 2000). Dennoch wurde nur eine Studie zum Zusammenhang von Compliance und Behandlungserfolg lokalisiert, die allen Einschlusskriterien entsprach (Kulhara et al. 1999). In dieser Studie an  $n = 118$  Patienten erwies sich eine gute Compliance als klinisch wie statistisch signifikanter Prädiktor für eine gute Lithiumresponse ( $r = 0.296$  (KFI = [0.122, 0.453],  $p = 0.001$ ). Da keine Anhaltspunkte für einen Bias vorliegen, wird eine gute Patientencompliance im folgenden als Positivprädiktor für eine gute Lithiumresponse betrachtet.

#### 4.6 Responseprädiktoren – Zusammenfassung der Metaanalyse

Im Rahmen dieser Metaanalyse wurden Primärstudien zu insgesamt 68 möglichen Prädiktoren für ein gutes Ansprechen auf eine Lithiumprophylaxe untersucht. Da es in einem ersten Schritt wesentlich erschien, die Primärstudien zur Responseprädiktion mit einem nicht allzu großen Informationsverlust zu kategorisieren und zu aggregieren, wurde im Zweifelsfall so verfahren, daß die Ergebnisse der Primärstudien für verwandte oder einander überlappende Variablen getrennt aggregiert wurden. Beispielsweise wurden der Einfluss psychotischer Symptome in der Indexepisode und der Einfluss stimmungskongruenter psychotische Symptome auf das Rezidivrisiko getrennt untersucht.

Für jede dieser 68 Variablen wurde die Information aus allen Studien in einem Effektstärkemaß zusammenfasst, welches die Stärke des Zusammenhangs zwischen der Prädiktorvariablen und der Lithiumresponse beschreibt. Die Zusammenfassung der Einzelergebnisse und die Signifikanzprüfung des Gesamtergebnis wurden im Rahmen eines Modells mit zufälligen Effekten (DerSimonian-Laird 1986) durchgeführt. Bei Zusammenfassung aller Studien erwiesen sich die in Tabelle 6 zusammengefassten 21 Variablen als signifikante Prädiktoren für das Ansprechen auf Lithium. Das Vorzeichen des Effektstärkemaßes  $\bar{r}$  ist stets so gewählt, daß Risikofaktoren ein negatives Vorzeichen für einen Risikofaktor steht und ein positives Vorzeichen für einen protektiven Faktor. Der Absolutbetrag  $|\bar{r}|$  gibt die Stärke des Zusammenhangs an (vgl. Tab. 6).

Tabelle 6: Responseprädiktoren für das Ansprechen auf eine Lithiumprophylaxe.

<b><i>Protective Faktoren</i></b>	<b><i>Risikofaktoren</i></b>
<p>Demographische Variablen: <i>keine signifikanten Prädiktoren</i></p> <p>Diagnostik und Psychopathologie:</p> <p>Anamnese:</p> <ul style="list-style-type: none"> <li>• Verlaufsmuster MDI (<math>\bar{r} = 0.203</math>, <math>p = 0.0004</math>)</li> <li>• Isolierte Krankheitsepisoden (<math>\bar{r} = 0.324</math>, <math>p = 0.0006</math>)</li> <li>• Höheres Alter bei Ersterkrankung (<math>\bar{r} = 0.114</math>, <math>p = 0.003</math>)</li> </ul> <p>Psychiatrische Komorbidität:</p> <p>Psychosoziale Variablen:</p> <ul style="list-style-type: none"> <li>• Hohe soziale Schicht (<math>\bar{r} = 0.280</math>, <math>p = 0.000001</math>)</li> <li>• Compliance (<math>\bar{r} = 0.296</math>, <math>p = 0.001</math>)</li> <li>• Soziale Unterstützung (<math>\bar{r} = 0.171</math>, <math>p = 0.001</math>)</li> <li>• Dominanz (<math>\bar{r} = 0.352</math>, <math>p = 0.0048</math>)</li> </ul>	<p>Demographische Variablen: <i>keine signifikanten Prädiktoren</i></p> <p>Diagnostik und Psychopathologie:</p> <ul style="list-style-type: none"> <li>• Depressivität oder Verlangsamung (<math>\bar{r} = -0.307</math>, <math>p = 0.02</math>)</li> </ul> <p>Anamnese:</p> <ul style="list-style-type: none"> <li>• Hohe Anzahl stationärer Episoden (<math>\bar{r} = -0.347</math>, <math>p &lt; 0.0000000001</math>)</li> <li>• Jemals stimmunginkongruente Wahnhinhalte (<math>\bar{r} = -0.409</math>, <math>p = 0.0000003</math>)</li> <li>• Verlaufsmuster CC (<math>\bar{r} = -0.213</math>, <math>p = 0.00002</math>)</li> <li>• Länge des ersten freien Intervalls (<math>\bar{r} = -0.162</math>, <math>p = 0.002</math>)</li> <li>• Verlaufsmuster DMI (<math>\bar{r} = -0.122</math>, <math>p = 0.02</math>)</li> <li>• Hohe Phasenfrequenz (<math>\bar{r} = -0.284</math>, <math>p = 0.03</math>)</li> <li>• Krankheitsdauer (<math>\bar{r} = -0.160</math>, <math>p = 0.03</math>)</li> </ul> <p>Psychiatrische Komorbidität:</p> <ul style="list-style-type: none"> <li>• Persönlichkeitsstörung (<math>\bar{r} = -0.272</math>, <math>p = 0.001</math>)</li> </ul> <p>Psychosoziale Variablen:</p> <ul style="list-style-type: none"> <li>• Stress (<math>\bar{r} = -0.254</math>, <math>p = 0.005</math>)</li> <li>• High expressed emotions (<math>\bar{r} = -0.200</math>, <math>p = 0.009</math>)</li> <li>• Hoher Neurotizismus (<math>\bar{r} = -0.259</math>, <math>p = 0.03</math>)</li> <li>• Arbeitslosigkeit (<math>\bar{r} = -0.118</math>, <math>p = 0.03</math>)</li> <li>• Life events (<math>\bar{r} = -0.085</math>, <math>p = 0.045</math>)</li> </ul>

Diese 21 klinischen und psychosozialen Variablen haben sich bei einer systematischen Literaturübersicht als Prädiktoren der Lithiumresponse erhärtet. Grundlage waren alle bis 3/2003 zu bipolaren Störungen publizierten und im Rahmen eines umfangreichen Rechercheprozesses lokalisierten Prophylaxestudien. Diese Studien wurden in der vorliegenden Arbeit nach metaanalytischen Methoden aggregiert.

Das Hauptthema des zweiten Teils dieser Arbeit ist, die im ersten Teil identifizierten Prädiktoren in geeigneter Weise zu verknüpfen. Das leitende Ziel hierbei ist, eine für die klinische Anwendung optimierte Prädiktion der Lithiumresponse zu erreichen. Zu diesem Zweck werden verschiedene, theoretisch naheliegende, Varianten diskutiert und empirisch an einer umfangreichen prospektiven Langzeitstudie auf ihre Eignung geprüft.

## **5 Daten und Messinstrumente**

### **5.1 Beschreibung der MAP-Studie**

#### **5.1.1 Generelle Charakterisierung der MAP-Studie**

Die MAP-Studie ist eine vom ehemaligen Bundesministerium für Forschung und Technologie (BMFT) geförderte „multizentrische prospektive kontrollierte Therapiestudie zum Vergleich ambulanter Langzeitbehandlung affektiver und schizoaffektiver Psychosen“. Die Studie wurde an 9 psychiatrischen Universitätskliniken in Deutschland durchgeführt (Aachen, Berlin, Düsseldorf, Heidelberg, Lübeck, München, Münster, Tübingen und Würzburg) und an der LMU München durch ein von Prof. Greil geleitetes Forschungsteam koordiniert. Die Datenerhebung fand in den späten 1980er und frühen 1990er Jahren statt und umfasste eine Rekrutierungsphase von 2.5 Jahren, eine Stabilisierung der einzelnen Patienten von bis zu 6 Monaten sowie eine Behandlungsdauer von 2.5 Jahren. Die Patienten wurden während der Stabilisierungsphase auf Lithium, Carbamazepin oder Amitriptylin randomisiert und während der Beobachtungszeit von 2.5 Jahren in regelmäßigen Abständen zu einer ausführlichen Untersuchung einbestellt oder kontaktiert. Die Randomisierung erfolgte aufgrund vorgefertigter Listen im Studienzentrum. Die Studie wurde unverblindet durchgeführt.

#### **5.1.2 Diagnostische Gruppen und Studienmedikation**

Insgesamt wurden drei diagnostische Gruppen im Rahmen der MAP-Studie untersucht:

unipolare Depression (ICD-9: 296.1, 296.2, 296.3, 296.4, 295.7)

bipolar affektive Störung (ICD-9: 296.2, 296.3, 296.4)

schizoaffektive Psychose (ICD-9: 295.7)

Patienten mit einer unipolaren Depression wurden auf Lithium oder Amitriptylin randomisiert. Die Patienten der beiden anderen Gruppen erhielten entweder Lithium oder Carbamazepin als Studienmedikation. Das generelle Studiendesign der MAP-Studie ist in Abb. 53 zusammengefasst.



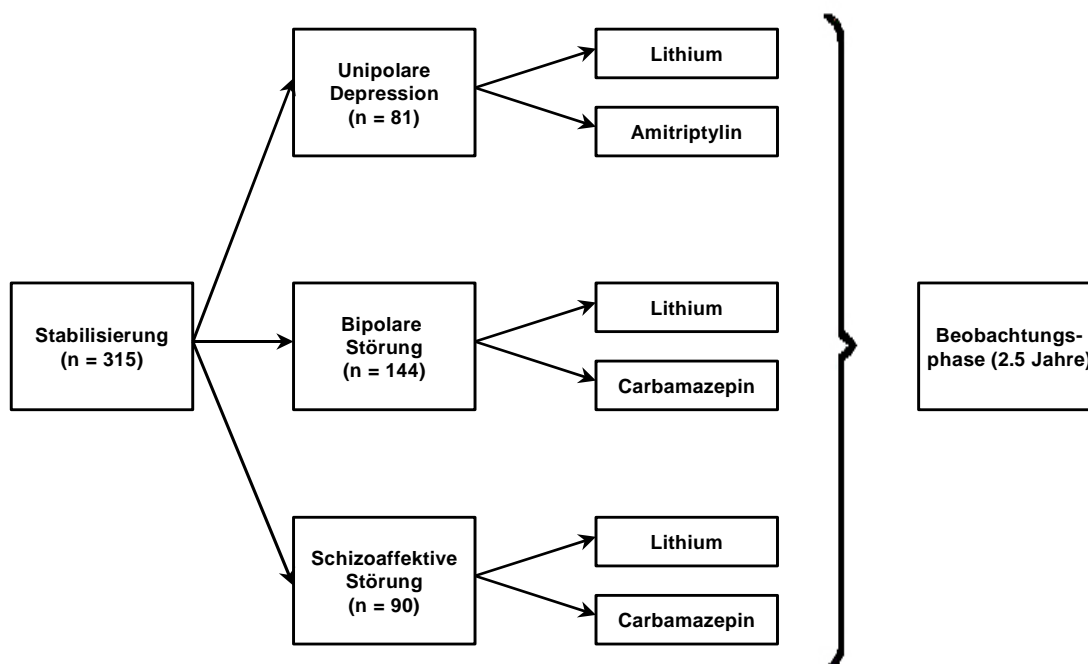


Abbildung 53. Design der MAP-Studie

Neben den Diagnosen nach ICD-9 wurden im Rahmen eines polydiagnostischen Ansatzes bereits während der Rekrutierung und bei Einschluss in die Studie Diagnosen nach RDC (Spitzer et al. 1982) und auf Grundlage eines SKID (Wittchen et al. 1987) Diagnosen nach DSM-III-R gestellt.

### 5.1.3 Rekrutierung der Studienpatienten

Da ein sorgfältig durchgeführter und dokumentierter Rekrutierungsprozess mit entscheidend für die Verallgemeinerbarkeit der Studienergebnisse ist, wurde hierauf in der MAP-Studie besonderer Wert gelegt. Patienten, die schließlich in die Studie aufgenommen wurden, wurden in einer eingehenden Untersuchung (Greil et al. 1993) mit den Nicht-Studienpatienten verglichen, die zwar prinzipiell für die Studie in Frage kamen, aber schließlich nicht auf eine der Studienmedikationen randomisiert wurden - sei es aufgrund von Ausschlusskriterien (z.B. mangelnde Sprachkenntnis) oder sei es, daß der Patient nicht an der Studie teilnehmen wollte. Im Ergebnis fanden sich in kritischen klinischen und soziodemographischen Variablen (wie z.B. Anzahl bisheriger Krankheitsphasen oder Alter) keine signifikanten Unterschiede zwischen Studienpatienten und Nicht-Studienpatienten (Greil et al. 1993).

Das Screening der Studienpatienten wurde in den jeweiligen Studienzentrum während der stationären Behandlung durchgeführt und umfasste alle hospitalisierten Patienten mit einem affektiven Syndrom. Für eine Aufnahme in die MAP-Studie wurden die folgenden Haupteinschlusskriterien gefordert:

1. Diagnose einer endogenen Depression, bisher nur monopolar, einer manisch-depressiven Psychose oder einer schizoaffektiven Psychose nach ICD-9
2. Mindestens eine frühere Episode in den vergangenen
  - 5 Jahren bei unipolarer Depression
  - 4 Jahren bei bipolarer affektiver Störung
  - 3 Jahren bei schizoaffektiver Psychose(Indikation einer Prophylaxe nach Angst (1981))
3. Keine Kontraindikationen gegen die Behandlung mit Lithium oder Amitriptylin bzw. gegen Lithium oder Carbamazepin
4. Keine prophylaktische Behandlung unmittelbar vor Einsetzen der derzeitigen Episode
5. Alter zwischen 18 und 65 Jahren
6. Kein Alkohol- oder Drogenmissbrauch (aktuell)
7. Einwilligung der Patienten nach Aufklärung ("informed consent").

#### **5.1.4 Stabilisierungsphase**

Um die prophylaktische Wirkung der Studienmedikationen zu untersuchen (vgl. Kap. 2.2.1), wurde zwischen Rekrutierung und Hauptuntersuchungsphase eine Stabilisierungsphase eingeschaltet. Hierdurch sollten die Patienten erst nach dem Ablauf der Indexepisode (aufgrund derer sie hospitalisiert waren) in die MAP-Studie eingeschlossen werden. Für die Stabilisierung war ein Zeitraum von maximal 6 Monaten angesetzt. Gemäß dem Studienprotokoll galt ein Patient als stabilisiert, falls er während dieser 6 Monate bei ausschleichender Dosierung psychotroper Medikation für mindestens zwei Wochen stabil war (GAF > 70). War dies der Fall, so wurde der Patient auf eine der Studienmedikationen randomisiert. An dieser Stelle erwies sich das Studienprotokoll als zu strikt - diese Vorgehensweise (Randomisierung nach Stabilisierung) konnte in der Umsetzung des Studienprotokolls nicht in jedem Fall durchgeführt werden. In der klinischen Praxis wird ein fließender Übergang von der Akuttherapie zur Erhaltungstherapie und schließlich zur Rezidivprophylaxe bevorzugt - insbesondere um negative Effekte des Ab- und Umsetzens von psychotropen Substanzen zu vermeiden. Dementsprechend wurde ein Teil der Patienten noch während des stationären Aufenthalts randomisiert und auf die Rezidivprophylaxe eingestellt. In einem Zeitrahmen von durchschnittlich vier Monaten konnten insgesamt 315 Patienten stabilisiert und auf eine der Therapiearme randomisiert werden.

### 5.1.5 Beobachtungsphase

Während der Beobachtungszeit von 2½ Jahren fanden vier Hauptuntersuchungen (T0-T3) statt:

- T0: Eingangsuntersuchung zu Studienbeginn
- T1: erste follow-up Untersuchung nach einem Jahr
- T2: zweite follow-up Untersuchung nach zwei Jahren
- T3: Abschlussuntersuchung nach 2½ Jahren

In der Eingangsuntersuchung erfolgte eine polydiagnostische Klassifikation des Krankheitsbildes (nach ICD-9, DSM-III-R und RDC), sowie eine umfangreiche körperliche und labortechnische Untersuchung. Zu allen Hauptuntersuchungszeitpunkten erfolgte ein ausführliches Rating der Psychopathologie, der Persönlichkeit, sowie von weiteren für die Krankheit und Befindlichkeit relevanten Daten. Hierfür wurden unter anderem die folgenden Messinstrumente verwendet (vgl. Tab. 7):

- Brief Psychiatric Rating Scale (BPRS) zur allgemeinen Erfassung der Psychopathologie
- Psychiatrische Anamnese (AMDP-100) zur allgemeinen Erfassung der Psychopathologie
- Bech-Rafaelsen Manie Skala (BRMaS) zur gezielten Erfassung manischer Syndrome
- Bech-Rafaelsen Melancholie Skala (BRMeS) zur gezielten Erfassung depressiver Syndrome
- Befindlichkeitsskala (BfS) zur Erfassung (hauptsächlich depressiver) Befindlichkeit
- Global Assessment of Functioning (GAF = GAS) zur globalen Beurteilung von Krankheits- und Funktionsgrad des Patienten
- Munich Personality Test (MPT = PPI) zur Erfassung der präorbiden Persönlichkeit
- Nebenwirkungsfragebogen
- Krankheitskonzept Skala (KK) zur Erfassung von für die Compliance wichtigen Einstellungen und Konzepten
- Globaleinschätzung der Compliance
- Social Interview Schedule (SIS) zur detaillierten Erfassung der Lebenssituation
- Camberwell Family Interview (CFI) zur Erfassung des emotionalen Familienklimas ("expressed emotions")
- Frühsymptome
- Visuelle Analogskalen (VAS) unter anderem zur Erfassung des Allgemeinbefindens, der Zufriedenheit mit der Therapie, der subjektiven Beeinträchtigung durch unerwünschte Arzneimittelwirkungen (UAW)
- Detaillierte Erfassung von unerwünschten Arzneimittelwirkungen (UAW)

Neben diesen Hauptkontakten fanden regelmäßige Zwischenkontakte statt. In den ersten 3 Monaten erfolgten diese Kontakte in 4-wöchigen, später in 8- bis 12-wöchigen Abständen. Hierbei wurden Veränderungen in Psychopathologie, Medikation, Laborbefunden und sozialer Situation sowie unerwünschter Arzneimittelwirkungen dokumentiert (vgl. Tab. 7). Zusätzlich wurde während dieser Haupt- und Zwischenkontakte jeweils der Krankheitsverlauf für die letzten Wochen retrospektiv erfragt (ein 4-stufiges Rating analog dem morbidity index für jede Woche). Diese Befragung bildet die Grundlage für die Beurteilung des Verlaufs.

Die Erhebungen wurden von einheitlich geschulten Ärzten und Psychologen durchgeführt. Sämtliche Daten wurden zeitnah nach der Erhebung gemäß den Standards der guten klinischen Praxis ("good clinical practice" - GCP) in der Münchner Studienzentrale auf Vollständigkeit und Plausibilität geprüft.

Tabelle 7: Erhebungen während der Beobachtungsphase der MAP-Studie (Auswahl).

	Rekrutierungsphase	T0	T1	T2	Zwischenkontakte
Diagnostik (ICD-9, RDC, DSM-III-R)	X				
Diagnostik (SKID für DSM III-R)		X			
Morbidity-Index (MI)		X	X	X	X
AMDP-100		X	X	X	X
BPRS		X	X	X	X
BRMeS, BRMaS		X	X	X	X
BfS		X	X	X	X
GAF		X	X	X	X
MPT		X	X	X	
KK		X	X	X	
Globaleinschätzung der Compliance		X	X	X	
SIS		X		X	
CFI (Familieninterview, war nicht strikt an das Erhebungsschema gebunden)					
Detaillierte Krankheitsanamnese	X				
AMDP-Anamnese		X			
Aktueller sozialer Status		X	X	X	X
VAS		X	X	X	X
Frühsymptome		X	X	X	X
Detaillierte Erfassung von UAW		X	X	X	X
Labor		X	X	X	X

## **5.2. Erhebungsinstrumente**

In der MAP-Studie wurde eine sehr umfangreiche Datenerhebung durchgeführt (vgl. Tab. 7). In der folgenden detaillierten Darstellung der Erhebungsinstrumente werden nur jene Instrumente aufgeführt, die im Rahmen dieser Auswertung von Bedeutung sind. Besonderer Wert wurde darauf gelegt, daß für alle im folgenden Kapitel 6 verwendeten Prädiktorvariablen ( $X_1$  bis  $X_{19}$ , vgl. Kap. 6.3.4.1) deutlich wird, mit welchen Messinstrumenten sie erhoben wurden.

### **5.2.1 Diagnostik**

Die Studie wurde im Rahmen eines polydiagnostischen Ansatzes durchgeführt. Die der vorliegenden Auswertung zugrunde liegenden Achse-I Diagnosen nach DSM-III-R wurden zu Beginn der Beobachtungsphase mit Hilfe des Strukturierten Klinischen Interviews (SKID-I für DSM-III-R, vgl. Wittchen et al. 1987) erhoben. Da die Symptome, aufgrund derer die Verschiebungen in der Diagnostik bipolarer Störungen beim Übergang von DSM-III-R zu DSM-IV vorgenommen wurden, durch den verwendeten SKID-I detailliert erfasst wurden, war es möglich, die DSM-III-R Diagnosen in exakte DSM-IV-Diagnosen zu übersetzen (vgl. Tab. 1, Kap. 2.1).

Ein Strukturiertes Klinisches Interview für Persönlichkeitsstörungen wurde nicht durchgeführt. Die Einschätzung, ob eine Persönlichkeitsstörung vorlag ( $X_5$ ) wurde während der Rekrutierungsphase vom Studienarzt vorgenommen, beruht also auf einer klinischen, nicht streng operationalisierte Diagnose.

### **5.2.2 Verlaufsbeurteilung**

Die in dieser Arbeit angewandte Verlaufsbeurteilung basiert auf den zu allen Haupt- und Zwischenkontakten retrospektiv erfassten 4-stufigen morbidity indices (MI) mit den Ausprägungen

0 = keine affektive Symptome

1 = nicht behandlungsbedürftige Symptome

2 = Symptome, die eine zusätzliche Behandlung erforderlich machen

3 = Symptome, die eine Hospitalisierung erforderlich machen

Der Verlauf wurde als „ungünstig“ beurteilt, falls während der Beobachtungszeit Symptome auftraten, die eine Hospitalisierung erforderlich machten (Ausprägung 3). Andernfalls galt der Verlauf im Rahmen dieser Arbeit als „günstig“. Für detailliertere Analysen zu den Verläufen, insbesondere zur Berücksichtigung von Intervallsymptomatik vgl. Buchberger (1997) und Kleindienst et al. (2002).

### **5.2.3 Drop-outs**

Im Falle eines vorzeitigen Studienabbruchs (drop-out) wurde ein detaillierter standardisierter drop-out Bogen im jeweiligen Studienzentrum ausgefüllt. Im Rahmen dieser Studie wurde keine outcome-relevante Klassifizierung der drop-out Gründe vorgenommen. Für eine einerseits kompakte und übersichtliche und andererseits informative Darstellung wurden folgende drop-out Gründe unterschieden:

- Non-compliance
- Unzureichende Wirksamkeit der Therapie
- Unerwünschte andere Wirkungen (UAWs)
- Schwangerschaft
- Weitere, vorwiegend externe Gründe

### **5.2.4 AMDP-100**

Im Rahmen der MAP-Studie wurde großer Wert auf eine detaillierte Erhebung der Psychopathologie gelegt. Mit dem Dokumentationssystem der Arbeitsgemeinschaft für Methodik und Dokumentation in der Psychiatrie (1979) wurde ein Breitband-Screening für 100 wichtige psychiatrische Symptome vorgenommen (AMDP-100). Diese Erhebung wurde zu Studienbeginn, sowie allen folgenden Haupt- und Zwischenuntersuchung durchgeführt (vgl. Tab. 7). Für eine fokussierte Erfassung der im Rahmen dieser Untersuchung im Vordergrund stehenden affektiven Syndrome und produktiv-psychotischer Syndrome wurde die AMDP-100 Skala durch die im folgenden beschriebenen Skalen (BPRS, BRMeS und BRMaS) ergänzt.

Die in der Literatur als Risikofaktor diskutierte Variable „deprimierte Stimmung oder Antriebsarmut“ ( $X_7$ ) wurde auf Grundlage der jeweils dichotomisierten items 63 („deprimiert“ – vorhanden vs. nicht vorhanden) und 80 („antriebsarm“ – vorhanden vs. nicht vorhanden) des zu Studienbeginn erhobenen AMDP-100 erfasst.

### **5.2.5 BPRS**

Die Brief Psychiatric Rating Scale (BPRS) ist ein besonders in der Schizophrenieforschung weit verbreitetes Instrument zur Erhebung psychopathologischer Symptomatik. Diese mehrdimensionale Skala eignet sich insbesondere zur Erfassung von Symptomen und Syndromen aus dem schizophrenen Formenkreis (vgl. Overall & Gorham 1970, CIPS 1981).

### **5.2.6 BRMeS und BRMaS**

Dem Fokus der MAP-Studie entsprechend wurde eine detaillierte Erfassung depressiver (BRMeS) und manischer (BRMaS) Symptomatik durchgeführt. Beide Skalen umfassen 11 in Fremdbeurteilung erhobene items. Sie erlauben eine Beschreibung sowohl der Schwere des aktuellen depressiven bzw. manischen Zustandes als auch eine Beschreibung des klinischen Bildes (vgl. Bech 1981).

### **5.2.7 BfS**

Die Befindlichkeitsskala (BfS) ist ein Selbstbeurteilungsbogen zur der Erfassung des subjektiven Empfindens des Patienten. Der Fokus liegt hierbei auf der momentanen Beeinträchtigung des Wohlbefindens. Insbesondere geht es um das Erkennen subtiler Veränderungen während der Remission oder im freien Intervall (vgl. von Zerssen 1973, 1981).

### **5.2.8 GAF**

Die GAF-Skala (Global Assessment of Functioning, auch GAS = Global Assessment Scale) dient der globalen Erfassung des psychosozialen Funktionsniveaus und der Leistungsfähigkeit psychiatrischer Patienten.

### **5.2.9 MPT**

Die Persönlichkeit wurde mit dem Münchner Persönlichkeits Test (MPT, auch PPI = Premorbid Personality Inventory) erfasst. Genauer gesagt zielt der MPT auf die prä-morbide Persönlichkeit, d.h. auf die Persönlichkeit – so wie sie vor dem Beginn der affektiven Störung bestand. Hierzu wird der Patienten ausdrücklich darauf hingewiesen, daß nur nach der psychisch gesunden Zeit gefragt wird ("usual self"). Bei der Erhebung beantwortet der Patient



in einem Selbstrating 51 items, die aufgrund von Faktorenanalysen 6 Hauptbereichen der Persönlichkeit zugeordnet werden:

- Neurotizismus
- Extraversion
- Frustrationstoleranz
- Rigidität
- Schizoidie
- Soziale Erwünschtheit

Diese 6 Faktoren werden durch eine Kontrollskala („Lügenskala“), Motivation, ergänzt. Der Faktor Soziale Erwünschtheit entspricht einerseits dem Bestreben nach sozialer Anerkennung und Billigung und kann daher als Persönlichkeitsmerkmal aufgefasst werden. Andererseits kann dieses Streben nach Anerkennung das gesamte Antwortverhalten beeinflussen (falls der Patient bestrebt ist, sich bei seinen Antworten in ein gutes Licht zu rücken). Daher erfüllt die im MPT erfasste Dimension der sozialen Erwünschtheit die Funktion einer zweiten Kontrollskala (vgl. von Zerssen 1988).

Im Rahmen dieser Auswertung ist vor allem der Neurotizismus-Score zu Studienbeginn ( $X_6$ ) von Interesse, da sich ein hoher Neurotizismus-Score aufgrund der Metaanalysen als Risikofaktor für die Lithiumresponse erhärtet hat (vgl. Kap. 4.5.9). Der Neurotizismus-Score umfasst insbesondere die folgenden Bereiche: Labilität, Überempfindlichkeit, starke Tendenz zur Niedergeschlagenheit bei Mißerfolgen, starke Abhängigkeit von der Beurteilung durch andere Personen, Angst vor Ablehnung.

### **5.2.10 KK**

Die Krankheitskonzepte wurden aufgrund der Krankheitskonzeptskala (KK-Skala) erhoben. Die KK-Skala umfasst 29 items, die aufgrund von Faktorenanalysen den folgenden 7 Dimensionen zugeordnet wurden (vgl. Linden et al. 1988):

- Arztvertrauen
- Medikamentenvertrauen
- Negativerwartungen
- Schuld
- Zufallskontrolle
- Anfälligkeit
- Idiosynkratische Annahmen

Neben einer Erfassung einzelner Krankheitskonzepte und von Erwartungen wurde von Linden et al. (1988) ein KK-Summenscore entwickelt, mit dem die Bereitschaft des Patienten zur Kooperation, und damit die erwartete Compliance erfasst werden kann. Der KK-Summenscore zu Studienbeginn wurde im Ergebnisteil als Prädiktor der Lithiumresponse getestet ( $X_{17}$ ). Zwar wird die tatsächliche Compliance durch die im folgenden Kapitel (5.2.11) besprochene "Globaleinschätzung der Compliance" direkter erfasst als durch die KK-Skala, allerdings eignet sich nur eine Einschätzung der erwarteten Compliance zur Prognose des Erfolgs einer Lithiumtherapie, nicht die später tatsächlich beobachtete Compliance. Daher wurde die Prädiktion der Lithiumresponse aufgrund des KK-Summscores zu Studienbeginn vorgenommen.

### **5.2.11 Globaleinschätzung der Compliance**

Die jeweils aktuelle Medikamentencompliance und Behandlungsmotivation wurde aufgrund von zwei 5-stufigen Lickert-Skalen zu den 4 Hauptuntersuchungszeitpunkten vom jeweiligen Studienarzt eingeschätzt.

### **5.2.12 SIS**

Mit dem Social Interview Schedule (SIS) werden für wichtige Lebensbereiche (Wohnen, Einkommen, Freizeit, Kontakte, ...) die objektive Situation, sowie das Zurechtkommen (Management) und die Zufriedenheit mit diesen Lebensbereichen erfasst. Die Erhebung beruht auf einem ausführlichen, methodisch sehr heterogenen Interview.

Von den im Ergebnisteil untersuchten Prädiktoren für die Lithiumresponse basieren zwei auf dem SIS: Das Vorliegen von Stress ( $X_{14}$ ) und die Soziale Unterstützung ( $X_{18}$ ). Soziale Unterstützung ist hierbei die unproblematischere Variable – das SIS enthält ein direktes 4-stufiges Fremdrating des „Social Support“, das für die Auswertung in dichotomisierter Form (angemessen vs. leicht-stark unangemessen) verwendet wurde. Problematischer ist die Operationalisierung des Konstrukts Stress. Auf diese Variable wurde im Rahmen der MAP-Studie nicht fokussiert. Der globale Management-Score über alle Bereiche, d.h. das Zurechtkommen mit wichtigen Lebensbereichen kam diesem Konstrukt am nächsten und wurde daher zur Beurteilung des Ausmaßes an Stress verwendet. Für eine ausführliche Beschreibung des SIS und die Ergebnisse einer psychometrischen Testung vgl. Faltermaier et al. (1985).

### 5.2.13 CFI

Für einen Teil der Patienten wurde mit mindestens einer relevanten Bezugsperson ein halbstandardisiertes Interview, das Camberwell Family Interview (CFI) zur Erfassung des emotionalen Familienklimas durchgeführt. Das CFI zielt insbesondere auf die Dimensionen "Kritische Bemerkungen" und "Emotionales Überengagement", die sich als prognostisch ungünstig erwiesen haben. Liegt in einer dieser beiden Dimensionen ein Wert über einem empirisch definierten Schwellenwert vor (vgl. Vaughn & Leff 1976, 1981), so spricht man von "high expressed emotions (high EE)". Die im Rahmen der MAP-Studie erhobene Variable high expressed emotions ( $X_{15}$ ) entspricht exakt dieser Definition.

### 5.2.14 Detaillierte Krankheitsanamnese

Eine detaillierte somatische wie psychiatrische Krankheitsanamnese wurde vom jeweiligen Projektarzt während der Rekrutierungsphase erhoben. Die üblicherweise während des Klinikaufenthaltes erstellten Krankengeschichte wurde durch eine standardisierte Erfassung jeder einzelnen affektiven Krankheitsphase ergänzt. Das Erhebungsschema wurde in der Münchner Studienzentrale entwickelt. Insbesondere wurden für jede rekonstruierbare affektive Phase insbesondere die folgenden Charakteristika erfasst:

- Beginn und Ende der Phase
- vorherrschende Symptomatik (detaillierte Erfassung mit folgenden Merkmalsausprägungen:
  - 1) depressiv, 2) manisch, 3) manisch-depressiver Mischzustand, 4) Wechsel von der Manie in die Depression, 5) Wechsel von der Depression in die Manie, 6) mehrere Wechsel, 7) Wechsel – Reihenfolge unklar, 8) andere Symptomatik ohne nähere Angaben, 9) schizophorenform, 10) schizoaffektiv, manisch, 11) schizoaffektiv, depressiv, 12) schizoaffektiv, gemischt oder zyklisch, 13) schizoaffektiv, unklar
- Phase psychotisch? ... mit den folgenden Merkmalsausprägungen: 1) psychotisch, stimmungskongruente Wahninhalte, 2) psychotisch, stimmungsinkongruente Wahninhalte, 3) psychotisch, unklar
- Schweregrad
- Hospitalisierung nötig?
- Qualität des anschließenden Intervalls

Aufgrund dieser detaillierten standardisierten Erfassung aller erinnerten Krankheitsphasen war es möglich, in der MAP-Studie die prognostische Kraft der folgenden anamnestischen Variablen zu untersuchen:

- (1) Die Anzahl stationärer Episoden in der Vorgeschichte ( $X_1$ ). Die Indexepisode wird hierbei nicht mitgezählt.
- (2) Die Krankheitsdauer ( $X_2$ ).
- (3) Das Vorliegen stimmungsinkongruente Wahninhalte in der Anamnese ( $X_3$ ). Bei dieser Variablen wird geprüft, ob im Rahmen irgendeiner affektiven Episode stimmungsinkongruente Wahninhalte aufgetreten sind. Da ein Schwerpunkt bei der Diagnostik auf der Unterscheidung zwischen affektiven und schizoaffektiven Störungen lag, wurde diese Variable besonders sorgfältig erfasst.
- (4) Die Verlaufsmuster MDI ( $X_5$ ), DMI ( $X_6$ ) und CC ( $X_7$ ). Dieses Verlaufsmuster bezieht sich auf die der Indexepisode unmittelbar vorausgehende Krankheitsphase. Es ist zu erwarten, daß diese jüngste anamnestische Phasen mit Abstand am besten erinnert wird und daher die zuverlässigsten Angaben enthält. Vor allem aus diesem Grunde wurde die jüngste anamnestische Phase als für alle Phasen prototypisch angenommen.
- (5) Auch für die Beurteilung, ob isolierte Krankheitsepisoden ( $X_{11}$ ) vorliegen, wurde die der Indexepisode unmittelbar vorausgehende Krankheitsphase herangezogen.
- (6) Das Alter bei Ersterkrankung ( $X_{12}$ ) ließ sich aus einem Vergleich des Datums der ersten Episode und des Geburtsdatums errechnen.
- (7) Die Phasenfrequenz ( $X_{13}$ ) wurde berechnet als Quotient aus der Anzahl bisheriger Krankheitsphasen und der Zeit seit der Ersterkrankung.
- (8) Soweit die Krankengeschichte ausreichend rekonstruiert werden konnte, ließ sich auch die Länge des ersten freien Intervalls ( $X_{19}$ ) direkt bestimmen.

### 5.2.15 AMDP-Anamnese

Die detaillierte Krankheitsanamnese wurde durch eine umfassende, dem Dokumentationssystem der Arbeitsgemeinschaft für Methodik und Dokumentation in der Psychiatrie (1979) entsprechende Anamnese ergänzt. Hierbei wurden unter anderem soziodemographische Variablen wie Familienstand, Ausbildung, etc. erhoben. Diese Anamnese umfasste auch eine strukturierte Erfassung kritischer Lebensereignisse mit 13 items zu vermuteten krankheitsfördernden Ereignissen und 32 items zu Veränderungen der Lebenssituation. Die items zu den vermuteten krankheitsfördernden Einflüssen wurden sowohl vom Arzt als auch vom Patienten eingeschätzt. Die vom Patienten als zutreffend eingeschätzten items wurden aufaddiert und bei der Auswertung als "life events" ( $X_{16}$ ) bezeichnet.

### **5.2.16 Aktueller sozialer Status**

Zu allen Haupt- und Zwischenkontakten wurde der aktuelle soziale Status des Patienten kurz erfasst (12 items). Im Rahmen dieser Erhebung wurde auch erhoben, ob der Patient aktuell beim Arbeitsamt als arbeitslos gemeldet ist. Arbeitslosigkeit zu Studienbeginn wurde bei der Auswertung als Prädiktor in Betracht gezogen ( $X_8$ ).

### **5.2.17 Weitere Erhebungsinstrumente**

Neben den hier beschriebenen Erhebungsinstrumenten wurden eine Reihe weiterer Messungen, Fremd- und Selbstratings durchgeführt. Diese umfassen insbesondere eine Reihe von Laboruntersuchungen, eine detaillierte Erhebung von unerwünschten Arzneimittelwirkungen und von Frühsymptomen sowie eine Reihe regelmäßig vom Patienten ausgefüllten visuellen Analogskalen (VAS), etwa zur globalen Zufriedenheit mit der Behandlung. Diese Erhebungsinstrumente sind Gegenstand früherer und aktueller Auswertungen, spielen aber im Rahmen dieser Arbeit keine wesentliche Rolle.

## 6 Ergebnisse aus der MAP-Studie

### 6.1 Studienpatienten

Insgesamt erhielten 171 Patienten die SKID-Diagnose einer bipolaren Störung nach DSM-IV (vgl. Kap. 5.2.1). 86 dieser Patienten wurden auf Lithium randomisiert, 85 auf Carbamazepin. Die ab dem zweiten Behandlungsmonat über die restliche Studienzeit gemittelte Tagesdosis für Lithium lag bei  $26.8 \pm 6.76$  mmol. Die entsprechende Dosis für Carbamazepin lag bei  $635 \pm 190$  mg. Dies entsprach einem gemittelten Plasmaspiegel von  $0.61 \pm 0.12$  mmol/l für Lithium und von  $6.12 \pm 1.27$  µg/ml für Carbamazepin.

Beim Vergleich demographischer und klinischer Variablen (vgl. Tabelle 8) zeigten sich keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsgruppen. Der Vergleich demographischer Variablen umfasste Alter und Geschlecht. Der Vergleich klinischer Faktoren beinhaltete eine Vielzahl grundlegender Variablen wie etwa die diagnostische Untergruppe, eine genaue Charakterisierung der Indexepisode und die Anzahl bisheriger Episoden. Beim Vergleich der prämorbidem Persönlichkeit zeigte sich in der Carbamazepin-Gruppe ein signifikant höherer Wert auf der Extraversions-Skala ( $13.4 \pm 6.4$  vs.  $11 \pm 6.4$ ,  $p=0.0089$ ). Alle anderen Faktoren des MPT, Neurotizismus, Frustrationstoleranz, Rigidity und Schizoidie zeigten keinen signifikanten Unterschied zwischen den beiden Behandlungsgruppen.

Tabelle 8: Demographische, klinische and psychologische Variablen zu Studienbeginn

Variable	Lithium (n=86)	Carbamazepin (n=85)	p-Wert
Alter (Mean±SD)	41±13	39±13	p=0.27
Geschlecht (% weiblich)	55%	58%	p=0.76
Diagnostische Untergruppe (% bipolar I)	67%	66%	p=0.87
Typ der Indexepisode (% Manie oder Mischzustand)	31%	23%	p=0.23
Stimmungsinkongruente Wahnhinhalte während der Indexepisode (%)	26%	22%	p=0.72
Anzahl bisheriger Episoden (Mean±SD)	4.3±2.3	4.1±2.2	p=0.34
GAF-Score zu Studienbeginn (Mean±SD)	79±10	79±10	p=0.62
Psychiatrische Komorbidität (hauptsächlich Lebenszeitdiagnose einer Angststörung oder eines Substanzmissbrauchs)	16%	16%	p=1.00
Neurotizismus	10±5.3	10±5.6	p=0.97
Frustrationstoleranz	6.6±3.9	6.9±4.1	p=0.72
Extraversion	11±6.6	13±5.7	p=0.01
Rigidität	10±4.4	11±4.6	p=0.21
Schizoidie	3.0±1.6	2.7±1.9	p=0.13

## 6.2 Therapieabbruch - Häufigkeit und Gründe

Bei 36% der Patienten (61 von 171) kam es zu einem vorzeitigen Studienabbruch. Unter Carbamazepin war das Abbruchrisiko deutlich höher als unter Lithium (relatives Risiko = 1.79,  $p = 0.0067$ ). Ein Vergleich der Gründe, die zu einem Therapieabbruch führten, ergab keinen signifikanten Unterschied zwischen den beiden Therapiegruppen ( $\chi^2(4) = 1.809$ ,  $p = 0.77$ ). In beiden Gruppen war Non-compliance der häufigste drop-out Grund, gefolgt von

einem Abbruch der Therapie aufgrund unzureichender Wirksamkeit und UAWs (vgl. Tabelle 9).

Tabelle 9: Gründe für einen vorzeitigen Therapieabbruch (drop-out):

Drop-out Grund	Lithium (n=86)	Carbamazepin (n=85)
Non-compliance	12	15
Unzureichende Wirksamkeit der Therapie	4	8
Unerwünschte andere Wirkungen (UAWs) <sup>1</sup>	3	8
Schwangerschaft	1	4
Weitere, vorwiegend externe Gründe <sup>2</sup>	2	4
Gesamt	22	39

<sup>1</sup>Unter Lithium: Akne und Gewichtszunahme (1 Patient), Potenzstörungen (1), Schwindel, Übelkeit und Kopfschmerz (1). Unter Carbamazepin: Exanthem, Lymphom und Kopfschmerz (1 Patient), Allergie und Ausschlag (1), generalisiertes Ekzem (1), allergische Hautreaktion (1), Hepatopathie (1), Lymphom und Diarrhoe (1), Exanthem (2).

<sup>2</sup>Z.B. Wegzug, oder versehentliche Verletzung des Studienprotokolls durch den Rater.

### 6.2.1 Zeitpunkt des Therapieabbruchs

Unter Lithium fand bei 11 der 22 Patienten mit vorzeitigem Therapieabbruch der drop-out vor einem (möglichen) stationären Rezidiv statt und bei den anderen 11 Patienten nach oder während der Rehospitalisierung. Für diese letzten 11 Patienten war eine vollständige Beurteilung des Gesamtverlaufs trotz des drop-outs möglich und dementsprechend gingen sie auch in die Analysen „according to protocol“ (ATP) wie z.B. die logistische Regressionen oder Diskriminanzanalysen ein. Die „ATP-Quote“ in der Lithiumgruppe betrug demnach 87% (75/86).

Unter Carbamazepin fanden die meisten der drop-outs zu einem frühen Zeitpunkt statt. In dieser Gruppe ereignete sich insgesamt bei 29 von insgesamt 39 Studienabbrechern der drop-out vor einem möglichen stationären Rezidiv, so daß diese Patienten nicht in den ATP-Analysen nicht berücksichtigt werden konnten. Insgesamt beziehen sich die ATP-Analysen also nur auf 66% der Carbamazepinpatienten (56/85).



## 6.3 Response-Prädiktion für Lithium

### 6.3.1 Auswertungsstrategien

Bei der Auswertung der MAP-Studie werden drei Modelltypen zur Prädiktion des Verlauf unter Lithium verwendet (vgl. Abb. 54):

- 1) Univariate Modelle (Kap. 6.3.2)
- 2) Klassische multivariate Modelle (Kap. 6.3.3)
- 3) „Additiv univariate Modelle“ (Kap. 6.3.4)

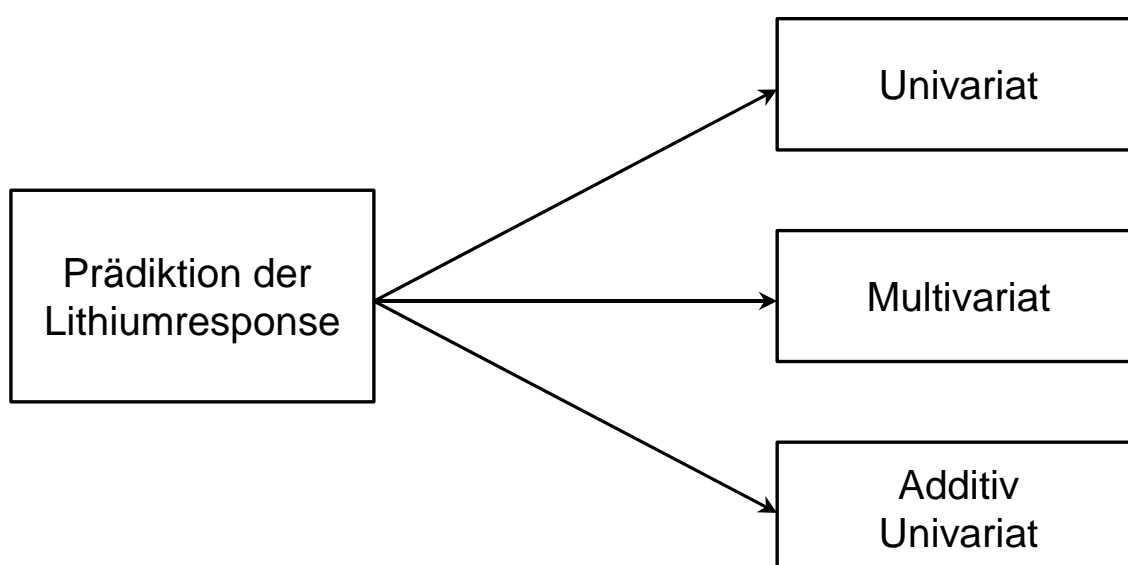


Abbildung 54. Strategien zur Prädiktion des Verlaufes unter Lithium.

Zu 1) Einfache univariate Modelle wie die Korrelation zwischen Therapieresponse und den verschiedenen Einflussvariablen entsprechen genau den in der Metaanalyse dargestellten Ergebnissen. Sie dienen der Modellierung des isoliert betrachteten Einflusses einzelner Prädiktoren auf die Lithium-Response. Wie in der vorgehenden Metaanalyse werden Korrelationen zwischen Prädiktorvariablen und Response berechnet und auf Signifikanz geprüft. Je nach Skalenniveau der Prädiktorvariablen (metrisch, ordinal oder dichotom) wurde eine Punkt-biseriale Korrelation, eine biseriale Rangkorrelation bzw. ein Phi-Koeffizient berechnet. Univariate Modelle haben den Nachteil, immer nur den Einfluss eines Prädiktors isoliert zu betrachten. Dies führt nicht nur zu einer Inflationierung des Fehlers 1. Art, sondern wird konzeptionell der multifaktoriellen Bedingtheit des Ansprechens auf eine Medikation wie Lithium nicht gerecht. Zum einen werden die sich überlagernden Einflüsse verschiedener Prädiktoren bei isolierter Betrachtung nicht berücksichtigt. Zum anderen werden mögliche Interaktionen zwischen verschiedenen unabhängigen Variablen nicht beachtet. Dennoch

erscheint es sinnvoll, univariate Ergebnisse zu berichten, weil die Kenntnis univariater Zusammenhänge ein erster wichtiger Schritt beim Aufbau komplexerer Modelle ist. Univariate Ergebnisse bieten auch den Vorteil, daß sie relativ leicht über verschiedene Studien hinweg verglichen und z.B. im Rahmen von Metaanalysen aggregiert werden können. Einige wesentliche Vor- und Nachteile univariater Methoden sind in Tabelle 10 zusammengestellt.

Zu 2) Konzeptionell die beste unter den hier vorgestellten Alternativen sind multivariate Modelle (z.B. logistische Regression, Diskriminanzanalyse oder Cox-Regression). Sie modellieren den simultanen Einfluss verschiedener Prädiktoren unter Berücksichtigung von Interaktionen. Dies entspricht genau dem in der Natur der unterstellten Zusammenhänge zwischen einem Bündel erklärender Variablen und dem Ansprechen auf Lithium. Dementsprechend wird in der Literatur immer wieder gefordert, die Response-Prädiktion auf multivariate Modelle zu basieren (z.B. Grof et al. 1993, S.16: "Statistical analysis should be commensurate with the complexity of the issue [sc. der Prädiktion der Lithium-Response], which is multivariate in nature.").

Alles scheint für multivariate Verfahren zu sprechen. Aber diese Verfahren haben auch Ernst zu nehmende Nachteile:

i) Zunächst bestehen erhebliche technische Schwierigkeiten beim Umgang mit fehlenden Werten. Entweder müssen fehlende Werte geschätzt und aufgefüllt werden, oder ein einziger fehlender Wert führt dazu, daß eine komplette Beobachtung von der multivariaten Analysen ausgeschlossen wird. Fehlende Werte sind in der psychiatrischen Forschung (und Praxis) jedoch die Regel.

ii) Von der statistischen Theorie her ist eine adäquate Modellierung des Einflusses von Variablen im Kontext anderer Variablen und eine richtige Modellierung nur dann zu erwarten, wenn das Modell (z.B. eine logistische Regression) vollständig und korrekt spezifiziert ist (vgl. z.B. Schneeweiß 1990, S. 148ff). Dies bedeutet, daß alle Variablen einer Erklärungsebene im Modell enthalten sein müssen (Vollständigkeit), und daß die funktionellen Zusammenhänge (z.B. Linearität) im Modell korrekt wiedergegeben sind. Kritisch ist vor allem die in der psychiatrischen Forschung nie gegebene Vollständigkeit. In der Forschungspraxis bedeutet dies, daß insbesondere bei untereinander stark korrelierenden erklärenden Variablen der Einfluss mancher Prädiktoren völlig falsch modelliert werden kann.

iii) Ein zu wenig beachteter Nachteil multivariater Methoden resultiert gerade aus ihrer großen Flexibilität. Dadurch, daß insbesondere bei der simultanen Berücksichtigung vieler Variablen eine Vielzahl von Parametern an den zur Modellbildung verwendeten Datensatz angepasst werden, kommt es bei zur Prädiktion verwendeten Verfahren wie der logistischen Regression

zu unrealistisch hohen Trefferquoten, die sich bei der praktischen Anwendung der Ergebnisse nicht mehr replizieren lassen. Diesem Problem lässt sich jedoch weitgehend dadurch begegnen, daß eine Kreuzvalidierung (Trennung der zur Modellbildung verwendeten Daten und der aufgrund des so ermittelten Modells zugeordneten Beobachtungen). Es bleibt jedoch zu beachten, daß auch bei Verwendung einer Kreuzvalidierung die Gefahr besteht, daß ein multivariates Modell zu spezifisch auf die Eichstichprobe optimiert wurde und bei der Anwendung auf einen anderen Datensatz sehr enttäuschende Vorhersagen liefert.

Zu 3) Gewissermaßen in der Mitte zwischen univariaten und multivariaten Modellen stehen „additiv univariate Modelle“. Sie sind dadurch charakterisiert, daß der Einfluss mehrerer erklärender Variablen gleichzeitig erfasst wird (wie in multivariaten Modellen), aber die Interaktionen zwischen diesen Variablen nicht modelliert werden (wie in univariaten Modellen). Der konzeptionelle Nachteil, die Variableninteraktionen nicht zu modellieren bietet mathematisch-technische Vorteile. Insbesondere eine größere Robustheit bei der Anwendung dieser Modelle auf neue Datensätze (also z.B. bei der Anwendung in der klinischen Praxis) und bei der Behandlung fehlender Werte. So können die Variablen von Beobachtungen mit fehlenden Werten direkt zur Modellbildung herangezogen werden und auch bei der Modellanwendung ist es ohne weiteres möglich, den Verlaufs von Patienten zu prognostizieren, über die nur unvollständige Informationen vorliegen.

Aufwendigere Verfahren dieses Typs (wie z.B. eine Diskriminanzanalyse, bei der anstatt der Varianz-Kovarianz-Matrix der unabhängigen Variablen nur die Diagonalmatrix der Varianzen modelliert wird), werden nur sehr selten verwendet. Ein sehr einfaches und v.a. in der Diagnostik und Erfassung von Syndromen weit verbreitetes Verfahren dieses Typs sind Skalen. Hierbei werden verschiedene items erfasst und zu einem Summenscore addiert. Vor allem in Hinblick auf die praktische Verwertbarkeit der Ergebnisse wird im Rahmen dieser Arbeit auf aufwendige additiv univariate Modelle zugunsten einer einfach anzuwendenden Skala verzichtet. Im Vergleich mit den üblicherweise verwendeten Skalen (z.B. YMRS oder die für in der Prädiktion verwendete Newcastle-Scale) wird die in dieser Arbeit definierte „Lithium-Response-Skala“ so konstruiert, daß der Summenscore auch bei fehlenden Werten definiert und aussagekräftig ist (vgl. Kap. 6.3.4).

Tab. 10 Vorteile und mögliche Nachteile der unterschiedlichen Modelltypen

	Vorteile	Nachteile oder Schwierigkeiten
Univariate Modelle	<ul style="list-style-type: none"> <li>• Direkte Verwendbarkeit der Ergebnisse in folgenden Metaanalysen und in der klinischen Praxis</li> </ul>	<ul style="list-style-type: none"> <li>• Isolierte Betrachtung, fehlende Integration des Einflusses verschiedener Einflussvariablen</li> <li>• Inflationierung des <math>\alpha</math>-Fehlers durch eine Vielzahl von Analysen</li> </ul>
Multivariate Modelle	<ul style="list-style-type: none"> <li>• Simultane Betrachtung des Einflusses mehrerer Variablen unter Berücksichtigung der Interaktionen zwischen den Einflussvariablen</li> </ul>	<ul style="list-style-type: none"> <li>• Erhebliche Schwierigkeiten beim Auftreten fehlender Werte</li> <li>• Erhöhte Gefahr der Fehleinschätzung des Einflusses bestimmter Prädiktoren in unvollständig spezifizierten Modellen</li> <li>• Überschätzung der Prädiktionskraft der Modelle aufgrund einer unrealistisch guten Anpassung an den zur Modellbildung verwendeten Datensatz</li> </ul>
„Additiv univariate Modelle“	<ul style="list-style-type: none"> <li>• Große Robustheit der Ergebnisse</li> <li>• Flexibler und einfacher Umgang mit fehlenden Werten</li> <li>• Integration des Einflusses verschiedener Einflussvariablen</li> </ul>	<ul style="list-style-type: none"> <li>• Fehlende Berücksichtigung der Variableninteraktionen</li> </ul>

### 6.3.2 Univariate Analysen für einzelne Prädiktoren

Im folgenden wird für jeden einzelnen Prädiktor der Einfluss auf den Verlauf unter Lithium untersucht. Je nach Datenstruktur werden hierfür entweder Punkt-biseriale Korrelationen oder Korrelationen zwischen zwei dichotomen Variablen (=  $\phi$ -Koeffizienten) berechnet. Die Kodierung der Variablen wurde durchgehend so gewählt, daß positive Korrelationen für eine positive Beeinflussung des Verlaufs stehen.

Für die graphische Darstellung wurden Histogramme und Box-Plots gewählt. Die Darstellung durch Box-Plots dient auch der visuellen Kontrolle von Ausreißern. Falls extreme Ausreißer aufgetreten sind (Abstand mehr als 3.5 Interquartilsdistanzen vom Median), wurden alle Analysen nach Ausreißerbeseitigung wiederholt (vgl. Kap. 6.3.2.19). Insgesamt kam es zu keinen Diskrepanzen zwischen den auf den Originaldaten beruhenden Ergebnissen und den Ergebnissen nach Ausreißerbeseitigung. Bei dichotomen Variablen wurden Histogramme verwendet. Hierbei wurde jeweils die Häufigkeit einer Merkmalsausprägung (wie z.B. das Vorhandensein einer Persönlichkeitsstörung) für die Gruppe mit günstigem und mit ungünstigem Verlauf verglichen.

### 6.3.2.1 Anzahl stationärer Episoden

In der MAP-Studie war eine hohe Zahl bisheriger affektiver Episoden, die zu einer Hospitalisierung führten, tendenziell prognostisch ungünstig. Die Punkt-biserial Korrelation zwischen der Anzahl stationärer Rezidive im bisherigen Krankheitsverlauf und einem günstigen Verlauf (keine Rehospitalisierung) unter Lithium betrug  $r = -0.24$ , ist allerdings statistisch nicht signifikant ( $p = 0.06$ ,  $n = 74$ , vgl. Abb. 55).

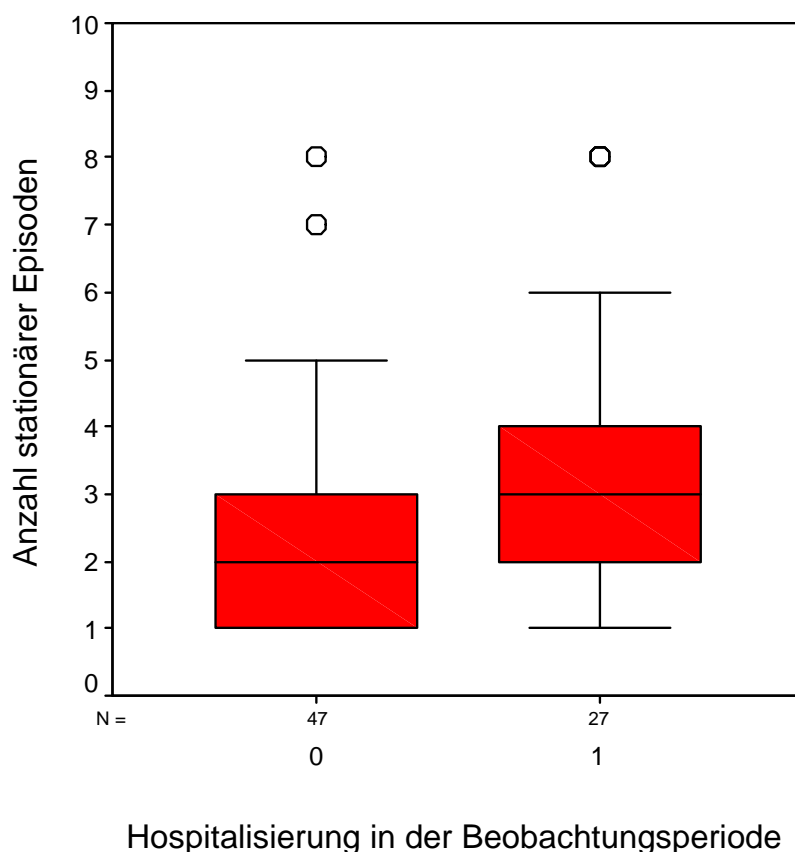


Abbildung 55. Anzahl stationärer Episoden bei Lithiumpatienten ohne bzw. mit Hospitalisierung in der Beobachtungsperiode ( $n = 47$  vs.  $n = 27$ ).

Erwartungsgemäß zeigt sich ein ähnlicher Trend für die Carbamazepingruppe. Auch hier fand sich eine Korrelation zwischen der Anzahl stationärer Rezidive im bisherigen Krankheitsverlauf und einem günstigen Verlauf von  $r = -0.24$  ( $p = 0.09$ ,  $n = 52$ ).

Zusammenfassend lässt sich sagen, daß eine hohe Anzahl affektiver Episoden in der Vorgeschichte tendenziell ein für Lithium unspezifischer Negativprädiktor für den Verlauf ist. Dieses Ergebnis lässt sich allerdings nicht statistisch absichern und ist bei isolierter Betrachtung daher sehr zurückhaltend zu interpretieren.

### 6.3.2.2 Krankheitsdauer

Sowohl unter Lithium wie unter Carbamazepin fand sich kein signifikanter Zusammenhang zwischen der Krankheitsdauer und einem günstigen Verlauf ( $r = 0.13$ ,  $p = 0.32$ ,  $n = 75$  unter Lithium (vgl. Abb. 56) und  $r = -0.04$ ,  $p = 0.74$ ,  $n = 56$  unter Carbamazepin).

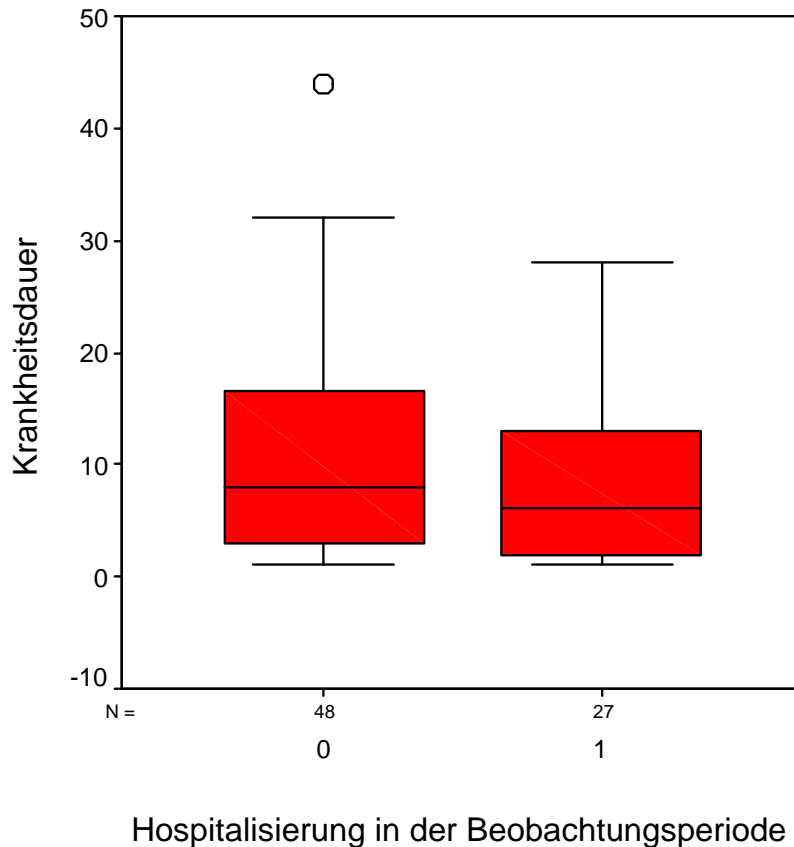


Abbildung 56. Anzahl stationärer Episoden bei Lithiumpatienten ohne bzw. mit Hospitalisierung in der Beobachtungsperiode ( $n = 47$  vs.  $n = 27$ ).

### 6.3.2.3 Stimmungsinkongruente Wahninhalte in der Anamnese

Unter keinem der beiden Medikamenten, weder unter Lithium noch unter Carbamazepin, zeigte sich ein signifikanter Einfluss stimmungsinkongruenter Wahninhalte in der Anamnese auf den Verlauf ( $r = 0.06$ ,  $p = 0.63$ ,  $n = 75$  unter Lithium (vgl. Abb. 57) und  $r = -0.07$ ,  $p = 0.51$ ,  $n = 56$  unter Carbamazepin). Bei diesem Ergebnis ist allerdings zu beachten, daß diese Variable nur routinemäßig mit dem des AMDP-100 erhoben wurde und aufgrund dieser wenig sensitiven Erhebung mit der Möglichkeit falsch negativer Resultate zu rechnen ist.

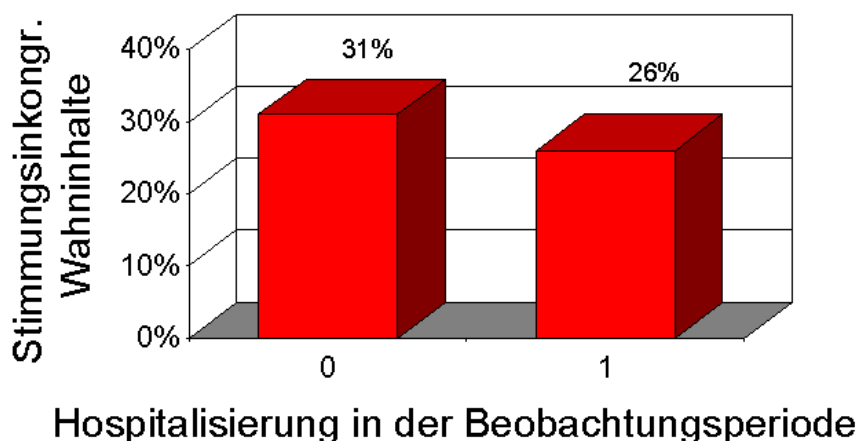


Abbildung 57. Anteil von Lithiumpatienten mit stimmungsinkongruenten Wahninhalten in den Gruppen ohne bzw. mit Hospitalisierung ( $n = 48$  vs.  $n = 27$ ).



#### 6.3.2.4 Verlaufsmuster MDI

Das Verlaufsmuster MDI (Manie – Depression – freies Intervall) war signifikant mit einem guten Lithiumresponse korreliert ( $r = 0.24$ ,  $p = 0.04$ ,  $n = 75$ , vgl. Abb. 58). Unter Carbamazepin hingegen fand sich kein signifikanter Zusammenhang zwischen dem Verlaufsmuster MDI und der Response ( $r = -0.180$ ,  $p = 0.18$ ,  $n = 56$ ). Der Unterschied zwischen den Punkt-biserialen Korrelationskoeffizienten unter Lithium und Carbamazepin ist statistisch signifikant ( $p = 0.04$ ). Dies spricht dafür, daß das Verlaufsmuster MDI ein für Lithium spezifischer Positivprädiktor ist. Einschränkend ist allerdings zu sagen, daß für insgesamt nur  $n = 7$  auf Lithium randomisierte Patienten ein Verlaufsmuster MDI festgestellt wurde, und die Ergebnisse daher vorsichtig interpretiert werden sollten.

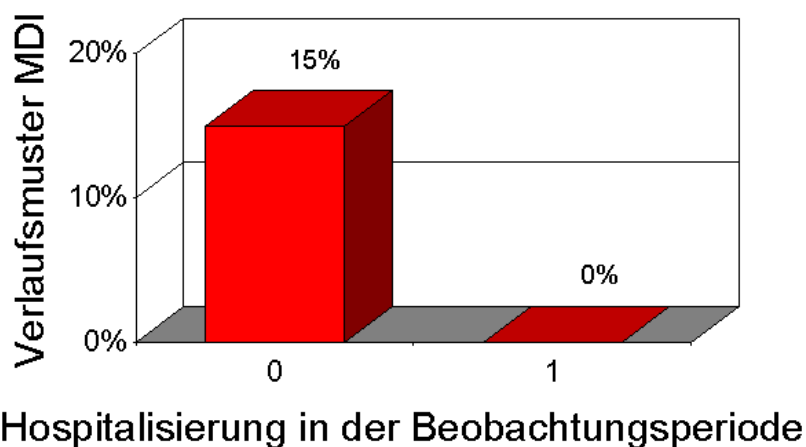


Abbildung 58. Anteil von Lithiumpatienten mit dem Verlaufsmuster MDI in den Gruppen ohne bzw. mit Hospitalisierung ( $n = 48$  vs.  $n = 27$ ).

### 6.3.2.5 Verlaufsmuster DMI

Unter Lithium erwies sich eine Episodenfolge DMI (Depression – Manie – freies Intervall) als tendenziell günstig für den Verlauf. Dieser Befund lag allerdings noch im Zufallsbereich ( $r = 0.19$ ,  $p = 0.10$ ,  $n = 75$ , vgl. Abb. 59). Unter Carbamazepin hingegen war der Zusammenhang zwischen dem Vorliegen einer Episodenfolge DMI und einem günstigen Verlauf statistisch signifikant ( $r = 0.30$ ,  $p = 0.02$ ,  $n = 56$ ). Der Unterschied zwischen den Punkt-biserialen Korrelationskoeffizienten unter Lithium und Carbamazepin war nicht signifikant.

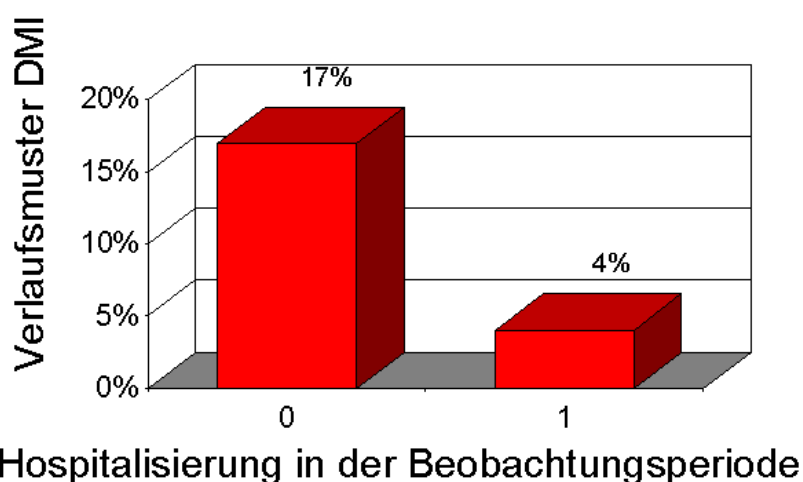


Abbildung 59. Anteil von Lithiumpatienten mit dem Verlaufsmuster DMI in den Gruppen ohne bzw. mit Hospitalisierung ( $n = 48$  vs.  $n = 27$ ).

### 6.3.2.6 Verlaufsmuster CC

Es zeigte sich kein klarer Zusammenhang zwischen einem continuous cycling (Verlaufsmuster CC) und dem Ansprechen auf eine Lithiumprophylaxe ( $r = -0.19$ ,  $p = 0.11$ ,  $n = 75$ , vgl. Abb. 60). Unter Carbamazepin fand sich ein tendenziell negativer Zusammenhang zwischen dem Verlaufsmuster CC und dem Verlauf unter einer Carbamazepinprophylaxe ( $r = -0.24$ ,  $p = 0.08$ ,  $n = 56$ ). Der Unterschied zwischen den Punkt-biserialen Korrelationskoeffizienten unter Lithium und Carbamazepin war nicht signifikant.

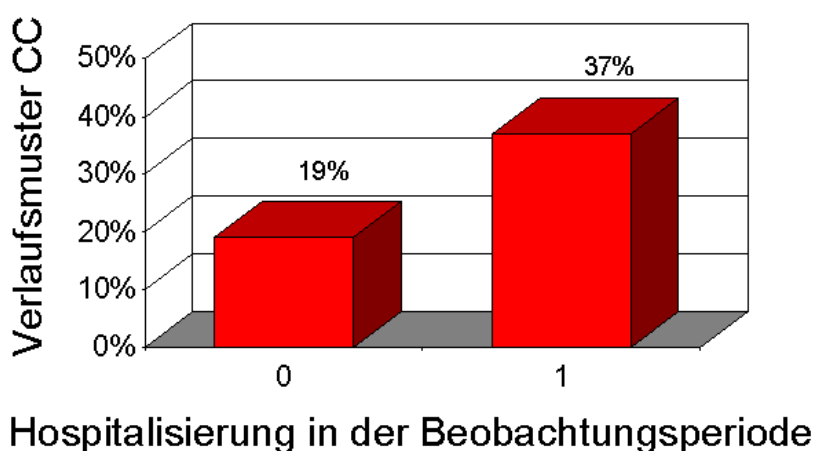


Abbildung 60. Anteil von Lithiumpatienten mit dem Verlaufsmuster CC in den Gruppen ohne bzw. mit Hospitalisierung ( $n = 48$  vs.  $n = 27$ ).

### 6.3.2.7 Komorbide Persönlichkeitsstörung

Das Vorliegen einer komorbiden Persönlichkeitsstörung erwies sich weder für den Verlauf unter Lithium noch für den Verlauf unter Carbamazepin als prädiktiv ( $r = 0.07$ ,  $p = 0.55$ ,  $n = 70$  unter Lithium (vgl. Abb. 61), bzw.  $r = -0.23$ ,  $p = 0.10$ ,  $n = 51$  unter Carbamazepin).

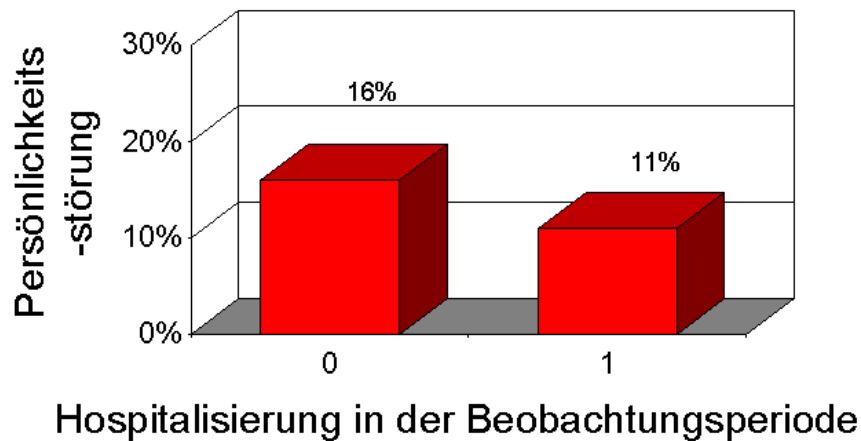


Abbildung 61. Anteil von Lithiumpatienten mit der Diagnose einer Persönlichkeitsstörung in den Gruppen ohne bzw. mit Hospitalisierung ( $n = 43$  vs.  $n = 27$ ).

### 6.3.2.8 Neurotizismus

Ein hoher Neurotizismus-Score zu Studienbeginn hatte keinen signifikanten Einfluss auf den Verlauf unter Lithium ( $r = -0.01$ ,  $p = 0.93$ ,  $n = 75$ , vgl. Abb. 62). Im Gegensatz hierzu war ein hoher Neurotizismus unter Carbamazepin mit einem ungünstigen Verlauf assoziiert ( $r = -0.36$ ,  $p = 0.01$ ,  $n = 55$ ). Der Unterschied der (Punkt-biserialen) Korrelationskoeffizienten zwischen Lithium und Carbamazepin war signifikant ( $Z = 2.05$ ,  $p = 0.04$ ).

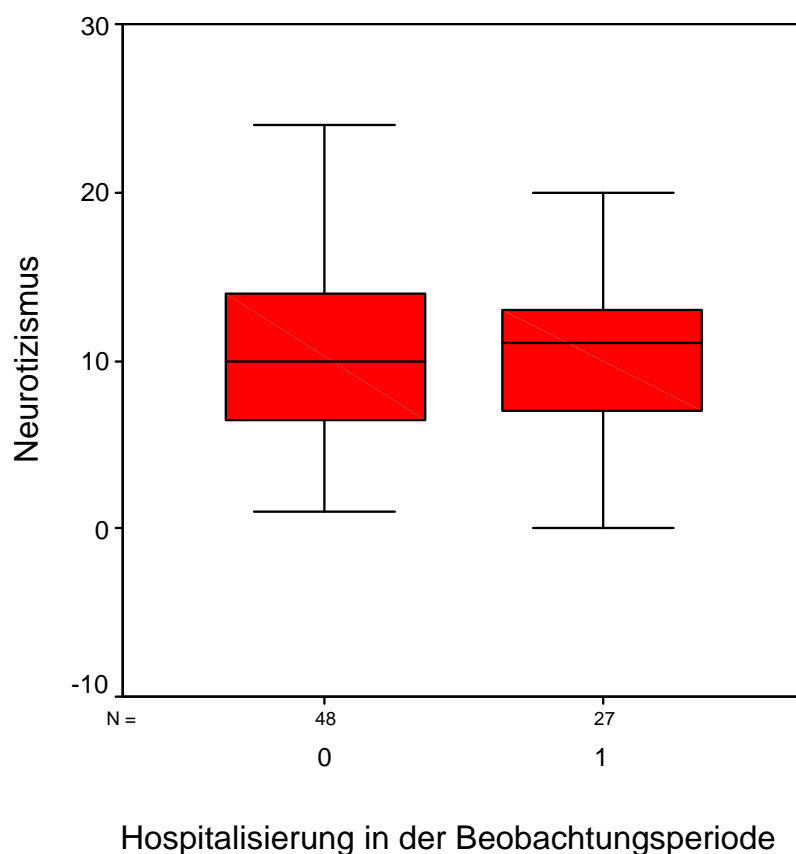


Abbildung 62. Neurotizismus-Score bei Lithiumpatienten ohne bzw. mit Hospitalisierung in der Beobachtungsperiode ( $n = 48$  vs.  $n = 27$ ).

### 6.3.2.9 Deprimiertheit oder Antriebsarmut

Unter Lithium erwies sich das Vorliegen eines der Kernsymptome einer depressiven Episode zu Studienbeginn (Deprimiertheit oder Antriebsarmut) als Risikofaktor für den weiteren Verlauf ( $r = -0.28$ ,  $p = 0.01$ ,  $n = 75$ , vgl. Abb. 63). Unter Carbamazepin fand sich tendenziell ein ähnlicher Effekt, jedoch lag die Korrelation von  $r = -0.25$  noch im Zufallsbereich ( $p = 0.06$ ,  $n = 56$ ). Der Unterschied der  $\phi$ -Koeffizienten zwischen Lithium und Carbamazepin war statistisch nicht signifikant ( $Z = 0.18$ ,  $p = 0.86$ ).

Zusammenfassend lässt sich sagen, daß Deprimiertheit oder Antriebsarmut zu Studienbeginn ein Negativprädiktor für den Verlauf unter Lithium ist, wobei aufgrund der Studiendaten unklar ist, ob dieser Prädiktor für Lithium spezifisch ist.

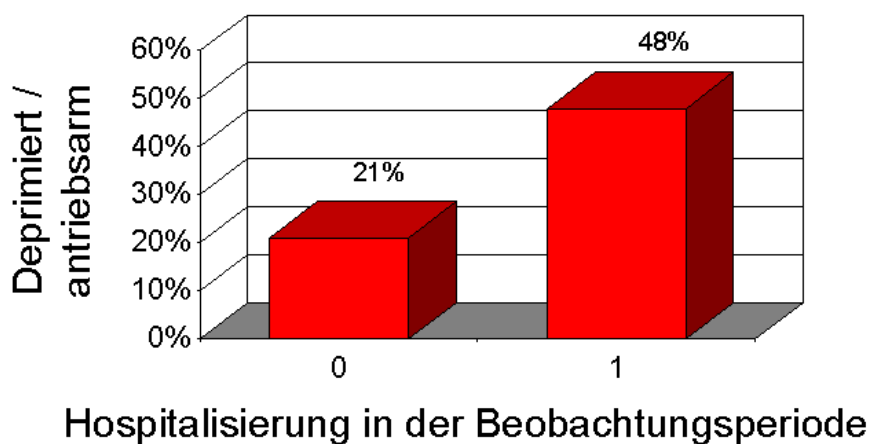


Abbildung 63. Anteil von Patienten mit depressiver Kernsymptomatik (Deprimiertheit oder Antriebsarmut) in den Gruppen ohne bzw. mit Hospitalisierung ( $n = 48$  vs.  $n = 27$ ).

### 6.3.2.10 Arbeitslosigkeit

Sowohl unter Lithium als auch unter Carbamazepin fand sich ein signifikanter und quantitativ gleich starker Zusammenhang zwischen Arbeitslosigkeit und Verlauf. In beiden Fällen war Arbeitslosigkeit ein Risikofaktor für den weiteren Verlauf ( $r = -0.28$ ,  $p = 0.01$ ,  $n = 75$  unter Lithium (vgl. Abb. 64), bzw.  $r = -0.28$ ,  $p = 0.03$ ,  $n = 56$  unter Carbamazepin).

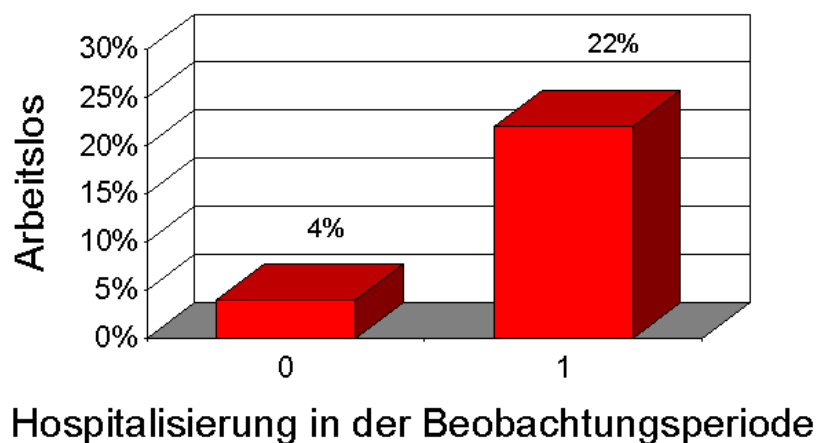


Abbildung 64. Anteil arbeitsloser Lithiumpatienten in den Gruppen ohne bzw. mit Hospitalisierung ( $n = 48$  vs.  $n = 27$ ).

### 6.3.2.11 Isoliertes Auftreten der Episoden

Das isolierte Auftreten der Episoden während der letzten Krankheitsphase hatte bei beiden Medikationen keinen signifikanten Einfluss auf den Verlauf ( $r = -0.09$ ,  $p = 0.44$ ,  $n = 75$  unter Lithium (vgl. Abb. 65), bzw.  $r = 0.14$ ,  $p = 0.29$ ,  $n = 56$  unter Carbamazepin).

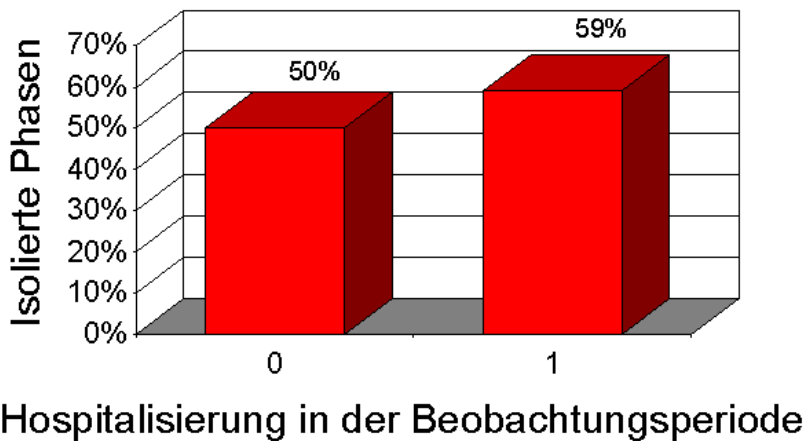


Abbildung 65. Anteil von Lithiumpatienten mit isolierten Auftreten der Phasen in den Gruppen ohne bzw. mit Hospitalisierung ( $n = 48$  vs.  $n = 27$ ).



### 6.3.2.12 Alter bei Ersterkrankung

Sowohl unter Lithium als auch unter Carbamazepin erwies sich ein hohes Ersterkrankungsalter als tendenziell protektiv gegen ein stationäres Rezidiv unter Lithium ( $r = 0.22$ ,  $p = 0.06$ ,  $n = 75$  unter Lithium (vgl. Abb. 66), bzw.  $r = 0.26$ ,  $p = 0.0502$ ,  $n = 56$  unter Carbamazepin). Dieses Ergebnis lässt sich statistisch allerdings weder für Lithium noch für Carbamazepin auf dem  $\alpha$ -Niveau von 5% absichern und ist daher nur mit Vorsicht zu interpretieren.

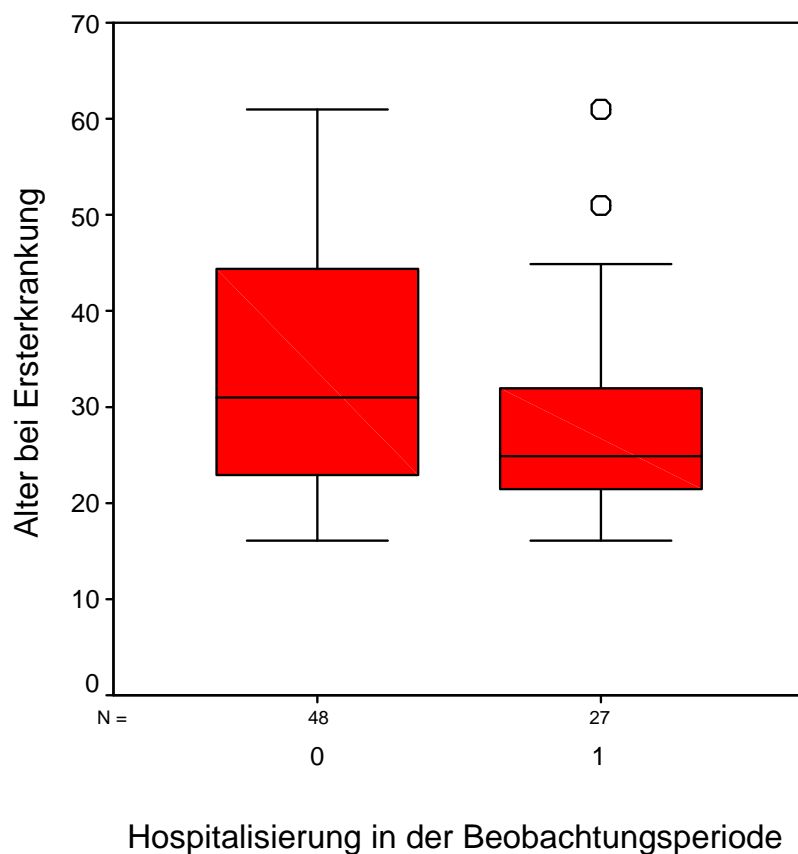


Abbildung 66. Ersterkrankungsalter bei Lithiumpatienten ohne bzw. mit Hospitalisierung in der Beobachtungsperiode ( $n = 48$  vs.  $n = 27$ ).

### 6.3.2.13 Phasenfrequenz

Eine hohe Frequenz affektiver Episoden (Anzahl der Episoden pro Zeiteinheit) war unter Lithium tendenziell ( $p = 0.065$ ) mit einem ungünstigen Verlauf assoziiert ( $r = -0.28$ ,  $n = 54$ ). Bei Beseitigung des einen Ausreißers in der Lithiumgruppe ohne Rehospitalisierung würde das Ergebnis signifikant werden ( $r = -0.28$ ,  $p = 0.013$ ,  $n = 53$ ). Die Beseitigung dieses Ausreißers entspräche zwar dem Schema zur Beseitigung möglicher Ausreißer. Aufgrund einer visuellen Kontrolle würde man den Ausreißer (die mit \* gekennzeichnete Beobachtung in Abb. 67) aber wohl nicht eliminieren, da er lediglich für die Gruppe von Patienten ohne Rezidiv untypisch ist, nicht aber für die Gesamtgruppe. Im Sinne einer konservativen Auswertungsstrategie wird daher in diesem speziellen Fall auf eine Ausreißerelimination verzichtet.

Unter Carbamazepin fand sich kein Zusammenhang zwischen Phasenfrequenz und Verlauf ( $r = -0.03$ ,  $p = 0.84$ ,  $n = 40$ ). Der Unterschied zwischen den Punkt-biserialen Korrelationskoeffizienten unter Lithium und Carbamazepin ist statistisch nicht signifikant ( $Z = 1.21$ ,  $p = 0.25$ ).

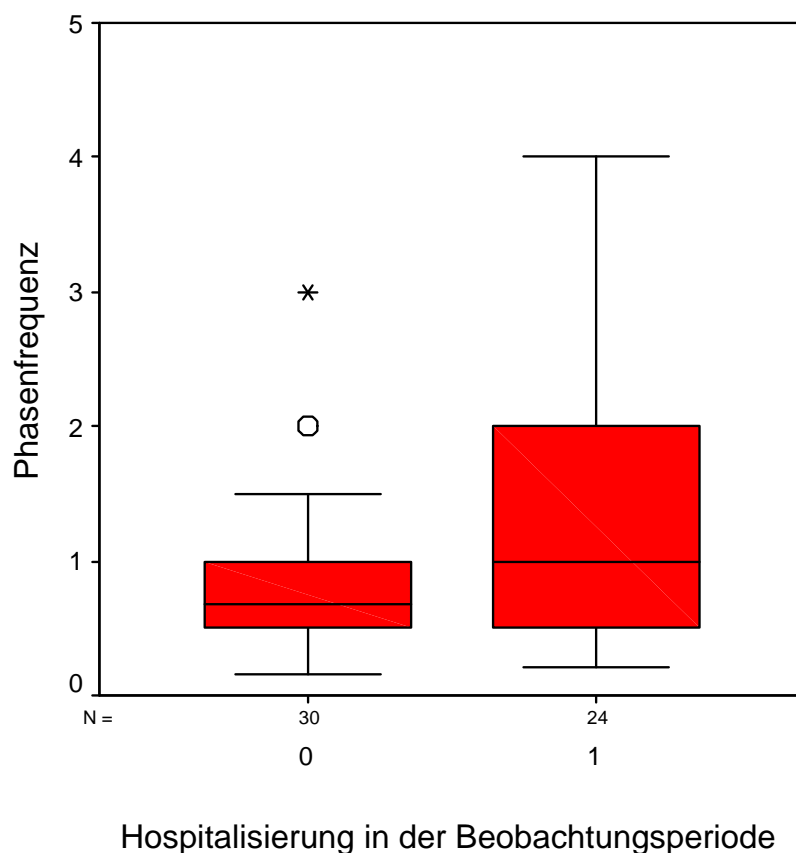


Abbildung 67. Phasenfrequenz bei Lithiumpatienten ohne bzw. mit Hospitalisierung in der Beobachtungsperiode ( $n = 30$  vs.  $n = 24$ ).

## 6.3.2.14 Stress

Der Zusammenhang zwischen Stress und Krankheitsverlauf war nur sehr schwer beurteilbar. Dies lag zum einen an der großen Zahl fehlender Werte und zum anderen an der durchaus problematischen Operationalisierung des Begriffs Stress als Summenscore des Bereichs M (Management) aus dem Social Interview Schedule (vgl. Kap. 5.2.12). Im Ergebnis zeigte sich auch hier kein statistisch signifikanter Zusammenhang zwischen Stress und Verlauf – weder für Lithium (vgl. Abb. 68), noch für Carbamazepin ( $r = 0.25$ ,  $p = 0.26$ ,  $n = 22$  unter Lithium bzw.  $r = 0.30$ ,  $p = 0.23$ ,  $n = 18$  unter Carbamazepin).

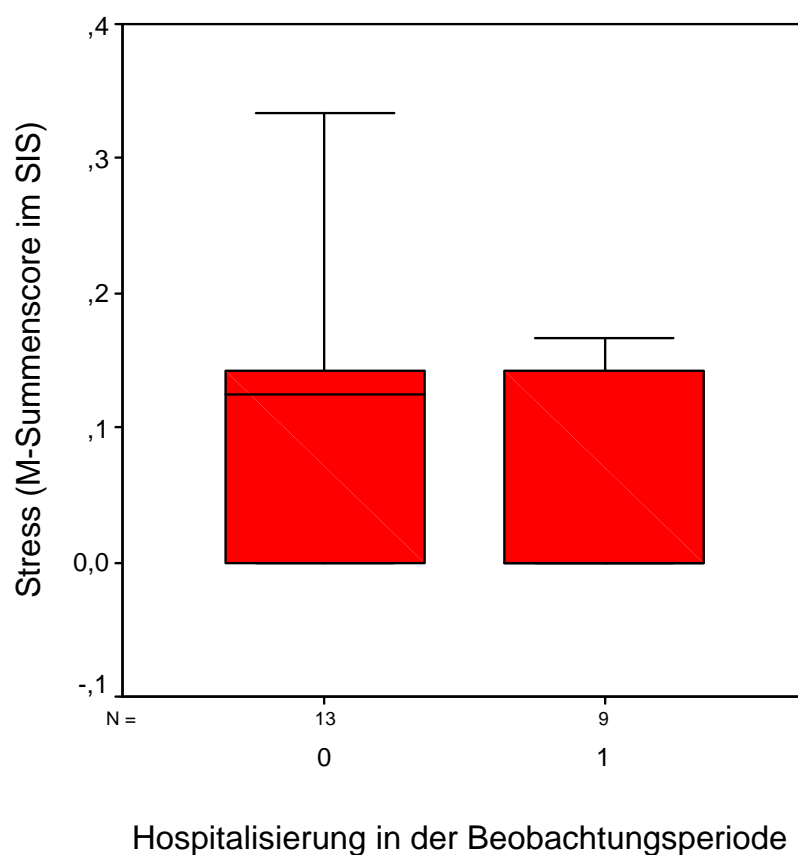


Abbildung 68. Ausmaß des Stresses bei Lithiumpatienten ohne bzw. mit Hospitalisierung in der Beobachtungsperiode ( $n = 13$  vs.  $n = 9$ ).

### 6.3.2.15 Expressed emotions

Bei nur 12 gültigen Beobachtungen unter Lithium und 9 gültigen Beobachtungen unter Carbamazepin waren keine aussagekräftigen Ergebnisse zur Responseprädiktion aufgrund von expressed emotions zu erwarten. Unter Lithium ergab sich eine Punkt-biseriale Korrelation von  $r = -0.03$  zwischen high expressed emotions und einem günstigen Verlauf ( $p = 0.93$ ,  $n = 12$ , vgl. Abb. 69). Für Carbamazepin lag diese Korrelation bei  $r = -0.32$  ( $p = 0.41$ ,  $n = 9$ ).

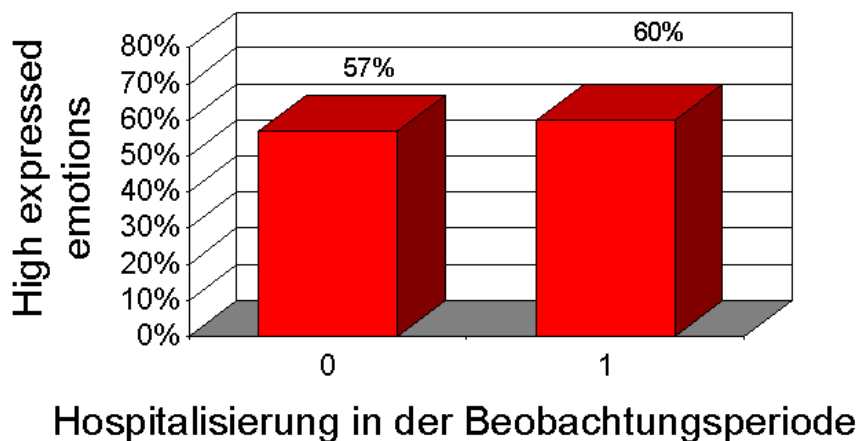


Abbildung 69. Anteil von Patienten mit high expressed emotions in den Gruppen ohne bzw. mit Hospitalisierung ( $n = 5$  vs.  $n = 7$ ).

## 6.3.2.16 Life events

Weder unter Lithium noch auch unter Carbamazepin fand sich ein signifikanter Zusammenhang zwischen life events und Verlauf ( $r = -0.02$ ,  $p = 0.88$ ,  $n = 59$  unter Lithium – (vgl. Abb. 70) und  $r = -0.02$ ,  $p = 0.90$ ,  $n = 46$  unter Carbamazepin). Dieses Ergebnis wurde auch nach Beseitigung der beiden Ausreißer in der Lithiumgruppe ohne Rehospitalisierung in vollem Umfang bestätigt ( $r = -0.02$ ,  $p = 0.88$ ,  $n = 57$ ).

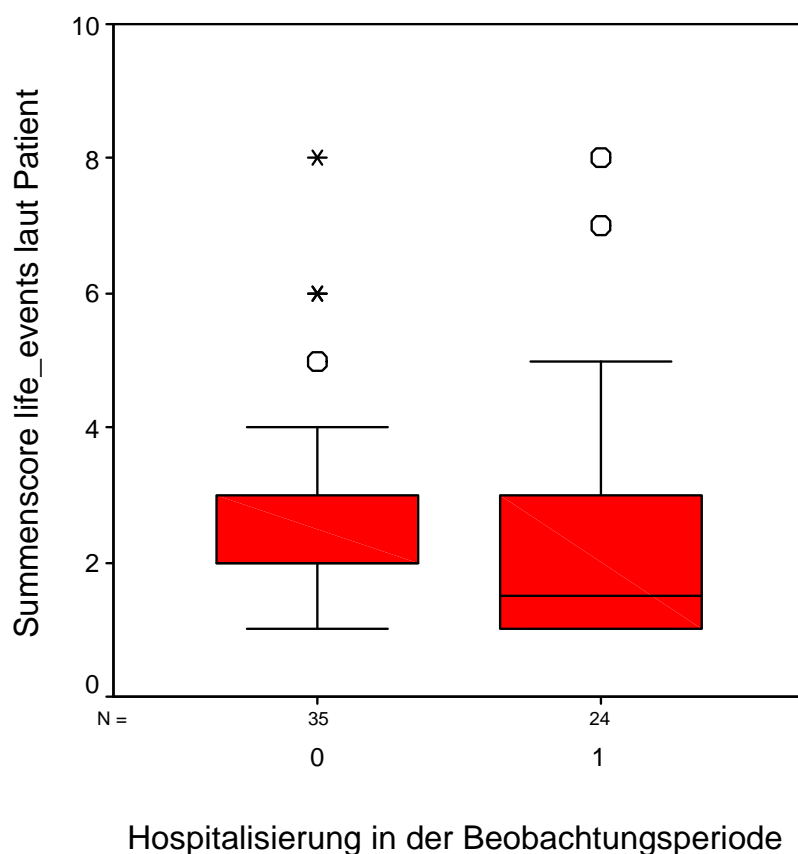


Abbildung 70. Life-events bei Lithiumpatienten ohne bzw. mit Hospitalisierung in der Beobachtungsperiode ( $n = 35$  vs.  $n = 24$ ).

### 6.3.2.17 Krankheitskonzepte

Der Summenscore der KK-Skala (Krankheitskonzept-Skala) erwies sich weder unter Lithium noch unter Carbamazepin als prädiktiv für den Verlauf. Die (Punkt-biserial) Korrelation zwischen KK-Summenscore und günstigem Verlauf war  $r = 0.13$  ( $p = 0.32$ ,  $n = 58$  unter Lithium, vgl. Abb. 71) und  $r = -0.02$ ,  $p = 0.88$ ,  $n = 50$  unter Carbamazepin).

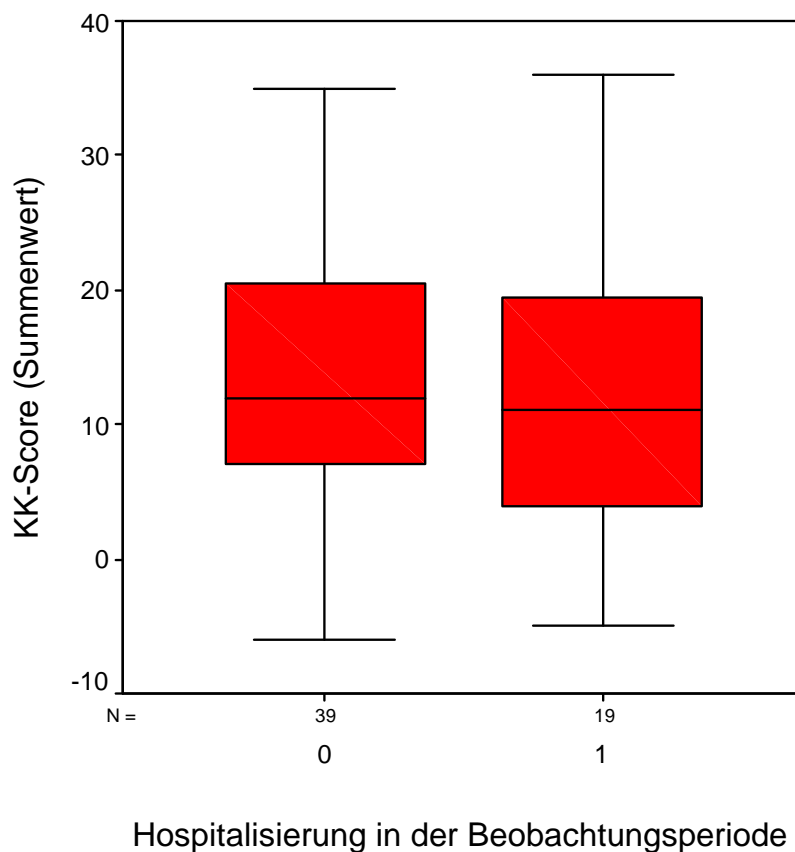
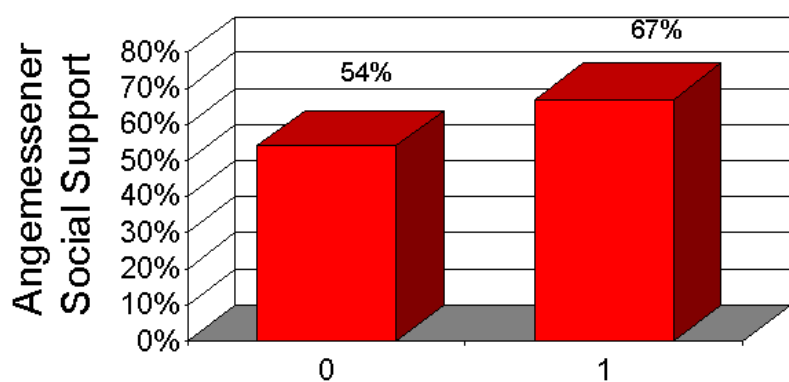


Abbildung 71. KK-Summenscore bei Lithiumpatienten ohne bzw. mit Hospitalisierung in der Beobachtungsperiode ( $n = 39$  vs.  $n = 19$ ).

### 6.3.2.18 Soziale Unterstützung

Unter keiner der beiden Studienmedikationen fand sich ein signifikanter Zusammenhang zwischen sozialer Unterstützung und Verlauf. Allerdings wurde die Variable soziale Unterstützung aus dem SIS nur bei einem Teil der Patienten erfasst, so daß die Aussagekraft dieser Auswertung eingeschränkt ist. Im Ergebnis ergab sich unter Lithium eine Punkt-biseriale Korrelation von  $r = -0.13$  zwischen sozialer Unterstützung und einem günstigen Verlauf ( $p = 0.57$ ,  $n = 22$ , vgl. Abb. 72). Für Carbamazepin lag diese Korrelation bei  $-0.18$ ,  $p = 0.49$ ,  $n = 18$ ).



Hospitalisierung in der Beobachtungsperiode

Abbildung 72. Anteil von Lithiumpatienten mit angemessener sozialer Unterstützung in den Gruppen ohne bzw. mit Hospitalisierung ( $n = 13$  vs.  $n = 8$ ).

### 6.3.2.19 Länge des ersten freien Intervalls

Die Länge freien Intervalls zwischen der ersten und zweiten Krankheitsepisode erwies sich unter keiner der beiden Studienmedikationen als prädiktiv ( $r = -0.06$ ,  $p = 0.67$ ,  $n = 51$  unter Lithium bzw.  $r = -0.02$ ,  $p = 0.89$ ,  $n = 45$  unter Carbamazepin). Unter Lithium war das Ergebnis numerisch stark von drei Ausreißern mit einem sehr langen freien Intervall zwischen der ersten und zweiten Krankheitsphase beeinflusst (vgl. Abb. 73). Allerdings erwies sich die Länge des ersten Intervalls auch nach Beseitigung (oder nach Stutzung) dieser drei Ausreißer als nicht prädiktiv ( $r = -0.22$ ,  $p = 0.14$ ,  $n = 48$  nach Ausreißerbereinigung, vgl. auch Abb. 74).

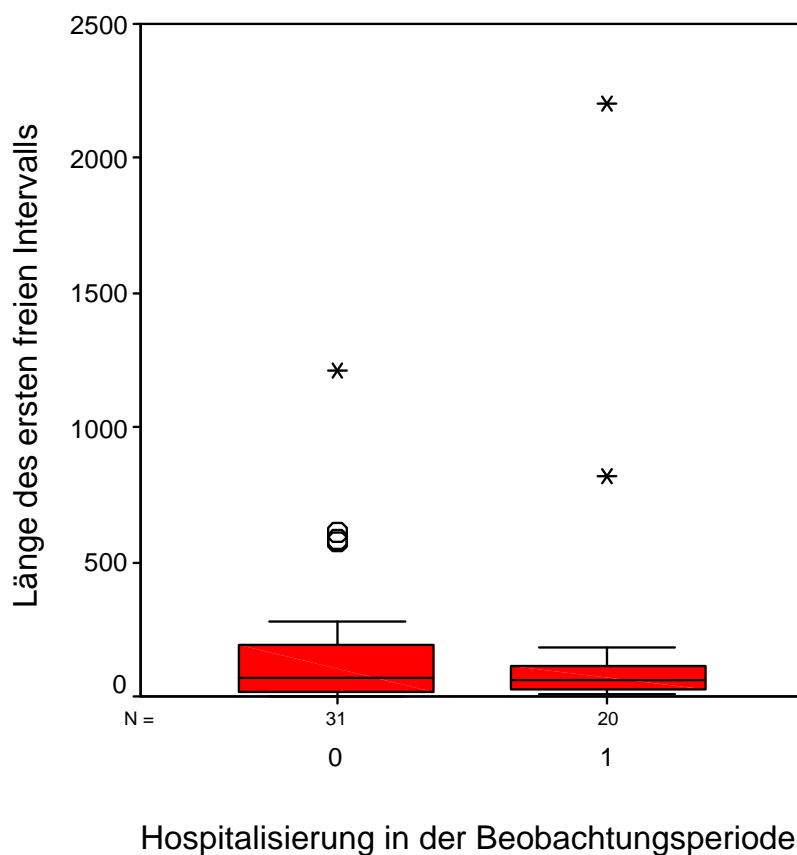


Abbildung 73. Länge des ersten freien Intervalls (in Tagen) bei Lithiumpatienten ohne bzw. mit Hospitalisierung in der Beobachtungsperiode ( $n = 31$  vs.  $n = 20$ ). Alle gültigen Beobachtungen.



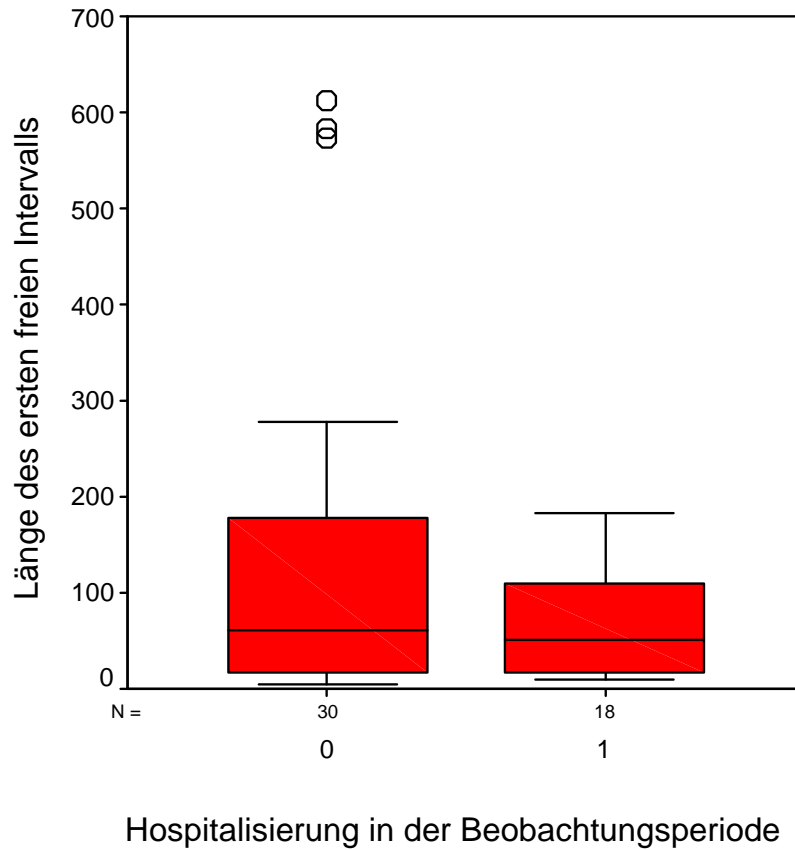


Abbildung 74. Länge des ersten freien Intervalls (in Tagen) bei Lithiumpatienten ohne bzw. mit Hospitalisierung in der Beobachtungsperiode (n = 30 vs. n = 18). Nach Entfernung der extremen Ausreißer.

### 6.3.2.20 Zusammenfassung der univariaten Analysen

Unter Lithium waren 3 von 19 Variablen signifikant mit dem Verlauf (günstig vs. ungünstig) korreliert: Das Verlaufsmuster MDI, das Vorhandensein depressiver Kernsymptome zu Studienbeginn (Antriebsarmut oder Depressivität) und Arbeitslosigkeit zu Studienbeginn. Die Episodenfolge MDI war ein Positivprädiktor, die beiden anderen Variablen waren Risikofaktoren. Alle Effektstärken  $|r|$  bewegten sich im Bereich von 0.24 bis 0.28. Somit stand keiner der aus der vorgeschalteten Metaanalyse extrahierten Prädiktoren mit dem Verlauf in einem starken Zusammenhang.

Bei der Zusammenfassung der Ergebnisse ist zu berücksichtigen, daß aufgrund der nicht durchgeführten  $\alpha$ -Fehler-Adjustierung bei den 19 durchgeführten Analysen im Durchschnitt mit einem falsch positiven Befund zu rechnen ist.

Ob diese Variablen generelle oder für Lithium spezifische Prädiktoren sind, lässt sich aufgrund der MAP-Studie nicht abschließend beantworten. Insgesamt sprechen die Befunde dafür, daß es sich bei depressiven Kernsymptomen und bei Arbeitslosigkeit eher um generelle Verlaufsprädiktoren handelt als um Lithium-spezifische Prädiktoren, da sich unter Carbamazepin für beide Prädiktorvariablen ähnlich große und dem Vorzeichen nach gleichsinnige Effekte fanden wie unter Lithium. Für das Verlaufsmuster MDI hingegen handelt es sich möglicherweise um einen für Lithium spezifischen Effekt.

Als Zwischenfazit bleibt festzuhalten, daß isolierte univariate Analysen keine befriedigende Prognose des Verlaufs unter Lithium erlauben. Die Ergebnisse bleiben insgesamt unsicher und die prognostische Kraft einzelner Variablen ist sehr begrenzt. Im folgenden wird versucht, die Prognose durch die simultane Betrachtung mehrerer Variablen entscheidend zu verbessern.

## 6.3.3 Multivariate Analysen

### 6.3.3.1 Logistische Regressionsmodelle

Im ersten Schritt werden verschiedene logistische Regressionsmodelle angewendet. Diese Analysen beziehen sich auf jene Patienten mit einem vollständig beurteilbaren Verlauf (ATP-Analysen, vgl. Kap. 6.2.1). Die drei verwendeten Modelle (vgl. Kap. 6.3.3.2-6.3.3.4) stellen verschiedene Ansätze zur Behandlung des Problems fehlender Werte in den unabhängigen Variablen (=Prädiktorvariablen) dar. Zunächst wurde versucht, die Lithiumresponse aufgrund des kompletten Satzes unabhängiger Variablen vorherzusagen (vgl. Kap. 6.3.3.2). Aufgrund massiver Schwierigkeiten mit fehlenden Prädiktorvariablen war es notwendig, die Prädiktorvariablen auf eine Teilmenge mit weitgehend vollständigen Variablen einzuschränken (vgl. Kap. 6.3.3.3) und/oder fehlende Werte zu ergänzen (vgl. Kap. 6.3.3.4).

In jedem Fall wurde die Wahrscheinlichkeit für einen günstigen Verlauf modelliert. Ein „günstiger Verlauf“ wurde operationalisiert als „2½ Jahre ohne stationäres Rezidiv“.

Bei insgesamt 11 der 86 Patienten (13%) war der Verlauf aufgrund eines vorzeitigen Studienabbruchs nicht vollständig dokumentiert und daher im Sinne einer ATP-Analyse nicht beurteilbar. Für anderen 75 von 86 Patienten (87%) unter Lithium konnte dieses outcome-Kriterium angewendet werden. Sie blieben entweder volle 2½ Jahre ohne Rezidiv (günstiger Verlauf), oder sie wurden unter Lithium als Studienmedikation rehospitalisiert (ungünstiger Verlauf).

#### 6.3.3.2 Logistische Regression mit vollem Variablensatz

Trotz der aufwendig und sorgfältig durchgeführten Datenerhebung in der MAP-Studie lag keine einzige Beobachtung mit gültigen Werten in allen 19 Prädiktorvariablen vor. Daher konnte keine der üblichen multivariaten Analysen (z.B. logistische Regression) mit allen 19 unabhängigen Variablen durchgeführt werden. Aufgrund des teilweise sehr hohen Anteils fehlender Werte wurde keine Datenimputation (Ersetzung fehlender Werte) durchgeführt. In einem nächsten Schritt wurden nur jene Variablen verwendet, für die mindestens 90% gültige Fälle vorlagen.

### 6.3.3.3 Logistische Regression mit weitgehend den vollständig vorhandenen Variablen

Dieses Kriterium von maximal 10% fehlender Werte war für 12 der 19 Prädiktorvariablen erfüllt:

- 1) Anzahl stationärer Episoden (n=74)
- 2) Krankheitsdauer (n=75)
- 3) Stimmungsinkongruente Wahnhalte in irgend einer Episode (n=75)
- 4) Verlaufsmuster MDI (n=75)
- 5) Verlaufsmuster DMI (n=75)
- 6) Verlaufsmuster CC (n=75)
- 7) Persönlichkeitsstörung (n=70)
- 8) Neurotizismus-Score (n=75)
- 9) Deprimiertheit oder Antriebsarmut (n=75)
- 10) Arbeitslosigkeit (n=75)
- 11) Isolierte Phasen (n=75)
- 12) Alter bei Ersterkrankung (n=75)

Bezüglich dieser 12 Variablen lag für n=70 Patienten unter Lithium ein vollständiger Variablensatz vor. Bei der Kollinearitätsanalyse dieser 12 Variablen ergab sich eine exakte lineare Abhängigkeit zwischen den Variablen „Verlaufsmuster MDI“, „Verlaufsmuster DMI“, „Verlaufsmuster CC“ und „Isolierte Phasen“. Nach der aufgrund dieser linearen Abhängigkeit notwendigen Entfernung einer dieser Variablen (standardmäßig wurde die letzte Variable entfernt) wurde eine erneute Kollinearitätsanalyse für die verbleibenden 11 Variablen durchgeführt. Hierbei ergab sich bei diesen 70 Beobachtungen eine Konditionzahl von 13.8 und Toleranzwerte für die einzelnen Variablen von 0.608 bis 0.835. Daher kann die Matrix der unabhängigen Variablen als gut konditioniert betrachtet werden (vgl. Belsley et al. 1980), es liegt also nach Beseitigung der exakten Linearität keine ausgeprägte Multikollinearitätsproblematik mehr vor. Der Hosmer-Lemeshow Goodness-of-Fit Test zeigte keine signifikante Abweichung der empirischen Daten vom Modell der logistischen Regression an ( $\chi^2 = 9.81$ ,  $p = 0.28$ ).

Bei der Durchführung der logistischen Regression ergibt sich ein statistisch signifikantes Gesamtmodell zur Erklärung der Wahrscheinlichkeit für einen günstigen Verlauf ( $\chi^2 = 33.6$ ,  $p = 0.001$ ). Keine der einzelnen Prädiktorvariablen des Modells war signifikant (p-Werte von 0.06 bis 1.00, vgl. Tab. 10). Insgesamt stehen die Ergebnisse weitgehend in Einklang mit den Ergebnissen aus den univariaten Analysen – für die beiden Variablen, die sich in den univariaten Analysen als Risikofaktoren erhärtet hatten, Depressivität/Antriebsarmut (DEPR\_AA) und Arbeitslosigkeit (ARB\_LOS) zeigte sich in der logistischen Regression eine gleichsinnige, allerdings statistisch nicht signifikante Tendenz ( $p = 0.11$  bzw.  $0.17$ , vgl. Tab.

10). Für die im Rahmen der univariaten Analysen dritte signifikant mit dem Verlauf korrelierte Variable, das Verlaufsmuster MDI, ergab sich allerdings im Rahmen der logistischen Regression aufgrund des großen Standardfehlers keine klare Tendenz (vgl. Tab. 10).

Tabelle 10: Modellkoeffizienten bei der Prognose der Lithiumresponse (logistische Regression).

Variable	Regressions- koeffizient B	Standard- fehler	Wald	df	p-Wert	Exp(B)
STATEPI	-0.40	0.21	3.66	1	0.06	0.67
KRANKDAU	0.07	0.05	2.63	1	0.11	1.08
INK_WAHN	-0.27	0.78	0.12	1	0.73	0.77
MDI	1.24	1.54	0.64	1	0.42	3.44
DMI	21.93	12764	0.00	1	1.00	3.3E+09
CC	1.25	1.24	1.01	1	0.31	3.50
PERS_ST	0.02	0.07	0.07	1	0.79	1.02
NEUROT_0	-1.48	0.80	3.41	1	0.07	0.23
DEPR_AA	-2.30	1.42	2.62	1	0.11	0.10
ARB_LOS	0.000	0.00	1.86	1	0.17	1.00
ALT1ERKR	0.20	0.77	0.07	1	0.80	1.22
Konstante	0.66	0.37	3.22	1	0.07	1.07

Insgesamt wurden im Rahmen dieses Modells 77.1% der Beobachtungen korrekt vorhergesagt (vgl. Tab. 11). Hierbei wurde ein Trennwert von 0.5 zugrunde gelegt, d.h. eine Beobachtung wird genau dann ein günstiger Verlauf prognostiziert, falls die aufgrund des Modells geschätzte Wahrscheinlichkeit für einen günstigen Verlauf  $\geq 50\%$  ist. Zu beachten ist allerdings, daß bei der Übertragung auf eine neue Stichprobe eine erheblich niedrigere Trefferquote zu erwarten ist – allein schon, weil die Daten, an der die Prädiktionsfunktion berechnet wurde gleichzeitig zur Bestimmung der Trefferquote verwendet wurde (fehlendes Re-sampling, vgl. hierzu Kap. 6.3.3.4).

Tabelle 11: Klassifizierungstabelle für die Prognose auf Grundlage einer logistischen Regression.

	Beobachtet		% korrekter Vorhersagen	
	+ (günstiger Verlauf)	- (ungünstiger Verlauf)		
Vorhergesagt <sup>1</sup>	+ (günstiger Verlauf)	19	8	70%
	- (ungünstiger Verlauf)	8	35	81%
Gesamt				77% (54/70)

<sup>1</sup>Bei einem Trennwert von 0.5

#### 6.3.3.4 Logistische Regression nach Datenimputation

Nach Imputation der fehlenden Werte (Schätzung der fehlenden Werte durch den Spaltenmittelwert) des in Kap. 6.3.3.3 beschriebenen Modells ergeben sich weitgehend unveränderte Ergebnisse. Auch nach der Datenimputation war die Prognose der Lithiumresponse im Gesamtmodell signifikant ( $\chi^2 = 31.7$ ,  $p = 0.001$ ), und ähnlich wie im Modell ohne Datenimputation wurden 76% (57/75) der Beobachtungen korrekt vorhergesagt. In diesem Modell waren zwei Prädiktoren signifikant – die Anzahl stationärer Episoden in der Vorgeschichte und depressive Kernsymptomatik (Depressivität/Antriebsarmut).

#### 6.3.3.5 Weitere multivariate Analysen, Jackknife-Methode

Um die Ergebnisse der logistischen Regressionsanalysen zu erhärten bzw. zu relativieren wurden eine Diskriminanzanalyse und eine Cox-Regression mit demselben Datensatz wie in Kap. 6.3.3.4. (bei dem – wie beschrieben – einige wenige fehlende Werte ergänzt wurden) gerechnet. Die Ergebnisse aus den logistischen Regressionsanalysen wurden hierdurch voll bestätigt.

Bei der Cox-Regression ergab sich ein p-Wert von  $p = 0.001$  ( $\chi^2 = 33.0$ ) für das Gesamtmodell. Zudem waren wieder die beiden signifikanten Einflussfaktoren aus der logistischen Regression nach Datenimputation signifikant (Anzahl stationärer Episoden und depressive Kernsymptomatik). Als weiterer Prädiktor kam im Rahmen der Cox-Regression die bisherige Krankheitsdauer hinzu.

Bei der Diskriminanzanalyse lag der p-Wert für die Prädiktionskraft des Gesamtmodells bei  $p = 0.005$  (Wilks  $\Lambda = 0.670$ ). Auch die Diskriminanzanalyse ergab eine vergleichbare Quote korrekter Zuordnungen (74.7%). Zu beachten ist jedoch, daß diese Quote auf 62.7% sinkt, falls die Diskriminanzanalyse mit einer Jackknife-Methode durchgeführt wird. Hierbei wird jeweils eine Beobachtung ausgeschlossen, dann wird die Diskriminanzfunktion aufgrund der übrigen Variablen bestimmt und am Ende die jeweils ausgeschlossene Beobachtung aus dem ermittelten Modell "vorhergesagt". Nur durch diese Trennung von Eichstichprobe und den zugeordneten Beobachtungen erhält man eine realistische Schätzung für den Anteil korrekt prognostizierter Verläufe in einer neuen, gleichartigen Stichprobe (Lachenbruch & Mickey 1968).

#### 6.3.3.6 Multivariate Auswertungen für die Carbamazepingruppe

Bei multivariater Auswertung der Daten für die Carbamazepingruppe fanden sich ähnliche Ergebnisse wie für die Lithiumgruppe. So lag unter Carbamazepin die Quote korrekter Zuordnungen aufgrund einer logistischen Regression nach Datenimputation bei 75%. Das (sich auf den Vektor der Regressionskoeffizienten beziehende) Gesamtmodell war zwar hochsignifikant, allerdings konnte für keine der Einzelvariablen ein signifikanter Beitrag zum Gesamtmodell ermittelt werden. Aufgrund einer Diskriminanzanalyse mit demselben Datensatz wurden 73% der Beobachtungen korrekt prognostiziert. Allerdings ist auch in diesem Fall davon auszugehen, daß diese hohe Trefferquote zu einem großen Teil mit der fehlenden Trennung von Eichstichprobe und den später zugeordneten Beobachtungen zusammenhängt. Analog zur Lithiumgruppe bestätigte sich diese Vermutung nachdem diese Trennung mittels Jackknife-Methode durchgeführt wurde – wieder sank die Trefferquote auf 63% ab. Insgesamt spricht der Vergleich der multivariaten Ergebnisse für Lithium und Carbamazepin nicht für spezifische Effekte in der Lithiumgruppe.

#### 6.3.3.7 Zwischenfazit zu den multivariaten Analysen

Die Durchführung multivariater Analysen war mit den erwarteten, vor allem durch das Auftreten fehlender Werte bedingten, technischen Schwierigkeiten verbunden. Nachdem eine für die vorliegende Datenstruktur adäquate Lösung dieser Schwierigkeiten verwendet wurde,

zeigte sich eine zunächst befriedigende Prognosegüte der aufgestellten Modelle. Eine nähere Analyse dieser Modelle zeigte jedoch, daß die Prognosegüte in der praktischen Anwendung bei einem unabhängigen Datensatz weit niedriger anzusetzen ist.

Zusammenfassend bleibt festzuhalten, daß die Verwendung multivariater Modelle bei dem vorliegenden Datensatz nicht den entscheidenden Durchbruch gebracht hat. Die Prognose war zwar numerisch etwas besser als aufgrund einzelner vielversprechenden Variablen wie z.B. der Anzahl bisheriger Episoden oder Phasenfrequenz. Insgesamt erscheinen multivariate Modelle jedoch für den Praxiseinsatz nur sehr bedingt geeignet – zum einen aufgrund der praktischen Schwierigkeiten, zum anderen ist aufgrund der Ergebnisse der MAP-Studie von einer nur mäßigen Prognosegüte auszugehen.

### **6.3.4 Prädiktion des Verlaufs aufgrund einer Lithium-Response-Skala**

#### 6.3.4.1 Definition der Lithium-Response-Skala (LRS)

Die Lithium-Response-Skala (LRS) wurde auf Grundlage jener klinischen und psychosozialen Variablen definiert, die sich in der Serie von Metaanalysen (vgl. Kap. 4.3 - 4.5) als Prädiktoren der prophylaktischen Wirksamkeit von Lithium erhärtet hatten. Zur Definition der Lithium-Response-Skala wurde jede dieser Variablen (sofern sie nicht schon dichotom vorlag) dichotomisiert. Für jede Variable mit einer prognostisch günstigen Ausprägung (z.B. hohes Ersterkrankungsalter) wird ein Punkt zum LRS-Score addiert, für jede prognostisch ungünstige Merkmalsausprägung (z.B. frühes Ersterkrankungsalter) wird ein Punkt abgezogen. Variablen mit fehlenden Werten gehen nicht in den LRS-Score ein.

Die Definition ist so gewählt, daß der LRS-Score stets wie folgt interpretiert werden kann: Je höher der LRS-Summenscore, um so besser ist die Prognose. Der Maximalwert für den LRS-Score ist 21, der Minimalwert ist -21. LRS-Summenscores  $> 0$  bedeuten eine günstige Prognose, LRS-Summenscores  $< 0$  bedeuten eine ungünstige Prognose (vgl. Abb. 74). Jeder Punkt bedeutet entweder das Vorhandensein eines protektiven Faktors oder das definitive Fehlen eines Risikofaktors. Jeder Punktabzug bedeutet das Fehlen eines protektiven Faktors oder das Vorhandensein eines Risikofaktors. Der LRS-Score kann auch bei unvollständiger Information (bei fehlenden Werten) bestimmt und interpretiert werden und kommt damit den Erfordernissen der klinische Praxis sehr entgegen.



$$\left. \begin{array}{l} LRS - Score < 0 \\ LRS - Score = 0 \\ LRS - Score > 0 \end{array} \right\} \longrightarrow \left\{ \begin{array}{l} \text{ungünstige Prognose} \\ \text{neutral} \\ \text{günstige Prognose} \end{array} \right.$$

Abbildung 74: Interpretation von LRS-Scores.

Die genaue Definition des LRS-Score lautet:

$$LRS - Score = \sum_{\substack{\text{beobachte} \\ \text{Variablen} \\ X_1 \dots X_k}} I_{(X_i \text{ ist prognostisch günstig})} - I_{(X_i \text{ ist prognostisch ungünstig})}$$

mit  $I_{(\dots)}$  = Indikatorfunktion und den in Tab. 12 definierten Variablen.

Tabelle 12: Prädiktoren der Lithium-Response-Skala.

Variable		prognostisch günstig	prognostisch ungünstig
X <sub>1</sub>	Anzahl stationärer Episoden in der Vorgeschichte	< 2	> 2
X <sub>2</sub>	Krankheitsdauer	< 6 Jahre	> 6 Jahre
X <sub>3</sub>	Jemals stimmungsinkongruente Wahnhinhalte während einer affektiven Episode?	nein	ja
X <sub>4</sub>	Verlaufsmuster MDI	ja	nein
X <sub>5</sub>	Verlaufsmuster DMI	nein	ja
X <sub>6</sub>	Verlaufsmuster CC (Continuous Cycling)	nein	ja
X <sub>7</sub>	Vorliegen einer Persönlichkeitsstörung?	nein	ja
X <sub>8</sub>	Hoher Neurotizismus	nein (Summenscore Neurotizismus des MPT < 10)	ja (Summenscore Neurotizismus des MPT > 10)

X <sub>9</sub>	deprimierte Stimmung oder Verlangsamung	nein (item 63 und item 80 des AMDP-100 sind 0)	ja (item 63 und/oder item 80 des AMDP-100 sind größer 0)
X <sub>10</sub>	Arbeitslosigkeit?	nein	ja
X <sub>11</sub>	Isoliertes Auftreten der Episoden	ja	nein
X <sub>12</sub>	Alter bei Ersterkrankung	> 28	< 28
X <sub>13</sub>	Phasenfrequenz	< 0.75 pro Jahr	> 0.75 pro Jahr
X <sub>14</sub>	Vorliegen von Stress	nein (M-Summenwert des SIS < 0.076)	ja (M-Summenwert des SIS > 0.076)
X <sub>15</sub>	High expressed emotions	nein	ja
X <sub>16</sub>	Life events	nein	ja
X <sub>17</sub>	Erwartete Compliance (günstige Krankheitskonzepte)	ja (KK-Summenscore > 10)	nein (KK-Summenscore < 10)
X <sub>18</sub>	Soziale Unterstützung	gut (mindestens 10 Kontakte zu nahen Verwandten in den letzten 3 Monaten laut SIS)	gering (weniger als 10 Kontakte zu nahen Verwandten in den letzten 3 Monaten laut SIS)
X <sub>19</sub>	Zeit zwischen der ersten und zweiten affektiven Phase	< 15 Monate	> 15 Monate
X <sub>20</sub>	Soziale Schicht	hoch	niedrig
X <sub>21</sub>	Dominante Charakterzüge	ja	nein

Von den 21 in der vorgehenden Metaanalyse als Prädiktoren identifizierten Variablen wurden 19 Variablen im Rahmen der MAP-Studie erfasst (X<sub>1</sub>–X<sub>19</sub>). Die Dichotomisierung ordinaler und metrischer Daten (z.B. von X<sub>1</sub>: Anzahl bisheriger Hospitalisierungen) wurde am Median der vorliegenden Stichprobe vorgenommen (Mediansplit), weil in den Primärstudien der Metaanalyse in der Regel keine einheitlichen Messinstrumente verwendet wurden, und daher keine klare Bestimmung des Trennwertes aus den Primärstudien möglich war. Diese

(gerundeten) Trennwerte sind zusammen mit den prognostisch günstigen vs. ungünstigen Merkmalsausprägungen der dichotomen Variablen in obenstehender Tabelle aufgelistet (Tab. 12).

#### 6.3.4.2 Ergebnisse bei Verwendung der Lithium-Response-Skala (LRS)

Von den 75 Patienten mit vollständig beurteilbarem Verlauf unter Lithium ergab sich aufgrund der LRS für 57 Patienten eine positive Prognose, für 10 Patienten eine negative Prognose und für 8 Patienten ein LRS-Summenscore von 0 (d.h. es konnte keine Prognose gestellt werden). Von den 10 Patienten mit guter Prognose (LRS > 0) hatten 7 tatsächlich einen günstigen Verlauf (70%). 43 der 57 Patienten mit negativer Prognose (LRS < 0) hatten tatsächlich einen ungünstigen Verlauf (75%). Damit ergibt sich eine globale Trefferquote von 75% (50/67) und ein Phi-Koeffizienten von  $\phi = 0.35$  ( $p = 0.008$ , vgl. Tab. 13). Für die (Punkt-biserialen) Korrelation zwischen dem LRS-Wert und dem Verlauf ergibt sich ein Wert von  $r = 0.37$  ( $p = 0.001$ ).

Tabelle 13: Klassifizierungstabelle auf Grundlage der Lithium-Response-Skala (LRS).

	Beobachtet		% korrekter Vorhersagen	
	+	-		
	(günstiger Verlauf) <sup>2</sup>	(ungünstiger Verlauf)		
Vorher- gesagt <sup>1</sup>	+	7	3	70%
	(günstiger Verlauf)			
	-	14	43	75%
	(ungünstiger Verlauf)			
Gesamt				75% (50/67)

<sup>1</sup>Bei einem Trennwert von 0

<sup>2</sup>Da in Tabelle 11 und Tabelle 13 leicht unterschiedliche Subgruppen der insgesamt 86 Lithiumpatienten eingegangen sind, ist die Anzahl von Patienten mit beobachtetem günstigen oder ungünstigen Verlauf nicht direkt vergleichbar.

Bei der Interpretation dieser Ergebnisse ist zu beachten, daß im Unterschied zu den vorgehenden multivariaten Analysen für jede Prädiktorvariable Betrag und Vorzeichen a priori aufgrund der Ergebnisse der Metaanalyse feststehen. Dadurch entfällt der Bias, der bei multivariaten Standardverfahren in der Regel zu einer erheblichen Überschätzung der Prognosegüte führt. Daher geben die Ergebnisse dieser Analyse ein realistisches Bild der Prognosegüte der LRS in der klinischen Anwendung.

#### 6.3.4.3 Ergebnisse für Carbamazepingruppe

Um zu testen, ob die signifikante Vorhersage der Response aufgrund der Lithium-Response-Skala für Lithium spezifisch ist, oder ob es sich bei den Prädiktoren eher um allgemeine Verlaufsprädiktoren handelt, wurden diese Analysen für die Carbamazepingruppe wiederholt. Im Ergebnis zeigte sich eine deutlich niedrigere, und statistisch nicht mehr signifikante Trefferquote von 63% ( $p = 0.20$ ). Dies ist ein Hinweis darauf, daß die signifikante Responseprädiktion zumindest zum Teil für Lithium spezifisch ist. Eine formelle Testung der Unterschiedlichkeit der Trefferquoten von 75% (unter Lithium) vs. 63% (unter Carbamazepin) ergab allerdings keine inferenzstatistische Absicherung dieses Befundes ( $p = 0.23$ ).

### 6.3.5 Zusammenfassung der empirischen Ergebnisse für Lithium

Die Ergebnisse der MAP-Studie zur Prädiktion der Lithiumresponse auf Grundlage der vorgeschalteten Metaanalysen lassen sich folgendermaßen zusammenfassen:

1) Bei isolierten univariaten Analysen einzelner Variablen waren drei Prädiktoren signifikant mit dem Verlauf unter Lithium korreliert:

- das Verlaufsmuster MDI
- das Vorhandensein depressiver Kernsymptome (Antriebsarmut oder Depressivität) zu Studienbeginn
- und Arbeitslosigkeit zu Studienbeginn

Die prognostische Kraft dieser isoliert betrachteten Variablen war allerdings begrenzt und insgesamt ließ sich die Lithium-Response aufgrund univariater Analysen nur sehr eingeschränkt vorhersagen ( $|r| = 0.24$  bis  $0.28$ ,  $p = 0.01$  bis  $0.04$ ).

2) Die verschiedenen multivariaten Verfahren erbrachten sehr konsistente Resultate. Im Ergebnis spielte es kaum eine Rolle, ob parametrische ATP-Analysen (logistische Regression, parametrische Diskriminanzanalyse) oder eine nonparametrische ITT-Analyse (Cox-

Regression) durchgeführt wurde. Stets war das Gesamtmodell signifikant und auch auf Ebene der einzelnen Modellvariablen zeigten sich dem Vorzeichen und relativen Gewicht der Variablen sehr ähnliche Ergebnisse. Auch der Anteil korrekt vorhergesagter Beobachtungen lag stets im Bereich von 75%-77% - unabhängig davon, ob eine Diskriminanzanalyse oder eine logistische Regression durchgeführt wurde, und ob fehlende Werte ersetzt wurden oder nicht. Hierbei ist allerdings zu beachten, daß diese hohe Trefferquote keine unverzerrte Schätzung darstellt, daß also in der Praxis mit einer wesentlich niedrigeren Trefferquote zu rechnen ist (vgl. Kap. 6.3.3.5). Weitere Probleme in der klinischen Anwendung multivariater Methoden sind die komplizierte (für die Anwendung an einen Rechner gebundene) Prädiktionsfunktion und vor allem das Problem der fehlenden Werte. Aufgrund des zuletzt genannten Problems könnte die Response nur für jene Patienten vorhergesagt werden, für die Merkmalsausprägungen aller Prädiktorvariablen bekannt sind oder geschätzt wurden. Dies stellt eine erhebliche Einschränkung für die praktische Anwendung dar.

3) Die Prädiktion des Verlaufs aufgrund einer Lithium-Response-Skala (additiv univariater Ansatz) erbrachte im Vergleich zu der isolierten Betrachtung einzelner Variablen eine erhebliche Verbesserung der Prognosegüte. Die Trefferquote lag bei 75% ( $r = 0.37$ ,  $p = 0.001$ ). Beim Vergleich mit den Ergebnissen für die multivariaten Verfahren ist zu beachten, daß die Prognose aufgrund der Ergebnisse der Metaanalyse gestellt und anhand eines neuen, unabhängigen, Datensatzes überprüft wurde – somit stellt die Trefferquote eine realistische Schätzung für die Prognosegüte im klinischen Einsatz dar.

Von großer Bedeutung für die Verwendung der Lithium-Response-Skala in der klinischen Praxis ist der unkomplizierte und flexible Umgang mit fehlenden Werten. Die Lithium-Response-Skala wurde so konstruiert, daß sich auch bei fehlenden Werten stets ein leicht zu interpretierender Wert ergibt, der für die Prognose des individuellen Verlaufs verwendet werden kann.

Zusammenfassend lässt sich festhalten, daß sich die Verwendung der Lithium-Response-Skala (additiv univariater Ansatz) am besten bewährt hat. Bei diesem Ansatz werden die verstreuten univariaten Einzelinformationen, die für sich genommen eine nur geringe prädiktive Kraft haben, gebündelt und so eine Verbesserung der Prognosegüte erreicht. Die Einbeziehung der Variableninteraktionen in die Modellbildung (multivariate Modelle) führte zu keiner weiteren Verbesserung der Trefferquote, sondern ist im Gegenteil mit erheblichen Schwierigkeiten bei der praktischen Anwendung verbunden. Auch der praktische Aspekt spricht eindeutig für die Lithium-Response-Skala – insbesondere aufgrund ihres einfachen und flexiblen Umgang mit fehlenden Werten.

## 7 Diskussion

Der erste Hauptteil der vorliegenden Arbeit umfasst eine quantitative Literaturübersicht zur Prädiktion der prophylaktischen Wirkung von Lithium aufgrund klinischer, demographischer und psychosozialer Variablen. Einerseits bereitet dieser Teil den zweiten Hauptteil (die Entwicklung und Evaluation eines Prognoseinstruments) vor. Andererseits erfüllt die Literaturübersicht auch einen „Zweck in sich“, nämlich die Vielzahl der verstreuten und teilweise vergessenen Einzelbefunde der Primärliteratur zusammenzustellen und systematisch und übersichtlich aufzubereiten. Die Darstellung des derzeitigen Wissensstandes zur Prädiktion der Lithiumresponse erfolgte im Rahmen einer Serie von Metaanalysen. Hierdurch werden die bisher stets informell durchgeführten Übersichtsarbeiten zur Responseprädiktion (insbesondere Carroll 1979, Grof 1983a,b,c, Abou-Saleh & Coppin 1986, Brochier & Gay 1991, Preisig & Angst 1991, Grof et al. 1993, Calabrese et al. 1996, Walden et al. 1999) durch einen systematischen Ansatz ergänzt. Ein Manko informeller Übersichtsarbeiten ist, daß ein größerer Teil der Primärstudien in der Regel keine Berücksichtigung findet. So wurde ein bedeutender Teil der Originalarbeiten offenbar nicht oder nur teilweise unter dem Aspekt der Responseprädiktion ausgewertet (z.B. die Arbeit von Faravelli et al. 1980). Ein anderer Teil wurde kaum rezipiert (so die Studie von Hanus & Zapletalek 1984), oder war unter dem Aspekt der Responseprädiktion in Vergessenheit geraten (etwa die Studie von Persson 1972). Offenbar hat sich innerhalb der Übersichtsarbeiten eine Tradition entwickelt, in der vorwiegend frühere Übersichtsarbeiten rezipiert und durch eigene Befunde, sowie einige wenige Originalarbeiten ergänzt wurden. Auf Basis der in der vorliegenden Arbeit durchgeführten Metaanalysen wurden insgesamt 21 Variablen definiert, die nach dem derzeitigen Forschungsstand als Responseprädiktoren gelten können. Wie stark diese Evidenz von den Ergebnissen der informellen Übersichtsarbeiten abweicht, zeigt ein exemplarischer Vergleich mit der jüngsten Übersichtsarbeit zur Lithiumresponse (Walden et al. 1999). In dieser Arbeit werden insgesamt 6 Wirksamkeitsprädiktoren identifiziert, von denen sich lediglich 2 (nämlich das Verlaufsmuster und das Vorliegen eines rapid cycling) mit den 21 im Rahmen der in aufgrund der durchgeführten Metaanalysen definierten Prädiktoren überschneiden.

Sicherlich bietet eine Metaanalyse gegenüber klassischen Übersichtsarbeiten einige Vorteile, etwa eine größere Transparenz der Schlussfolgerungen. Andererseits ist die Fokussierung auf quantitative Aspekte auch mit Nachteilen verbunden – vor allem die nur sehr begrenzte inhaltliche Vertiefung ist ein Manko der vorliegenden Arbeit. In dieser Hinsicht besitzt die vorliegende Arbeit vorbereitenden Charakter für weiterführende Analysen zu einzelnen Variablen (beispielsweise zu „expressed emotions“ oder zu „life events“). Der Verfasser würde sich weitergehende Erkenntnisse zur Responseprädiktion erwarten, wenn die einzelnen Metaanalysen dieser Arbeit durch vertiefte inhaltliche Analysen ergänzt würden – sei es zu den Hintergründen der einzelnen Variablen, sei es zu den Quellen der bei den Effektstärken

einiger Variablen aufgetretenen Heterogenität zwischen den Studien. Insgesamt wurden 68 klinische, demographische und psychosoziale Variablen identifiziert, die in mindestens einer Primärstudie mit dem Erfolg einer Lithiumprophylaxe bei bipolaren Patienten in einen quantifizierbaren Zusammenhang gebracht werden konnten. Vertiefte inhaltliche Analysen zu diesen Variablen hätten den Rahmen dieser Arbeit gesprengt. Zur Anzahl der Variablen ist anzumerken, daß diese Zahl insbesondere bei psychosozialen Faktoren nicht von vornherein klar definiert ist. Bei einigen Variablen stellt sich die Frage, ob Ergebnisse zu inhaltlich ähnlichen Variablen aus unterschiedlichen Studien zusammengefasst werden sollten, oder nicht. Im Zweifelsfall wurde eine solche Vereinheitlichung (etwa von Reizbarkeit und Aggressivität) nicht vorgenommen. Diese Vorgehensweise entspricht dem eher quantitativ orientierten Ansatz der Arbeit. Für eine stärkere Zusammenfassung ähnlicher Variablen hätte es einer generell weit stärkeren inhaltlichen Arbeit mit fundiertem theoretischen Hintergrund zu den einzelnen Variablen aus den Metaanalysen bedurft.

Zwei weitere Grundsatzentscheidungen, die für die Ergebnisse von weitreichender Bedeutung sind, betreffen die Frage einer  $\alpha$ -Fehler-Adjustierung und die Aggregation heterogener Effektstärken. Bei einem konservativen Ansatz ( $\alpha$ -Fehler-Adjustierung und Aggregation nach einem Modell zufälliger Effekte) werden weit weniger Variablen als Responseprädiktoren definiert als bei einem offensiveren Ansatz. Im Rahmen der Metaanalysen wurde keine  $\alpha$ -Fehler-Adjustierung durchgeführt – es wurde also jede Metaanalyse isoliert betrachtet. Bei insgesamt 68 möglichen Prädiktoren waren bei einer Aggregation über alle Studien 21 Variablen signifikant mit der Response auf eine Lithiumprophylaxe assoziiert. Aufgrund der fehlenden  $\alpha$ -Fehler-Adjustierung wäre von einer Größenordnung von ca. 3 (ca. 5% von 68) Zufallsbefunden auszugehen. Mit Hilfe einer Bonferroni-Adjustierung könnten diese falsch positiven Befunde zwar mit hoher Wahrscheinlichkeit eliminiert werden, doch wäre diese Elimination mit dem Nichterkennen von 12 weiteren, aufgrund von wahrscheinlichkeitstheoretischen Überlegungen vermutlich echten, Prädiktoren verbunden. Dieser Schaden wäre im Kontext dieser Arbeit wohl der größere, so daß die Bonferroni-Adjustierung nicht durchgeführt wurde. Bei der Aggregation der Effektstärken wurde hingegen konservativ vorgegangen. Aufgrund inhaltlicher Überlegungen und aufgrund der immer wieder aufgetretenen Heterogenität der für die einzelnen Studien ermittelten Effektstärken wurden die Daten grundsätzlich nach einem Modell zufälliger Effekte aggregiert (DerSimonian & Laird 1986). Dieser Ansatz umfasst auch den Fall, daß keine überzufällige Heterogenität vorliegt – in diesem Fall entspricht das verwendete Modell dem klassischen Ansatz, bei dem die einzelnen Studien auf Grundlage ihres Stichprobenumfangs gewichtet werden.

Insgesamt stellt sich die Frage, wie sicher für die 21 im Rahmen der Metaanalyse als Responseprädiktoren definierten Variablen ein Zusammenhang mit dem Ansprechen auf eine

Lithiumprophylaxe unterstellt werden kann. Die Frage kann allerdings nicht kurz und bündig beantwortet werden, vor allem, weil zu viele Unbekannte im Spiel sind. So ist es kaum möglich, den Einfluss verschiedener in den Primärstudien auftretender, aber sich natürlich in die Metaanalyse fortsetzender Biasarten abzuschätzen. Beispielsweise ist in retrospektiven Studien mit einem Selektionsbias zu rechnen – so dürften etwa für wenig complianten Patienten tendenziell weniger auswertbare Daten vorliegen, und bei einem vermutlich schlechteren Verlauf dieser Patienten würde in den Primärarbeiten der protektive Einfluss guter Compliance unterschätzt. Eine Vielzahl derartiger Überlegungen ließe sich anstellen. Leider würde bei all diesen Überlegungen das Problem bestehen bleiben, daß sich die Stärke der systematischen Verzerrung kaum abschätzen ließe – auch nicht der Größenordnung nach. Stets ist auch mit einem Publikationsbias zu rechnen – leider gelten signifikante Ergebnisse nach wie vor als publikationswürdiger als sogenannte negative Ergebnisse (vgl. etwa die aktuellen Hinweise für Autoren einer so renommierten Zeitschrift wie den *European Archives of Psychiatry and Clinical Neuroscience* vom Oktober 2003 (Vol. 253, No. 5, o.S.)). Von daher ist insgesamt damit zu rechnen, daß der Effekt einiger Prädiktoren überschätzt wird, oder tatsächlich gar nicht besteht. All diese Effekte pflanzen sich auf die Ebene der Metaanalyse fort. Zwar wurde versucht, einen möglichen Publikationsbias mit Hilfe von Funnelplots aufzuspüren, aber in aller Regel waren zu wenig Primärstudien vorhanden und in einigen Fällen waren die Effektstärken zu heterogen, um einen klaren Schluss zu ermöglichen. Um eine auf Ebene der Metaanalyse entstehende Verzerrung zu vermeiden, sehr viel Mühe für eine möglichst vollständige Lokalisation der Primärstudien betrieben. Insbesondere wurde die Lokalisation der Primärstudien mit sehr breiten Suchbegriffen durchgeführt und durch eine umfangreiche Handsuche ergänzt. Ob das Bestreben nach einer möglichst vollständigen Erfassung der Primärstudien geglückt ist, lässt sich naturgemäß schwer beurteilen. Als Anhaltspunkt für eine Evaluation dieses Schrittes wurde ein Vergleich mit der Metaanalyse von *Viguera et al. (2000)* zur Prädiktion des Erfolgs einer Lithiumprophylaxe aufgrund des Geschlechts angestellt. Im Ergebnis konnten bei vergleichbaren Einschlusskriterien wesentlich mehr als die 9 in *Viguera et al. (2000)* lokalisierten Studien zu bipolaren Patienten lokalisiert werden (21+3 Primärstudien – 21 sind in dem von *Viguera et al.* berücksichtigten Zeitraum erschienen, 3 weitere nach 1998).

Allein die eben geschilderte Problematik systematischer Verzerrungen in den Primärstudien macht deutlich, mit wie viel Vorsicht eine definitive Aussage zu einzelnen Variablen formuliert werden sollte – insbesondere bei der in dieser Arbeit gewählten eher globalen Betrachtungsweise, die eine sorgfältige Abwägung von Bias und Gegenbias bei den einzelnen Variablen lediglich vorbereitet. Dennoch sollen im folgenden einige inhaltliche Hinweise auf Grundlage der durchgeführten Metaanalyse zur Responseprädiktion gegeben werden.



- Die wichtigsten demographischen Variablen, Alter und Geschlecht, scheinen keinen bedeutenden Einfluss auf die Lithiumresponse zu haben. Insbesondere das Geschlecht scheint ohne Einfluss auf die Wirksamkeit von Lithium. Die Befunde zum Geschlecht aus der Primärliteratur sind konsistent und schließen aufgrund der insgesamt relativ großen Fallzahl eine klinisch bedeutsame Effektstärke mit hoher Wahrscheinlichkeit aus.
- Den deutlichsten Anhalt für eine Prognose gibt die Krankenanamnese. Insbesondere für das Verlaufsmuster und für die Anzahl bisheriger Episoden sind die Befunde sehr klar und konsistent. Ein klares Verlaufsmuster Manie – Depression – freies Intervall ist demnach für das Ansprechen auf eine Lithiumprophylaxe günstig. Ein durch fließenden Übergang einer Episode in die andere charakterisierter Krankheitsverlauf ist hingegen als Risikofaktor für ein Rezidiv unter Lithium anzusehen. Diese Ergebnisse stimmen voll mit dem Ergebnis der Metaanalyse von Faedda et al. (1991) zum Zusammenhang zwischen Verlaufsmuster und Lithiumresponse überein. Prognostisch ungünstig ist offenbar auch eine hohe Anzahl bisheriger stationärer Episoden. Auch für diese Variable sind die Befunde außerordentlich klar und konsistent.
- In zweiter Linie scheinen psychosoziale Variablen als Prädiktoren in Betracht zu kommen (in diesem Bereich wurden insgesamt 9 Variablen als Prädiktoren identifiziert). Prognostisch günstig scheinen insbesondere eine hohe soziale Schicht, gute Compliance und gute soziale Unterstützung. Insgesamt liegen allerdings vergleichsweise wenig Befunde zur Responseprädiktion aufgrund psychosozialer Faktoren vor. Gemessen an der prognostischen Bedeutung ist dieser Bereich möglicherweise zu wenig beachtet worden. Hinzu kommt, daß – etwa im Gegensatz zur Anamnese – ein Teil der psychosozialen Variablen (etwa Compliance oder soziale Unterstützung) zumindest prinzipiell beeinflussbar sind, und von dieser Seite her besondere Aufmerksamkeit verdienen.
- Etwas überraschend ist, daß nach derzeitigem Kenntnisstand Variablen aus den Bereichen Diagnostik, Psychopathologie und Komorbidität offenbar keinen deutlichen Anhalt für eine Prognose liefern. Diese Aussage ist allerdings erheblich einzuschränken, weil sie sich auf die im Rahmen dieser Arbeit untersuchte vergleichsweise homogene Gruppe bipolarer Patienten bezieht. Frühere Befunde, etwa daß eine klare Diagnose einer primären affektiven Störung prädiktiv für ein gutes Ansprechen auf Lithium ist (etwa Grof et al. 1983a), werden daher von der vorliegenden Arbeit nicht berührt.

Insgesamt sprechen die Ergebnisse der Metaanalyse dafür, daß es durchaus möglich ist, das Ansprechen auf eine Lithiumresponse vorherzusagen – alles in allem waren 21 Variablen signifikant mit einer Lithiumresponse korreliert, und bei einigen dieser Prädiktorvariablen war der Zusammenhang zur Lithiumresponse sehr klar. Wie gut sich das Ansprechen auf eine Lithiumprophylaxe tatsächlich vorhersagen lässt, wurde exemplarisch anhand eines umfangreichen prospektiven Datensatzes evaluiert. Diese Evaluation ist Gegenstand des zweiten Hauptteils der vorliegenden Arbeit. Hierzu wurde in einem ersten Schritt aufgrund der 21 im Rahmen der Metaanalyse identifizierten Responseprädiktoren ein Vorhersageinstrument konstruiert. Im Ergebnis konnten aufgrund dieses Prognoseinstruments, der Lithium-Response-Skala (LRS), 75% der Verläufe korrekt vorhergesagt werden ( $p = 0.001$ ). Die Trennung des ersten, explorativen, Teils vom zweiten, konfirmatorischen, Teil ist für die Interpretation dieses Ergebnisses wesentlich. Nur, weil die Variablenselektion ausschließlich anhand eines vorgeschalteten hypothesengenerierenden Schrittes – anhand der Metaanalysen – durchgeführt wurde, kann die Trefferquote von 75% als realistische Schätzung der prognostischen Kraft der Lithium-Response-Skala betrachtet werden.

Die Lithium-Response-Skala (LRS) wurde primär an den Anforderungen der klinischen Praxis ausgerichtet. Im Hinblick auf die klinische Praxis wurde die LRS so konstruiert, daß sie einfach und flexibel einsetzbar ist. Insbesondere ist sie aufgrund ihrer Normierung auch bei unvollständiger Information einsetzbar. In jedem Fall bedeutet ein Skalenwert  $> 0$  eine günstige Prognose, und ein Skalenwert  $< 0$  eine ungünstige Prognose – auch bei unvollständigen Angaben. Für die klinische Anwendung wäre eine Trefferquote von 75% sicherlich noch nicht optimal, aber sie wäre klinisch relevant und angesichts sehr pessimistischer Einschätzungen einiger Autoren ist diese Trefferquote als Erfolg zu werten (vgl. etwa Mander (1986, p.35): "It would seem that prediction of outcome and response to lithium cannot be based on clinical factors."). Die LRS beruht auf dem derzeitigen Wissenstand und bezieht sich nur auf klinische, demographische und psychosoziale Variablen. Es ist zu erwarten, daß die Prognosegüte dieses Instruments verbessert wird, wenn die Basis der LRS durch die Hinzunahme von Prädiktoren aus anderen Bereichen erweitert wird. Eine Erweiterung etwa durch biologische oder genetische Prädiktoren ist aufgrund der einfachen und flexiblen Konstruktion der LRS jederzeit möglich. Die Konstruktion der LRS folgt einem streng empirischen Ansatz – es sind genau jene Variablen enthalten, die bei einer quantitativen Analyse des aktuellen Forschungsstandes als Prädiktoren anzusehen sind. Daß sich der Forschungsstand laufend ändert, ist bereits in der Konstruktion der LRS eine Anpassung an den jeweiligen Forschungsstand angelegt.

Selbstverständlich ist in diesem Stadium noch keine definitive Aussage über die Prognosegüte der LRS in verschiedenen klinischen Settings möglich. Die Prognosegüte bezieht sich auf eine Ende der 1980er Jahren an insgesamt 10 deutschen Universitätskliniken rekrutierte

Stichprobe. Wie üblich, konnte schließlich nur ein geringer Teil (in diesem Fall 6%) aller gescreenten Patienten in die Studie eingeschlossen werden, und so stellt sich die Frage, ob die Studienpatienten vielleicht eine untypische Patientenauswahl darstellen. Eine Besonderheit der vorliegenden Studie liegt allerdings darin, daß der Rekrutierungsprozess sorgfältig dokumentiert wurde und die letztendlich eingeschlossenen Patienten so weit wie möglich auf Repräsentativität untersucht wurden (Greil et al. 1993). Im Ergebnis ergaben sich keine wesentlichen Unterschiede der randomisierten Patienten mit jenen, die zwar studiengerecht waren, aber aus verschiedenen Gründen nicht randomisiert werden konnten. Dieses Ergebnis spricht dafür, daß die Studienpatienten als repräsentativ gelten können für stationär behandelte bipolare Patienten mit der Indikation für eine Rezidivprophylaxe. Dennoch ist aus prinzipiellen Überlegungen damit zu rechnen, daß beispielsweise besonders motivierte und complianten Patienten in der Studie überrepräsentiert sind. Inwieweit und in welche Richtung dadurch die Güte der Prognose beeinflusst wird, lässt sich nicht zuverlässig abschätzen.

Die Frage, ob die Responseprädiktoren für Lithium spezifisch sind oder ob es sich eher um allgemeine Verlaufsprädiktoren handelt, lässt sich aufgrund der vorliegenden Arbeit nicht abschließend beantworten. Zwar zeigte sich bei einer Anwendung der LRS auf die Carbamazepingruppe eine im Vergleich zur Lithiumgruppe deutlich niedrigere und statistisch nicht mehr signifikante ( $p = 0.20$ ) Quote korrekter Prognosen, aber eine formelle Testung des Unterschieds der Trefferquoten (75% unter Lithium vs. 63% unter Carbamazepin) erlaubt keine inferenzstatistische Absicherung des Befundes ( $p = 0.23$ ). Die Vermutung liegt nahe, daß sich ein Teil der Prädiktoren bei Lithium spezifisch oder besonders stark ausprägen, während andere Variablen eher einen generell günstigen oder ungünstigen Verlauf anzeigen. Diese Vermutung wird durch die Ergebnisse von Greil et al. (1998) gestützt, wonach drei aufgrund inhaltlicher Überlegungen definierte Variablen (z.B. stimmungsinkongruente Wahninhalte) *spezifisch* das Ansprechen vs. Nichtansprechen auf eine Lithiumprophylaxe präzisieren. Auch zur Trennung spezifischer und allgemeiner Prädiktoren würde der Verfasser weitere Fortschritte von einer stärker inhaltlich orientierten Arbeit erwarten.

Die Ergebnisse der vorliegenden Arbeit sind in der aktuellen Diskussion um den Stellenwert von Lithium in der Prophylaxe bipolarer Störungen (vgl. Kap. 2.4.2.2) von Bedeutung. Seit den 1990er Jahren gibt es einen zunehmenden Trend weg von einer globalen Betrachtungsweise des Einsatzes von Lithium hin zu einem differenzierten, individuellerem Einsatz von Lithium (Grof 1998, Schou 1998). Nach den scharfen Extrempositionen – einerseits Lithium als die Standardtherapie zur Phasenprophylaxe schlechthin, andererseits die Behauptung der generellen Unwirksamkeit (zu dieser Kontroverse vgl. Kap. 2.4.2.2) zeichnet sich eine breit akzeptierte differenziertere Position ab, wonach der Erfolg einer Lithiumprophylaxe entscheidend von zwei Faktoren abhängt. Erstens von einer intensiven Kooperation mit dem Patienten und den Angehörigen des Patienten (vgl. etwa Grof 1997,

Greil et al. 2000), und zweitens von einer sorgfältigen Auswahl der Patienten (etwa Gershon & Soares 1997, Greil et al. 1998, Grof 1998, Schou 1998, Soares & Gershon 2000). Doch genau an dieser Stelle klafft eine Forschungslücke – aufgrund welcher Kriterien sollen die Patienten für eine Lithiumtherapie ausgewählt werden? Für eine wirklich rationale Entscheidung für diese oder jene Therapie wäre es nötig, die Chancen-Risiko-Verhältnisse für eine Lithiumtherapie, die wichtigsten alternativen Therapien (z.B. Lamotrigin, Olanzapin, Carbamazepin) sowie das Chancen-Risiko-Verhältnis für eine Nichtbehandlung auf evidenzbasierter Grundlage abzuwägen. Derzeit sind wir hiervon weit entfernt. Die vorliegende Arbeit versteht sich als Beitrag zu einer evidenzbasierten Einschätzung der Chancen einer Lithiumprophylaxe – für wen ist diese Therapie aussichtsreich, für wen ist sie wenig aussichtsreich. Vor allem aufgrund der großen praktischen Relevanz einer rationalen Entscheidung für die aussichtsreichste Prophylaxe wäre es außerordentlich wünschenswert, die vorliegende Arbeit in zumindest drei Richtungen fortzuführen:

1. durch die bereits angesprochene Erweiterung der Basis der LRS
2. durch Konstruktion ähnlicher Prognoseinstrumente für andere Phasenprophylaktika
3. und schließlich durch eine evidenzbasierte Prädiktion für die mit den verschiedenen prophylaktischen Therapien verbundenen Risiken.

Der für die Konstruktion der LRS gewählte additiv univariate Ansatz scheint am ehesten der aktuellen Forschungslage zu entsprechen. Bei diesem Ansatz werden die verstreuten Einzelinformationen, die für sich genommen eine nur geringe prädiktive Kraft haben, gebündelt und so eine im Vergleich zu Einzelbetrachtungen erheblich verbesserte Prognosegüte erreicht. Die Einbeziehung von Variableninteraktionen im Rahmen multivariater Modellen war hingegen mit keiner weiteren Verbesserung der Prognosegüte verbunden. Der Einfluss der Interaktionen von Prädiktorvariablen ist nur wenig erforscht und lässt derzeit angesichts eines fehlenden kausalen Modells keine sicheren Prognosen zu (vgl. hierzu auch Preisig & Angst 1991). Dies könnte sich in einigen Jahren oder Jahrzehnten ändern, und dann wäre ein multivariates Prognosemodell möglicherweise vielversprechender als eine Skala. Ein Nebenaspekt der vorliegenden Arbeit war, die Möglichkeit der Responseprädiktion aufgrund multivariater Modelle (logistische Regression, Diskriminanzanalyse, Cox-Regression) beim aktuellen Forschungsstand zu untersuchen. Im Vergleich mit der LRS haben sich diese Modelle weniger bewährt. Die Einbeziehung der Variableninteraktionen in die Modellbildung führte zu keiner weiteren Verbesserung der Prognosegüte, sondern ist im Gegenteil mit geringerer Transparenz und vor allem mit erheblichen Schwierigkeiten bei der praktischen Anwendung verbunden. Auch eine kritische Prüfung der publizierten multivariaten Modelle zur Prognose der Lithiumresponse unterstreicht diese Einschätzung. In den meisten Fällen (Kusalic & Engelsmann 1998, Kato 2000) sind aufgrund erheblicher methodischer Mängel – etwa einer fehlenden Kreuz-

validierung bei einem kleinen Datensatz und einer Vielzahl von Variablen – kaum beurteilbar. Die vermutlich sorgfältigste multivariate Auswertung (Grof et al. 1983a,b,c) ist leider methodisch wenig transparent. Der Versuch, das Prognosemodell dieser Arbeit anhand des hier vorliegenden Datensatzes zu replizieren, ist im Ergebnis nicht geglückt – die globale Trefferquote aufgrund einer Diskriminanzanalyse mit den beiden Variablen „Phasenfrequenz“ und „Residualsymptomatik im Intervall“ lag bei lediglich 59% (Ergebnis nicht berichtet). Die dritte Variable des Modells von Grof et al. („eindeutige Diagnose einer primären affektiven Störung“) hatte im vorliegenden Datensatz keine Varianz – alle Patienten hatten eine primäre bipolare Störung. Zusammenfassend lässt sich festhalten, daß die in einzelnen Prädiktorvariablen gegebene Information derzeit wohl am adäquatesten in einem additiv univariaten Ansatz zu einem Prognosemodell gebündelt wird.

Der Gegensatz zwischen univariaten und multivariaten Modellen ist allerdings nicht absolut zu sehen. Es wäre vielversprechend, eine umfassende Klasse von Prognosemodellen zu entwickeln, die beide Modelltypen als Extremfälle enthält. Ein Teil der Variableninteraktionen könnten so wie im additiv univariaten Ansatz a priori Null gesetzt werden, andere Variableninteraktionen könnten für eine Anpassung des Modells vollständig oder teilweise freigegeben werden. Dieser Modelltyp würde dem tatsächlichen Wissensstand in vielen empirischen Forschungsbereichen etwa der Medizin, Psychologie oder Soziologie möglicherweise am besten entsprechen – das Modell könnte dann auch dem Wissensfortschritt entsprechend gezielt bestimmte Variableninteraktionen modellieren.

Zusammenfassend lässt sich festhalten, daß der in dieser Arbeit beschrittene Weg, zur Schaffung einer breiten, evidenz-basierten Entscheidungsgrundlage für die Abwägung der Chancen und Risiken des Beginns einer Lithiumprophylaxe, zumindest teilweise erfolgreich verlaufen ist. Mit Hilfe des auf Grundlage einer Serie von Metaanalysen konstruierten Prognoseinstruments konnten in einem großen und vor allem unabhängigen Datensatz 75% der Verläufe korrekt vorhergesagt werden. Angesichts der bisherigen Ergebnisse zur Prognose des Ansprechens auf eine Lithiumprophylaxe ist dies als Teilerfolg zu werten. Das entwickelte Prognoseinstrument stellt in Hinsicht auf seine prognostische Kraft und im Hinblick auf die klinische Anwendbarkeit einen Fortschritt dar. Es spiegelt allerdings nur den derzeitigen Forschungsstand wider. Es ist zu vermuten, daß sich die Prognosegüte mit der Einbeziehung neuer Erkenntnisse zur Responseprädiktion und mit der Einbeziehung aktueller Erkenntnisse zu weiteren Prädiktoren, insbesondere von biologischen und genetischen Faktoren, weiter verbessert. Jeder kleine Schritt der Verbesserung eines in der klinischen Praxis nutzbaren Prognoseinstruments ist für einige Patienten mit einem rascheren Beginn einer individuell wirksamen Therapie verbunden. In der Konsequenz bedeutet das für diese Patienten das Vermeiden unnötiger manischer oder depressiver Episoden – mit all ihren schwerwiegenden Folgeerscheinungen.

## 8 Literaturverzeichnis

Abou-Saleh MT (1983) Platelet MAO, personality and response to lithium prophylaxis. *J Affect Disord* 5: 55-65

Abou-Saleh MT, Copen A (1986) Who responds to prophylactic lithium? *J Affect Disord* 10: 115-125

Akiskal HS (2003) Validating 'hard' and 'soft' phenotypes within the bipolar spectrum: continuity or discontinuity? *J Affect Disord* 73: 1-5

Alda M (2001) Genetic factors and treatment of mood disorders. *Bipolar Disord* 3: 318-324

Ananth J, Engelsmann F, Kiriakos R, Kolivakis T (1979) Prediction of lithium response. *Acta Psychiatr Scand* 60: 279-286

Ananth J, Pecknold JC (1978) Prediction of lithium response in affective disorders. *J Clin Psychiatry* 39: 95-100

Angst J (1966) Zur Ätiologie und Nosologie endogener depressiver Psychosen. Eine genetische, soziologische und klinische Studie. Springer. Berlin, Heidelberg, New York

Angst J (1978) The course of affective disorders. II. Typology of bipolar manic-depressive illness. *Arch Psychiatr Nervenkr* 226: 65-73

Angst J (1980) Verlauf unipolar depressiver, bipolar manisch-depressiver und schizo-affektiver Erkrankungen und Psychosen. Ergebnisse einer prospektiven Studie. *Fortschr Neurol Psychiatr* 48: 3-30

Angst J (1981) Ungelöste Probleme bei der Indikationsstellung zur Lithiumprophylaxe affektiver und schizo-affektiver Erkrankungen. *Bibl Psychiatr* 161: 34-44

Angst J (1986) The course of affective disorders. *Psychopathology* 19 Suppl 2: 47-52

Angst J (1992) How recurrent and predictable is depressive illness? In: Montgomery SA, Rouillon F (Hrsg.) *Long-term Treatment of Depression*. Wiley & Sons. Chichester, New York, 1-13

Angst J, Gamma A, Benazzi F, Ajdacic V, Eich D, Rössler W (2003) Diagnostic issues in bipolar disorder. *Eur Neuropsychopharmacol* 13 Suppl 2: 43-50

Angst J, Gamma A, Benazzi F, Ajdacic V, Eich D, Rössler W (2003) Toward a re-definition of subthreshold bipolarity: epidemiology and proposed criteria for bipolar-II, minor bipolar disorders and hypomania. *J Affect Disord* 73: 133-146

Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC (1992) A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *JAMA* 268: 240-248

Arbeitsgemeinschaft für Methodik und Dokumentation in der Psychiatrie (Hrsg.) (1979) Das AMDP-System. Springer. Berlin, Heidelberg, New York

Aronoff MS, Epstein RS (1970) Factors associated with poor response to lithium carbonate: a clinical study. *Am J Psychiatry* 127: 472-480

Ayuso-Gutierrez JL, Ramos-Brieva JA (1982) The course of manic-depressive illness. A comparative study of bipolar I and bipolar II patients. *J Affect Disord* 4: 9-14

Baastrup PC, Schou M (1967) Lithium as a Prophylactic Agent. Its Effect Against Recurrent Depressions and Manic-Depressive Psychosis. *Arch Gen Psychiatry* 16: 162-172

Baastrup, PC, Poulsen JC, Schou M, Thomsen K, Amdisen A (1970) Prophylactic lithium: Double-blind discontinuation in manic-depressive and recurrent depressive disorders. *Lancet* 7668: 326-330

Baethge C, Tondo L, Bratti IM, Bschor T, Bauer M, Viguera AC, Baldessarini RJ (2003) Prophylaxis latency and outcome in bipolar disorders. *Can J Psychiatry* 48: 449-455

Baldessarini RJ, Tondo L, Hennen J, Floris G (1999) Latency and episodes before treatment: response to lithium maintenance in bipolar I and II disorders. *Bipolar Disord* 1: 91-97

Baldessarini RJ, Tondo L, Floris G, Hennen J (2000) Effects of rapid cycling on response to lithium maintenance treatment in 360 bipolar I and II disorder patients. *J Affect Disord* 61: 13-22

Baron M, Gershon ES, Rudy V, Jonas WZ, Buchsbaum M (1975) Lithium carbonate response in depression. Prediction by unipolar/bipolar illness, average-evoked response, catechol-O-methyl transferase, and family history. *Arch Gen Psychiatry* 32: 1107-1111

Bech, P (1981) Rating scales for affective disorders: their validity and consistency. *Acta Psychiatr Scand Suppl* 295: 1-101

Belsley D, Kuh E, Welsch R (1980) *Regression diagnostics*. Wiley. New York

Bottlender R, Rudolf D, Strauss A, Möller HJ (2001) Mood-stabilisers reduce the risk of developing antidepressant-induced manic states in acute treatment of bipolar I depressed patients. *J Affect Disord* 63: 79-83

Bouman TK, Niemantsverdriet-van Kampen JG, Ormel J, Slooff CJ (1986) The effectiveness of lithium prophylaxis in bipolar and unipolar depressions and schizo-affective disorders. *J Affect Disord* 11: 275-280

Bowden CL, Calabrese JR, McElroy SL, Gyulai L, Wassef A, Petty F, Pope HG Jr, Chou JC, Keck PE Jr, Rhodes LJ, Swann AC, Hirschfeld RM, Wozniak PJ (2000) A randomized, placebo-controlled 12-month trial of divalproex and lithium in treatment of outpatients with bipolar I disorder. Divalproex Maintenance Study Group. *Arch Gen Psychiatry* 57: 481-489

Bowden CL (2003) Acute and maintenance treatment with mood stabilizers. *Int J Neuropsychopharmacol* 6: 269-275

Bowden CL, Calabrese JR, Sachs G, Yatham LN, Asghar SA, Hompland M, Montgomery P, Earl N, Smoot TM, DeVaugh-Geiss J (2003) Lamictal 606 Study Group. A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently manic or hypomanic patients with bipolar I disorder. *Arch Gen Psychiatry* 60: 392-400

Brochier T, Gay C (1991) Prédiction de l'efficacité des thymorégulateurs. *L'Encéphale* 16 : 391-398

Buchberger R (1997) Intervallsymptomatik affektiver und schizoaffektiver Psychosen unter Rezidivprophylaxe mit Lithium vs Carbamazepin vs Amitriptylin. Unveröffentlichte Dissertation aus der Psychiatrischen Klinik der Universität München



- Calabrese JR, Sullivan JR, Bowden CL, Suppes T, Goldberg JF, Sachs GS, Shelton MD, Goodwin FK, Frye MA, Kusumakar V (2002) Rash in multicenter trials of lamotrigine in mood disorders: clinical relevance and management. *J Clin Psychiatry* 63: 1012-1019
- Calabrese JR, Fatemi SH, Kujawa M, Woyshville MJ (1996) Predictors of response to mood stabilizers. *J Clin Psychopharmacol* 16 Suppl 1: 24-31
- Calabrese JR, Rapport DJ, Shelton MD (2002) Pursuit of the Ideal Mood Stabilizer: Time to Give Up and Move to Combination Trials. In: Maj M, Akiskal HA, Lopez-Ibor JJ, Sartorius N (Hrsg.) *Bipolar Disorder*. Wiley. New York
- Carroll BJ (1979) Prediction of treatment outcome with lithium. *Arch Gen Psychiatry* 36: 870-878
- Cazzullo CL, Sacchetti E, Smeraldi E (1979) Psychotropic drugs and their relationship with psychopathology of affective disorders. *Prog Neuropsychopharmacol* 3: 25-38
- CIPS (Collegium Internationale Psychiatriae Salarum) (1981) *Internationale Skalen für Psychiatrie*. Beltz. Weinheim
- Colom F, Vieta E, Martinez-Aran A, Reinares M, Benabarre A, Gasto C (2000) Clinical factors associated with treatment noncompliance in euthymic bipolar patients. *J Clin Psychiatry* 61: 549-555
- Coppen AM (1979) Peet: The long-term management of patients with affective disorders. In: Paykel, ES, Coppen A (Hrsg.) *Psychopharmacology of affective disorders*. University Press. Oxford, 248-256
- Coryell W, Akiskal H, Leon AC, Turvey C, Solomon D, Endicott J (2000) Family history and symptom levels during treatment for bipolar I affective disorder. *Biol Psychiatry* 47: 1034-1042
- Coryell W, Winokur G (1980) Diagnosis, family and follow-up studies. In: Belmaker RH, van Praag HM (Hrsg.) *Mania. An evolving concept*. Spectrum Publications. Jamaica, New York, 25-41
- Coryell W, Winokur G (1980) Predicting Lithium Responders and Non-Responders: Familial Indicators. In: Johnson FN (Hrsg.) *Handbook of Lithium Therapy*. MTP Press. Lancaster, 137-142

Coryell W, Winokur G, Solomon D, Shea T, Leon A, Keller MB (1997) Lithium and recurrence in a long-term follow-up of bipolar affective disorder. *Psychological Medicine* 27: 281-289

Cundall RL, Brooks PW, Murray LG (1972) A controlled evaluation of lithium prophylaxis in affective disorders. *Psychol Med* 2: 308-311

Darcourt, G (1995) Editorial. *L'Encéphale Suppl* 6: 1-2

Dardennes R, Even C, Bange F, Heim A (1995) Comparison of carbamazepine and lithium in the prophylaxis of bipolar disorders. A meta-analysis. *Br J Psychiatry* 166: 378-381

Deister A, Marneros A (1993) Predicting the long-term outcome of affective disorders. *Acta Psychiatr Scand* 88: 174-177

Deltito J, Martin L, Riefkohl J, Austria B, Kissilenko A, Corless C Morse P (2001) Do patients with borderline personality disorder belong to the bipolar spectrum? *J Affect Disord* 67: 221-228

Denicoff KD, Smith-Jackson EE, Disney ER, Ali SO, Leverich GS, Post RM (1997) Comparative prophylactic efficacy of lithium, carbamazepine, and the combination in bipolar disorder. *J Clin Psychiatry* 58: 470-478

DerSimonian R, Laird N (1986) Meta-Analysis in Clinical Trials. *Controlled Clinical Trials* 7: 177-188

Dick P (1983) L'effet curative du lithium peut-il permettre de prévoir son effet prophylactique? In: Dufour H, Pringuey D, Milech T (Hrsg.) *La Prédiction de la Réponse au Lithium*. Université de Marseille, 149-155

Dickersin K, Scherer R, Lefebvre C (1995) Identifying relevant studies for systematic reviews. In: Chalmers I & Altman DG (Hrsg.) *Systematic Reviews*. BMJ Publishing Group. London, 17-36

Donnelly EF, Goodwin FK, Waldman IN, Murphy DL (1978) Prediction of antidepressant responses to lithium. *Am J Psychiatry* 135: 552-556

Dorus E, Pandey GN (1983) Red Cell Membrane Lithium Transport: Implications for Response to Lithium Treatment. In: Dufour H, Pringuey D, Milech T (Hrsg.) *La Prédiction de la Réponse au Lithium*. Université de Marseille, 97-103

Double DB (1996) Lithium revisited. *Br J Psychiatry* 168: 381-382

Dunner DL, Fleiss JL, Fieve RR (1976) Lithium carbonate prophylaxis failure. *Br J Psychiatry* 129: 40-44

Dunner DL, Stallone F, Fieve RR (1976) Lithium carbonate and affective disorders. V: A double-blind study of prophylaxis of depression in bipolar illness. *Arch Gen Psychiatry* 33: 117-120

Dunner DL, Fieve RR (1974) Clinical factors in lithium carbonate prophylaxis failure. *Arch Gen Psychiatry* 30: 229-233

Engström C, Aström M, Nordqvist-Karlsson B, Adolfsson R, Nylander PO (1997) Relationship between Prophylactic Effect of Lithium Therapy and Family History of Affective Disorders. *Biol Psychiatry* 42: 425-433

Faedda GL, Baldessarini RJ, Tohen M, Strakowski SM, Waternaux C (1991) Episode sequence in bipolar disorder and response to lithium treatment. *Am J Psychiatry* 148: 1237-1239

Faltermaier T, Wittchen HU, Ellmann R, Lassel R (1985) The Social Interview Schedule (SIS). Content, structure and reliability. *Soc Psychiatry* 20: 115-124

Faravelli C, Marchetti G, Benvenuti P, Cabras PL (1980) Il Litio Nella Prevenzione Delle Turbe Cicliche Dell'Affettività. *Riv Patol Nerv Ment* 101: 261-271

Fava, GA (1999) Subclinical symptoms in mood disorders: pathophysiological and therapeutic implications. *Psychol Med* 29: 47-61

Franchini L, Zanardi R, Smeraldi E, Gasperini M (1999) Early onset of lithium prophylaxis as a predictor of good long-term outcome. *Eur Arch Psychiatry Neurol Sci* 249: 227-230

Freeman, MP, Stoll AL (1998) Mood stabilizer combinations: A review of safety and efficacy. *Am J Psychiatry* 155: 12-21

- Frye MA, Ketter TA, Leverich GS, Huggins T, Lantz C, Denicoff KD, Post RM (2000) The increasing use of polypharmacotherapy for refractory mood disorders: 22 years of study. *J Clin Psychiatry* 61: 9-15
- Gasperini M, Scherillo P, Manfredonia MG, Franchini L, Smeraldi E (1993) A study of relapses in subjects with mood disorder on lithium treatment. *Eur Neuropsychopharmacol* 3: 103-110
- Gelenberg AJ, Kane JM, Keller MB, Lavori P, Rosenbaum JF, Cole K, Lavelle J (1989) Comparison of standard and low serum levels of lithium for maintenance treatment of bipolar disorder. *N Engl J Med* 321: 1489-1493
- Gershon S, Soares JC (1997) Current therapeutic profile of lithium. *Arch Gen Psychiatry* 54: 16-20
- Goodwin FK, Murphy DL, Dunner DL, Bunney WE (1972) Lithium response in unipolar versus bipolar depression. *Am J Psychiatry* 129: 44-47
- Goodwin FK, Jamison RK (1990) *Manic-Depressive Illness*. Oxford University Press. New York, Oxford
- Goodwin FK, Fireman B, Simon GE, Hunkeler EM, Lee J, Revicki D (2003) Suicide risk in bipolar disorder during treatment with lithium and divalproex. *JAMA* 290: 1467-1473
- Greden JF, DeVigne JP, Ariavabala A, Tarika J, Buitenheim M, Eiser A, Carroll BJ (1983) Post-dexametasonde plasma cortisol levels among lithium-treated rapid cycling bipolar patients. In: Dufour H, Pringuey D, Milech T (Hrsg.) *La Prédiction de la Réponse au Lithium*. Université de Marseille, 105-112
- Greil W, Broucek B, Klein HE, Engel-Sittenfeld P (1982) Discontinuation of lithium maintenance therapy: reversibility of clinical, psychological and neuroendocrinological changes. In: Emrich HR, Aldenhoff JB, Lux HD (Hrsg.) *Basic mechanisms in the action of lithium*. Excerpta Medica. Amsterdam, Oxford, Princeton, 235-248
- Greil W, Schmidt ST (1985) Medikamentöse Rezidivverhütung von affektiven Psychosen. In: Hippus H, Greil W (Hrsg.) *Psychiatrie für die Praxis 2: Diagnostik und Therapie depressiver Störungen*. MMV Medizin Verlag. München, 54-71

Greil W, Ludwig-Mayerhofer W, Steller B, Czernik A, Giedke H, Müller-Oerlinghausen B, Osterheider M, Rudolf GAE, Sauer H, Tegeler J, Wetterling T (1993) The recruitment process for a multicenter study on the long-term prophylactic treatment of affective disorders. *J Affect Disord* 28: 257-265

Greil W, Sassim N, Ströbel-Sassim C (1996) Die manisch- depressive Krankheit: Therapie mit Carbamazepin. 2. überarbeitete u. erweiterte Auflage. Thieme. Stuttgart

Greil W, Ludwig-Mayerhofer W, Erazo E, Engel RR, Czernik A, Giedke H, Müller-Oerlinghausen B, Osterheider M, Rudolf GAE, Sauer H, Tegeler J, Wetterling T (1996) Comparative efficacy of lithium and amitriptyline in the maintenance treatment of recurrent unipolar depression: a randomised study. *J Affect Disord* 40: 179-190

Greil W, Ludwig-Mayerhofer W, Erazo N, Schöchlin N, Engel RR, Czernik A, Giedke H, Müller-Oerlinghausen B, Osterheider M, Rudolf GAE, Sauer H, Tegeler J, Wetterling T (1997). Lithium versus carbamazepine in the maintenance treatment of bipolar disorders - a randomised study. *J Affect Disord* 43: 151-161

Greil W, Kleindienst N: Rezidivprophylaxe affektiver Störungen mit Lithium (1997) In: Müller-Oerlinghausen B, Greil W, Berghöfer A (Hrsg.): Die Lithiumtherapie: Nutzen, Risiken, Alternativen. 2. Auflage. Springer. Berlin, Heidelberg, New York, 190-218

Greil W, Kleindienst N, Erazo N, Müller-Oerlinghausen B (1998) Differential response to lithium and carbamazepine in the prophylaxis of bipolar disorder. *J Clin Psychopharmacol* 18: 455-460

Greil W, Kleindienst N, Schlösser S (2000) Medikamentöse Rezidivprophylaxe affektiver Störungen. In: Möller HJ (Hrsg.) Therapie psychiatrischer Erkrankungen, 2. Auflage. Enke-Verlag. Stuttgart, 424-450

Greil W, Kleindienst N (2003) Concepts in the Treatment of Bipolar Disorder. *Acta Psychiatr Scand Suppl* 418: 41-46

Grof P, Hux M, Grof E, Arato M (1983a) Prediction of response to stabilizing lithium treatment. *Pharmacopsychiatry* 16: 195-200

Grof P (1983b) Response to Long-Term Lithium Treatment: Research Studies and Clinical Implications. In: Davis JM & Maas JW (Hrsg.) Affective Disorders. American Psychiatric Press. Washington, 357-366

Grof P, Lane J, Hux M, Daigle L, Grof E, Werstiuk E (1983c) Predicting the response to long-term lithium treatment... Interim results of a prospective study. In: Dufour H, Pringuey D, Milech T (Hrsg.) *La Prédiction de la Réponse au Lithium*. Université de Marseille, 43-53

Grof E, Haag M, Grof P, Haag H (1987) Lithium response and the sequence of episode polarities: preliminary report on a Hamilton sample. *Prog Neuropsychopharmacol Biol Psychiatry* 11: 199-203

Grof P, Alda M, Grof E, Fox D, Cameron P (1993) The challenge of predicting response to stabilising lithium treatment. The importance of patient selection. *Br J Psychiatry Suppl*:16-19

Grof P (1997) Selektionskriterien für kurative und prophylaktische Lithiumbehandlung. In: Müller-Oerlinghausen B., Greil W, Berghöfer A (Hrsg.) *Die Lithiumtherapie: Nutzen, Risiken, Alternativen*. 2. Auflage. Springer. Berlin, Heidelberg, New York, 232-240

Grof P (1998) Has the effectiveness of lithium changed? Impact of the variety of lithium's effects. *Neuropsychopharmacology* 19: 183-188

Grof P, Duffy A, Cavazzoni P, Grof E, Garnham J, MacDougall M, O'Donovan C, Alda M (2002) Is response to prophylactic lithium a familial trait? *J Clin Psychiatry* 63: 942-947

Guscott R, Taylor L (1994) Lithium prophylaxis in recurrent affective illness. *Br J Psychiatry* 164: 741-746

Haag M, Heidorn A, Haag H, Greil W (1986) Response to Stabilising Lithium Therapy and Sequence of Affective Polarity. *Pharmacopsychiatry* 19: 278-279

Haag M, Heidorn A, Haag H, Greil W (1987) Sequence of Affective Polarity and Lithium Response: Preliminary Report on Munich Sample. *Neuro-Psychopharmacol. & Biol. Psychiat* 11: 205-208

Hanus H, Zapletal M (1984) The prophylactic lithium treatment in affective disorders and the possibilities of the outcome prediction. *Sb Ved Pr Lek Fak Karlovy Univerzity Hradci Kralove* 27: 5-75

Hartong EG, Moleman P, Hoogduin CA, Broekman TG, Nolen WA; LitCar Group (2003) Prophylactic efficacy of lithium versus carbamazepine in treatment-naive bipolar patients. *J Clin Psychiatry* 64: 144-151

Hegerl U, Mavrogiorgou P (1997) Prädiktive Bedeutung evozierter kortikaler Potentiale. In: Müller-Oerlinghausen B, Greil W, Berghöfer A (Hrsg.) Die Lithiumtherapie: Nutzen, Risiken, Alternativen. 2. Auflage. Springer. Berlin, Heidelberg, New York, 241-250

Johnson FN (1980) Predicting Lithium Responders and Non-Responders: Psychological Indicators. In: Johnson FN (Hrsg.) Handbook of Lithium Therapy. MTP Press. Lancaster, 126-132

Johnston BB, Naylor GJ, Dick EG, Hopwood SE, Dick DA (1980) Prediction of clinical course of bipolar manic depressive illness treated with lithium. *Psychol Med* 10: 329-334

Kane JM, Quitkin FM, Rifkin A, Ramos-Lorenzi JR, Nayak DD, Howard A (1982) Lithium carbonate and imipramine in the prophylaxis of unipolar and bipolar II illness: a prospective, placebo-controlled comparison. *Arch Gen Psychiatry* 39: 1065-1069

Kato T, Inubushi T, Kato N (2000) Prediction of lithium response by 31P-MRS in bipolar disorder. *Int J Neuropsychopharmacol* 3: 83-85

Keck PE Jr, McElroy SL, Strakowski SM, Bourne ML, West SA (1997) Compliance with maintenance treatment in bipolar disorder. *Psychopharmacol Bull* 33: 87-91

Ketter TA, Wang PW (2002) Predictors of treatment response in bipolar disorders: evidence from clinical and brain imaging studies. *J Clin Psychiatry* 63 Suppl 3: 21-25

Klein HE, Broucek B, Greil W (1981) Lithium withdrawal triggers psychotic states *Br J Psychiatry* 139: 255-256

Kleindienst N, Greil W, Rüger B, Möller HJ (1999) The prophylactic efficacy of lithium-transient or persistent? *Eur Arch Psychiatry Neurol Sci* 249: 144-149

Kleindienst N, Greil W (2002) Inter-episodic morbidity and drop-out under carbamazepine and lithium in the maintenance treatment of bipolar disorder. *Psychol Med* 32: 493-501

Kleindienst N, Greil W (2003) Lithium in the long-term treatment of bipolar disorders. *Eur Arch Psychiatry Neurol Sci* 253: 120-125

Kocsis JH, Stokes PE (1979) Lithium maintenance: factors affecting outcome. *Am J Psychiatry* 136: 563-566

Koukopoulos A, Reginaldi D, Minnai G, Serra G, Pani L, Johnson FN (1995) The long term Prophylaxis of Affective Disorders. In: Gessa G, Fratta W, Pani L, Serra G (Hrsg.) *Depression and Mania: From Neurobiology to Treatment*. Raven Press. New York, 127-147

Kukopulos A, Reginaldi D, Laddomada P, Floris G, Serra G, Tondo L (1980) Course of the manic-depressive cycle and changes caused by treatment. *Pharmakopsychiatry* 13: 156-167

Kukopulos A, Reginaldi D (1980) Recurrences of Manie-Depressive Episodes During Lithium Treatment. In: Johnson FN (Hrsg.) *Handbook of Lithium Therapy*. University of Lancaster: 109-117

Kulhara P, Basu D, Mattoo SK, Sharan P, Chopra R (1999) Lithium prophylaxis of recurrent bipolar affective disorder: long-term outcome and its psychosocial correlates. *J Affect Disord* 54: 87-96

Kupfer, DJ (1991) Long-term treatment of depression. *Journal of Clinical Psychiatry* 52 Suppl 5: 28-34

Kusalic M, Engelsmann F (1998) Predictors of lithium treatment responsiveness in bipolar patients. A two-year prospective study. *Neuropsychobiology* 37: 146-149

Lachenbruch PA, Mickey MA (1968) Estimation of Error Rates in Discriminant Analysis. *Technometrics* 10: 1-10

Laupacis A, Sackett DL, Roberts RS (1988) An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 318: 1728-1733

Leverich GS, Post RM, Rosoff AS (1990) Factors associated with relapse during maintenance treatment of affective disorders. *Int Clin Psychopharmacology* 5: 135-156

Linden M, Nather J and Wilms HU (1988) Zur Definition, Bedeutung und Messung der Krankheitskonzepte von Patienten. Die Krankheitskonzeptskala (KK-Skala) für schizophrene Patienten. *Fortschr Neurol Psychiatr* 56: 35-43

Lusznat RM, Murphy DP, Nunn CM (1988) Carbamazepine vs lithium in the treatment and prophylaxis of mania. *Br J Psychiatry* 153: 198-204



Maj M, Starace F, Nolfi G, Kemali D (1986) Minimum plasma lithium levels required for effective prophylaxis in DSM III bipolar disorder: a prospective study. *Pharmacopsychiatry* 19: 420-423

Maj M, Pirozzi R, Kemali D (1989) Long-term outcome of lithium prophylaxis in patients initially classified as complete responders. *Psychopharmacology* 98: 535-538

Maj M, Pirozzi R, Starace F (1989) Previous pattern of course of the illness as a predictor of response to lithium prophylaxis in bipolar patients. *J Affect Disord* 17: 237-241

Maj M (1990) Clinical prediction of response to lithium prophylaxis in bipolar patients: the importance of the previous pattern of course of the illness. *Clin Neuropharmacol* 13 Suppl 1: 66-70

Maj M, Pirozzi R, Magliano L (1996) Late non-response to lithium prophylaxis in bipolar patients: prevalence and predictors. *J Affect Disord* 39: 39-42

Maj M, Pirozzi R, Magliano L, Bartoli L (1998) Long-term outcome of lithium prophylaxis in bipolar disorder: a 5-year prospective study of 402 patients at a lithium clinic. *Am J Psychiatry* 155: 30-35

Maj M, Arena F, Lovero N, Pirozzi R, Kemali D (1985) Factors associated with response to lithium prophylaxis in DSM III major depression and bipolar disorder. *Pharmacopsychiatry* 18: 309-313

Maj M (2000) The impact of lithium prophylaxis on the course of bipolar disorder: a review of the research evidence. *Bipolar Disord* 2: 93-101

Mander AJ (1986) Clinical prediction of outcome and lithium response in bipolar affective disorder. *J Affect Disord* 11: 35-41

Marini JL, Sheard MH, Bridges CI, Wagner E (1976) An evaluation of the double-blind design in a study comparing lithium carbonate with placebo. *Acta Psychiatr Scand* 53: 343-354

Marneros A (1999) *Handbuch der unipolaren und bipolaren Erkrankungen*. Thieme. Stuttgart, New York

Marneros A (2001) Expanding the group of bipolar disorders. *J Affect Disord* 62: 39-44

Melia PI (1970) Prophylactic lithium: a double-blind trial in recurrent affective disorders. *Br J Psychiatry* 116: 621-624

Mendels J (1976) Lithium in the treatment of depression. *Am J Psychiatry* 133: 373-378

Mendels J, Frazer A (1973) Intracellular lithium concentration and clinical response: towards a membrane theory of depression. *J Psychiatr Res* 10: 9-18

Mendlewicz J, Fieve RR, Stallone F (1973) Relationship between the effectiveness of lithium therapy and family history. *Am J Psychiatry* 130: 1011-1013

Mendlewicz J (1979) Prediction of Treatment Outcome: Family and Twin Studies in Lithium Prophylaxis and the Question of Lithium Red Blood Cell/Plasma Ratios. In: Cooper TB, Gershon S, Kline NS, Schou M (Hrsg.) *Lithium: Controversies and Unresolved Issues*. Excerpta Medica. Amsterdam

Miklowitz DJ, Goldstein MJ, Nuechterlein KH, Snyder KS, Mintz J (1988) Family factors and the course of bipolar affective disorder. *Arch Gen Psychiatry* 45: 225-231

Minsker EI, Panteleeva GP, Gindilis VM, Mikhailova NM, Beliaev BS (1977) K voprusu o znachenii klinicheskikh i farmakokineticheskikh parametrov v predskananii effektivnosti profilakticheskogo mimeniia solei litiia u bol'nykh affektivnymi psikhozmi. *Zh Nevropatol Psikhiatr Im S S Korsakova* 77: 1209-1214

Moncrieff J (1995) Lithium revisited. A re-examination of the placebo-controlled trials of lithium prophylaxis in manic-depressive disorder. *Br J Psychiatry* 167: 569-573

Moncrieff, J (1997) Lithium: evidence reconsidered. *Br J Psychiatry* 171: 113-119

Müller-Oerlinghausen B, Berghöfer A, Ahrens B (2003) The antisuicidal and mortality-reducing effect of lithium prophylaxis: consequences for guidelines in clinical psychiatry. *Can J Psychiatry* 48: 433-439

Mulrow CD (1987) The medical review article: state of the science. *Ann Intern Med* 107: 485-488

Mulrow CD (1995) Rationale for systematic reviews. In: Chalmers I, Altman DG (Hrsg.) *Systematic Reviews*. BMJ Publishing Group. London, 1-8

- Murray N, Hopwood S, Balfour DJ, Ogston S, Hewick DS (1983) The influence of age on lithium efficacy and side-effects in out-patients. *Psychol Med* 13: 53-60
- O'Connell RA, Mayo JA, Flatow L, Cuthbertson B, O'Brien B (1991) Outcome of bipolar disorder on long-term treatment with lithium. *British Journal of Psychiatry* 159: 123-129
- Okuma T (1993) Effects of carbamazepine and lithium on affective disorders. *Neuropsychobiology* 27: 138-145
- Overall JE, Gorham DR (1970) 06-BPRS. Brief Psychiatric Rating Scale. In: Guy W, Bonato RR (Hrsg.) *Manual for the ECDEU Assessment Battery*. Chevy Chase, Maryland, 06-1 – 06-19
- Oxtoby A, Jones A, Robinson M (1989) Is your "double-blind" design truly double-blind? *Br J Psychiatry* 155: 700-701
- Page C, Benaim S, Lappin F (1987) A long-term retrospective follow-up study of patients treated with prophylactic lithium carbonate. *Br J Psychiatry* 150: 175-179
- Pandey GN, Baker J, Chang S, Davis JM (1978) Prediction of in vivo red cell/plasma  $LI^+$  ratio by in vitro methods. *Clin Pharmacol Ther* 24: 343-349
- Perris C (1966) A study of bipolar (manic-depressive) and unipolar recurrent depressive psychoses. *Acta Psychiatr Scand Suppl* 194: 9-14
- Persson G (1972) Lithium prophylaxis in affective disorders. An open trial with matched controls. *Acta Psychiatr Scand* 90: 295-297
- Pies R (2002) Have we undersold lithium for bipolar disorder? *J Clin Psychopharmacol* 22: 445-449
- Post RM, Leverich GS, Pazzaglia PJ, Mikalauskas K, Denicoff K (1993) Lithium tolerance and discontinuation as pathways to refractoriness. In: Birch NJ, Padgham C, Hughes MS (Hrsg.) *Lithium in Medicine and Biology*. Marius Press. Huddersfield
- Post RM, Ketter TA, Denicoff K, Pazzaglia PJ, Leverich GS, Marangell LB, Callahan AM, George MS, Frye MA (1996) The place of anticonvulsant therapy in bipolar illness. *Psychopharmacology* 128: 115-129

Preisig M, Angst J (1991) Maladies dépressives. Facteurs du risque du récidence. *Encéphale* 17: 365-372

Prien, RF (1988) Maintenance treatment of depressive and manic states In: Georgotas A, Cancro R (Hrsg.) *Depression and mania*. Elsevier. New York, 439-451

Prien RF, Caffey EM Jr, Klett J (1973) Prophylactic efficacy of lithium carbonate in manic depressive illness. *Arch Gen Psychiatry* 28: 337-341

Prien RF, Caffey EM Jr, Klett CJ (1974) Factors associated with treatment success in lithium carbonate prophylaxis. Report of the Veterans Administration and National Institute of Mental Health collaborative study group. *Arch Gen Psychiatry* 31: 189-192

Prien RF, Kupfer DJ, Mansky PA, Small JG, Tuason VB, Voss CB, Johnson WE (1984) Drug therapy in the prevention of recurrences in unipolar and bipolar affective disorders. Report of the NIMH Collaborative Study Group comparing lithium carbonate, imipramine, and a lithium carbonate-imipramine combination. *Arch Gen Psychiatry* 41: 1096-1104

Rosenthal NE, Rosenthal LN, Stallone F, Fleiss J, Dunner DL, Fieve RR (1979) Psychosis as a predictor of response to lithium maintenance treatment in bipolar affective disorder. *J Affect Disord* 1: 237-245

Rosenthal R (1991) *Meta-analytic procedures for social research*. Sage. Newbury Park, London, New Delhi

Rybakowski & Strzyzewski (1976) Red-blood-cell lithium index and long-term maintenance treatment. *Lancet* 1(7974): 1408-1409

Rybakowski J, Frazer A, Mendels J, Ramsey TA (1977) Prediction of the lithium ratio in man by means of an in vitro test. *Clin Pharmacol Ther* 22: 465-469

Sachs GS, Thase ME (2000) Bipolar disorder therapeutics: maintenance treatment. *Biol Psychiatry* 48: 573-581

Sarantidis D, Waters B (1981) Predictors of lithium prophylaxis effectiveness. *Prog Neuropsychopharmacol* 5: 507-510

Schneeweiß H (1990) *Ökonometrie*. Physica-Verlag. Heidelberg

- Schou M, Baastrup PC, Grof P, Weis P, Angst J (1970) Pharmacological and clinical problems of lithium prophylaxis. *Br J Psychiatry* 116: 615-619
- Schou M (1997) The combat of non-compliance during prophylactic lithium treatment. *Acta Psychiatr Scand* 95: 361-363
- Schou M (1998) Has the time come to abandon prophylactic lithium treatment? A review for clinicians. *Pharmacopsychiatry* 31: 210-215
- Schurhoff F, Bellivier F, Jouvent R, Mouren-Simeoni MC, Bouvard M, Allilaire JF, Leboyer M (2000) Early and late onset bipolar disorders: two different forms of manic-depressive illness? *J Affect Disord* 58: 215-221
- Schwarzer G, Galandi D, Antes G, Schumacher M (2000) Meta-Analyse randomisierter klinischer Studien, Publikations-Bias und Evidence-Based Medicine (EBM). *Informatik, Biometrie und Epidemiologie in Medizin und Biologie* 31: 1-21
- Serreti A, Artioli P (2003) Predicting Response to Lithium in Mood Disorders. Role of Genetic Polymorphisms. *Am J Pharmacogenomics* 3: 17-30
- Sinclair JC, Bracken MB (1994) Clinically useful measures of effect in binary analyses of randomized trials. *J Clin Epidemiol* 47: 881-889
- Small JG, Klapper MH, Milstein V, Kellams JJ, Miller MJ, Marhenke JD, Small IF (1991) Carbamazepine compared with lithium in the treatment of mania. *Arch Gen Psychiatry* 48: 915-921
- Soares JC, Gershon S (2000) The psychopharmacologic specificity of the lithium ion: origins and trajectory. *J Clin Psychiatry* 61 Suppl 9: 16-22
- Solomon DA, Keitner GI, Miller IW, Shea MT, Keller MB (1995) Course of illness and maintenance treatments for patients with bipolar disorder. *J Clin Psychiatry* 56: 5-13
- Spitzer RL, Endicott J, Robins E (1982) Forschungsdiagnose-Kriterien (RDC) (translated by Klein HE). Weinheim: Beltz. (Orig: Spitzer RL, Endicott J, Robins E (1978) Research Diagnostic Criteria (RDC). Biometrics Research Department. New York State Psychiatric Institut, New York.)

- Stallone F, Shelley E, Mendlewicz J, Fieve RR (1973) The use of lithium in affective disorders. 3. A double-blind study of prophylaxis in bipolar illness. *Am J Psychiatry* 130: 1006-1010
- Stefos G, Bauwens F, Staner L, Pardoën D, Mendlewicz J (1996) Psychosocial predictors of major affective recurrences in bipolar disorder: a 4-year longitudinal study of patients on prophylactic treatment. *Acta Psychiatr Scand* 93: 420-426
- Stoll AL, Tohen M, Baldessarini RJ, Goodwin DC, Stein S, Katz S, Geenens D, Swinson RP, Goethe JW, McGlashan T (1993) Shifts in diagnostic frequencies of schizophrenia and major affective disorders at six North American psychiatric hospitals, 1972-1988. *Am J Psychiatry* 150: 1668-1673
- Suppes T, Baldessarini RJ, Faedda GL, Tohen M (1991) Risk of recurrence following discontinuation of lithium treatment in bipolar disorder. *Arch Gen Psychiatry* 48: 1082-1088
- Sussman N (2003) The implications of weight changes with antipsychotic treatment. *J Clin Psychopharmacol* 23 Suppl 1: 21-26
- Sutter J (1983) Conclusions. In: Dufour H, Pringuey D, Milech T (Hrsg.) *La Prédiction de la Réponse au Lithium*. Université de Marseille, 251-254
- Tellenbach H (1961) *Melancholie: Zur Problemgeschichte, Typologie, Pathogenese und Klinik*. Springer. Berlin
- Tohen M, Marneros A, Greil W, Cassano GB, Koukopoulos A, Belmaker H, Calabrese JR, Bowden CL, Yatham L, Licht R, Sachs GS, Evans ARA, Risser R, Baker RW, Grunze H, Crane H, Dossenbach MR, Müller-Oerlinghausen B (zur Publikation eingereicht) Olanzapine versus Lithium in Relapse Prevention in Bipolar Disorder: A Randomized Double-blind Controlled 12-month Clinical Trial.
- Tondo L, Baldessarini RJ, Hennen J, Floris G (1998) Lithium maintenance treatment of depression and mania in bipolar I and bipolar II disorders. *Am J Psychiatry* 155: 638-645
- Tondo L, Baldessarini RJ, Floris G (2001) Long-term clinical effectiveness of lithium maintenance treatment in types I and II bipolar disorders. *Br J Psychiatry Suppl* 41: 184-190

Tondo L, Hennen J, Baldessarini RJ (2003) Rapid-cycling bipolar disorder: effects of long-term treatments. *Acta Psychiatr Scand* 108: 4-14

Upadhyaya AK, Varma VK, Sankaranarayanan A, Goel A (1985) Lithium in prophylactic therapy of manic-depressive illness: biochemical correlates of response. *Biol Psychiatry* 20: 202-205

Van der Velde CD (1970) Effectiveness of lithium carbonate in the treatment of manic-depressive illness. *Am J Psychiatry* 127: 345-351

Vaughn C, Leff JP (1976) The measurement of expressed emotion in the families of psychiatric patients. *Br J Soc Clin Psychol* 15: 157-165

Vaughn CE, Leff JP (1981) Patterns of emotional response in relatives of schizophrenic patients. *Schizophr Bull* 7: 43-44

Vieta E, Gasto C, Otero A, Nieto E, Vallejo J (1997) Differential features between bipolar I and bipolar II disorder. *Compr Psychiatry* 38: 98-101

Viguera AC, Tondo L, Baldessarini RJ (2000) Sex differences in response to lithium treatment. *Am J Psychiatry* 157: 1509-1511

Viguera AC, Baldessarini RJ, Tondo L (2001) Response to lithium maintenance treatment in bipolar disorders: comparison of women and men. *Bipolar Disord* 3: 245-252

Grunze H (1999) Wirksamkeitsprädiktoren für Phasenprophylaktika (Stimmungsstabilisierer) bei bipolar affektiven Störungen. *Fortschr Neurol Psychiatr* 67: 75-80

Waters B, Lapierre Y, Gagnon A, Cahudhry R, Tremblay A, Sarantidis D, Gray R (1982) Determination of the Optimal Concentration of Lithium for the Prophylaxis of Manic-Depressive Disorder. *Biol Psychiatry* 17: 1323-1329

Watkins SE, Callender K, Thomas DR, Tidmarsh SF, Shaw DM (1987) The effect of carbamazepine and lithium on remission from affective illness. *Br J Psychiatry* 150: 180-182

Winokur G, Clayton PJ (1967) Family history studies: I. Two types of affective disorders separated according to genetic and clinical factors. In: Wortis J (Hrsg.) *Recent Advances in Biological Psychiatry*. Vol. 10. Plenum. New York

Wittchen HU, Zaudig M, Schramm E, Mombour W, Klug J, Horn R (1987) Strukturiertes Klinisches Interview für DSM-III-R (SKID - P/4/87). Weinheim: Beltz. (Orig: Spitzer RL, Williams JB, Gibbon M (1984) Structured Clinical Interview for DSM-III-R (SCID). Biometrics Research Department. New York State Psychiatric Institut. New York.)

Yazici O, Kora K, Ücok A, Tunali D, Turan N (1999) Predictors of Lithium prophylaxis in bipolar patients. *J Affect Disord* 55: 133-142

Yang YY (1985) Prophylactic efficacy of lithium and its effective plasma levels in Chinese bipolar patients. *Acta Psychiatr Scand* 71: 171-175

Young LT, Cooke RG, Robb JC, Levitt AJ, Joffe RT (1993) Anxious and non-anxious bipolar disorder. *J Affect Disord* 29: 49-52

Tellenbach H (1961) *Melancholie: Zur Problemgeschichte, Typologie, Pathogenese und Klinik*. Springer. Berlin, Heidelberg, New York

Zerssen D von (1973) Selbstbeurteilungs-Skalen zur Abschätzung des „subjektiven Befunds“ in psychopathologischen Querschnitt- und Längsschnitt-Untersuchungen. *Arch Psychiatr Nervenkr* 217: 299-314

Zerssen D von, Pfister H, Koeller DM (1988) The Munich Personality Test (MPT) a short questionnaire for self-rating and relatives' rating of personality traits: formal properties and clinical potential. *Eur Arch Psychiatry Neurol Sci* 238: 73-93

Zerssen D von (1981) Die Befindlichkeits-Skala. In: CIPS (Hrsg.) *Internationale Skalen für Psychiatrie*. Beltz. Weinheim



## Danksagung

Die vorliegende Arbeit hat sehr durch die Zusammenarbeit mit einer Reihe von Personen profitiert. Besonders hervorheben möchte ich die Unterstützung durch Prof. Dr. Waldemar Greil und Prof. Dr. Rolf R. Engel.

Prof. Dr. Greil hat nicht nur die Multizenterstudie, die Grundlage eines zentralen Kapitels meiner Dissertation war, in der Planungs- und Erhebungsphase geleitet. Vielmehr hält er seit mehr als 10 Jahren die Infrastruktur für eine umfassende Auswertung dieser Daten aufrecht. Nur aufgrund dieses Umstands war es überhaupt möglich, die vorliegende Arbeit in dieser Form zu schreiben. Darüber hinaus hat mich Herr Prof. Dr. Greil jederzeit unterstützt – etwa durch eine Vielzahl von Fachdiskussionen oder allgemein durch seine kritisch-wohlwollende Begleitung der Arbeit. Hierfür möchte ich mich herzlich bedanken.

Prof. Dr. Engel hat mir sehr wertvolle Anregungen und Hinweise gegeben, von der die vorliegende Arbeit außerordentlich profitiert hat. Ich konnte ihn jederzeit um Rat fragen und bekam immer eine Antwort, die mich wirklich weiter brachte. Darüber hinaus verdanke ich der Zusammenarbeit mit Prof. Dr. Engel meine Kenntnisse im Bereich metaanalytischer Methoden. Er half mir, dieses Gebiet überhaupt erst zu erschließen. Hiervon habe ich persönlich wie auch die vorliegende Arbeit sehr profitiert. Auch Herrn Prof. Dr. Rolf Engel gilt mein herzlicher Dank.

Die Bibliothekarinnen der Emil-Kraepelin-Bibliothek haben bei der Beschaffung schwer zugänglicher Literatur vieles möglich gemacht, was bei weitem nicht selbstverständlich ist. Bei Frau Elke Sund, Frau Marianne Kratzer und Frau Margarite Schuhmacher möchte ich mich hierfür bedanken.

Bedanken möchte ich mich auch bei Frau Sandra Dittmann und bei Frau Eliza Kozuch. Sie haben meine Arbeit aufmerksam und kritisch korrekturgelesen. Die Endfassung der Dissertation hat hiervon sehr profitiert.

Dem Leiter der Psychiatrischen Klinik der LMU, Prof. Dr. Hans-Jürgen Möller danke ich für die Bereitstellung hervorragender Rahmenbedingungen, die mir ein effektives Arbeiten ermöglichten.

Schließlich möchte ich mich herzlichst bei meinen Eltern bedanken, die mich über die Jahre meiner Ausbildung finanziell wie persönlich unterstützt haben.

## Lebenslauf

### PERSÖNLICHE DATEN

---

Geburtsdatum: 27.7.1965

Geburtsort: Lindenberg/Allgäu

Eltern: Walter Kleindienst, Schreiner  
Sunhild Kleindienst (geb. Fitz), Kauffrau

### AUSBILDUNG, BERUF

---

Schulbildung: 1971-1975: Grundschule in Kalzhofen  
1975-1984: mathematisch naturwissenschaftliches  
Gymnasium in Lindenberg/Allgäu

Hochschulstudium: 1985-1989: Lehramtsstudium (Französisch / Geschichte /  
Latein) in Konstanz und Innsbruck  
1987-1988: Lehrtätigkeit im Rahmen des  
Lehramtsstudiums in Armentières, Frankreich  
1989-1995: Studium der Statistik in München  
(Anwendungsgebiet: Psychologie)  
30.11.1995: Abschluß zum Diplom-Statistiker

Berufliche Tätigkeiten: 12/95-5/96: Wissenschaftlicher Assistent am Institut für  
Statistik (Arbeitsgruppe Prof. Rüger)

seit 6/96: Wissenschaftlicher Mitarbeiter an der  
Psychiatrischen Klinik der Ludwig-Maximilians-  
Universität München (Arbeitsgruppe Prof. Greil)

seit 2/99: Mitarbeit im Qualitätssicherungsprojekt der  
Psychiatrischen Privatklinik Sanatorium  
Kilchberg (CH)

9/02 – 5/03: Beratung der Gesundheitsdirektion des Kantons  
Zürich (Projekt „Outcome-Messung Psychiatrie“)