MS. IDA BEHRENDT-MØLLER (Orcid ID : 0000-0001-9716-720X)
DR. TERENCE KETTER (Orcid ID : 0000-0001-5132-453X)
DR. KEMING GAO (Orcid ID : 0000-0003-1130-2461)
DR. WILLIAM V BOBO (Orcid ID : 0000-0002-5674-3079)
DR. MAURICIO TOHEN (Orcid ID : 0000-0001-8049-4351)

Article type : Original Article

Patterns of changes in bipolar depressive symptoms revealed by trajectory analysis among 482 patients with bipolar disorder

By: Ida Behrendt-Møller[1*,] BSc; Trine Madsen[1], PhD; Holger Jelling Sørensen[1], MD, PhD; Louisa Sylvia[3,4], PhD; Edward S. Friedman[5], MD; Richard C. Shelton[6], MD; Charles L. Bowden[7]; MD; Joseph R. Calabrese[8], MD; Susan L. McElroy[9], MD; Terence A. Ketter[10], MD; Noreen A. Reilly Harrington[3,4], PhD; Keming Gao[8], MD; Michael Thase[11], MD; William V. Bobo[12], MD; Mauricio Tohen[13], MD, DrPH, MBA; Melvin McInnis[14], MD; Masoud Kamali[3], James H. Kocsis MD[15]; Thilo Deckersbach[3,4], PhD; Ole Köhler-Forsberg[1,2], MD; Andrew A. Nierenberg[3,4], MD

*** Corresponding author**: Ida Behrendt-Møller; Mental Health Centre Copenhagen, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark; Phone: +4520604181; Mail: ida.behrendt.moeller@gmail.com
§

Word count abstract: 249

Word count main text: 3152 words, 5 Tables & Figures, 39 references

[1] Mental Health Centre Copenhagen, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark
[2] Psychosis Research Unit, Aarhus University Hospital, Risskov, Denmark
[3] Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA
[4] Harvard Medical School, Boston, MA, USA
[5] Department of Psychiatry, University of Pittsburgh Medical Center, Pittsburgh, PA, USA
[6] Department of Psychiatry, University of Alabama at Birmingham, Birmingham, AL, USA
[7] Department of Psychiatry, University of Texas Health Science Center, San Antonio, TX, USA
[8] Department of Psychiatry, Case Western Reserve University, Cleveland, OH, USA
[9] Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati College of Medicine, Cincinnati, OH and Lindner Center of HOPE, Mason, OH, USA
[10] Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA
[11] Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA
[12] Department of Psychiatry & Psychology, Mayo Clinic, Rochester, MN, USA
[13] Department of Psychiatry, University of New Mexico Health Science Center, Albuquerque, NM, USA
[14] Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA
[15] Department of Psychiatry, Weill Cornell Medical College, NY, USA

Author Disclosure: Potential Conflicts of Interest

Ida Behrendt-Møller, Holger Jelling Sørensen, Ole Köhler-Forsberg and Trine Madsen have nothing to disclose.

Dr. Sylvia was a shareholder in Concordant Rater Systems and serves as a consultant for United Biosource Corporation and Clintara. She receives royalties from New Harbinger. Dr. Sylvia has also received grant funding from AFSP, PCORI and NIMH.

Dr. Bowden is conducting a biological study in mood disorders sponsored by Myriad Inc. He has no competing interests.

Dr. Gao has been on a speakers bureau and an advisory board for Sunovion and has also received grant supports from AstraZeneca, Brain and Behavior Research Foundation and Cleveland Foundation.

Dr. Bobo has received research support from NIMH, AHRA, and the Mayo Foundation for Medical Education and Research.

Dr. Trivedi has been an advisor/consultant and received fees from: Alkermes Inc., Allergan, AstraZeneca, Brintellix, BMS, Cerecor, Eli Lilly & Company, Forest, Health Research Associates, Johnson & Johnson, Lundbeck, Medscape, MSI Methylation Sciences Inc., Merck, Naurex Inc., Nestle Health Science – Pamlab Inc., One Carbon Therapeutics, Otsuka America Pharmaceuticals Inc., PamLab, Pfizer Inc., Roche, SHIRE Development, Takeda Pharmaceuticals Inc. In addition, he has received grants/research support from: Agency for Healthcare Research and Quality (AHRQ), National Institute of Mental Health (NIMH), Johnson and Johnson, National Institute of Diabetes and Digestive and Kidney diseases (NIDDK) and National Institute on Drug Abuse (NIDA).

Dr. Calabrese has received federal funding from the Department of Defense, Health Resources Services Administration and National Institute of Mental Health as well as grant support from: Abbott Laboratories; AstraZeneca; Bristol-Myers Squibb Company; Cephalon, Inc. (now Teva Pharmaceutical Industries Ltd.); Dainippon Sumitomo Pharma Co., Ltd.; GlaxoSmithKline; Janssen Pharmaceuticals, Inc.; Eli Lilly and Company; Intra-Cellular Therapies, Inc.; Pfizer, Inc; H. Lundbeck A/S; Sunovion Pharmaceuticals Inc.; Takeda Pharmaceutical Company Limited. Dr. Calabrese has served as a consultant/advisory board member/speaker for: Abbott Laboratories; Allergan; AstraZeneca; Bristol-Myers Squibb Company; Cephalon, Inc. (now Teva Pharmaceutical Industries Ltd.); Dainippon Sumitomo Pharma Co., Ltd.; GlaxoSmithKline; Janssen Pharmaceuticals, Inc.; H. Lundbeck A/S,;Merck & Co., Inc.; Otsuka Pharmaceutical Co., Ltd.; Pfizer, Inc; Repligen Corporation; Servier; Sunovion Pharmaceuticals Inc.; Solvay Pharmaceuticals, Inc.; Takeda Pharmaceutical Company Limited.

Dr. Thase has been an advisor/consultant: Alkermes; Allergan; AstraZeneca; Bristol-Myers Squibb Company; Cerecor, Inc.; Eli Lilly & Co.; Fabre-Kramer Pharmaceuticals, Inc.; Forest Laboratories; Gerson Lehrman Group; GlaxoSmithKline; Guidepoint Global; H. Lundbeck A/S; MedAvante, Inc.; Merck and Co. Inc. (formerly Schering Plough and Organon); Moksha8; Naurex, Inc.; Neuronetics, Inc.; Novartis; Ortho-McNeil Pharmaceuticals (Johnson & Johnson; Janssen); Otsuka; Pamlab, L.L.C. (Nestle); Pfizer (formerly Wyeth Ayerst Pharmaceuticals); Shire US Inc.; Sunovion Pharmaceuticals, Inc.; Takeda; and Trius Therapeutical, Inc. Dr. Thase receives grant funding from the Agency for Healthcare Research and Quality; Alkermes; AssureRx; Avanir; Forest Pharmaceuticals; Janssen; National Institute of Mental Health; Otsuka Pharmaceuticals. He has

equity holdings in MedAvante, Inc. and receives royalty income from American Psychiatric Foundation, Inc., Guilford Publications, Herald House, Oxford University Press, and W.W. Norton & Company. His wife is employed as the Group Scientific Director for Peloton Advantage which does business with Pfizer.

Dr. Shelton has been a consultant for: Allergan, Inc., Cerecor, Inc., Cyberonics, Inc., Janssen Pharmaceutica, Medtronic, Inc., Naurex, Inc., Nestle Health Science, Pfizer, Inc., Takeda Pharmaceuticals. Grant/Research Support: Allergan, Inc., Assurex, Inc., Avanir Pharmaceuticals, Cerecor, Inc., Genomind, Janssen Pharmaceutica , Naurex, Inc., Novartis Pharmaceuticals, Otsuka America, Nestle Health Science, Takeda Pharmaceuticals.

Dr. McInnis has received grants for research support from NIMH, the Heinz C Prechter Research Fund, and the Michigan Institute for Clinical Health Research (MICHR). MM has received consulting income from the Qatar National Research Foundation, Janssen, and Merck Pharmaceuticals.

Dr. Tohen was an employee of Lilly (1997 to 2008) and has received honoraria from or consulted for Abbott, AstraZeneca, Alkermes, Bristol Myers Squibb, GlaxoSmithKline, Lilly, Johnson & Johnson, Otsuka, Merck, Sunovion, Forest, Roche, Elan, Lundbeck, Teva, Pamlab, Minerva, Pfizer, Wyeth and Wiley Publishing; his spouse was a full time employee at Lilly (1998-2013).

Dr. Ketter has the following financial interests/arrangements or affiliations that could be perceived as real or apparent conflicts of interest: Grant/Research Support from the AstraZeneca Pharmaceuticals LP, Cephalon Inc., Eli Lilly and Company, Pfizer Inc., and Sunovion Pharmaceuticals; Consultant Fees from Allergan, Inc., Avanir Pharmaceuticals, Bristol-Myers Squibb Company, Cephalon Inc., Forest Pharmaceuticals, Janssen Pharmaceutica Products, LP, Merck & Co., Inc., Sunovion Pharmaceuticals, Teva Pharmaceuticals; Lecture Honoraria from Abbott Laboratories, Inc., AstraZeneca Pharmaceuticals LP, GlaxoSmithKline, and Otsuka Pharmaceuticals; and Publication Royalties from American Psychiatric Publishing, Inc. In addition, Dr. Ketter's spouse is an employee of and holds stock in Janssen Pharmaceuticals.

Dr. Friedman receives royalties from Springer.

Dr. Deckersbach's research has been funded by NIH, NIMH, NARSAD, TSA, IOCDF, Tufts University, DBDAT, Otsuka Pharmaceuticals and Cogito, Inc. He has received honoraria, consultation fees and/or royalties from the MGH Psychiatry Academy, BrainCells Inc., Clintara, LLC., Systems Research and Applications Corporation, Boston University, the Catalan Agency for Health Technology Assessment and Research, the National Association of Social Workers Massachusetts, the Massachusetts Medical Society, Tufts University, NIDA, NIMH, and Oxford University Press. He has also participated in research funded by DARPA, NIH, NIMH, NIA, AHRQ, PCORI, Janssen Pharmaceuticals, The Forest Research Institute, Shire Development Inc., Medtronic, Cyberonics, Northstar, Takeda, and Sunovion.

Dr. McElroy has been a consultant to or member of the scientific advisory boards of Allergen, Alkermes, Corcept, Ironshore, MedAvante, Mitsubishi Tanabe Pharma, Naurex, NovoNordisk, Shire, Sunovian, and Teva. She has been a principal or co-investigator on studies sonsored by the Agency for Healthcare Research & Quality (AHRQ), Alkermes, Allergen, AstraZeneca, Azevan, Cephalon, Eli Lilly and Company, Marriott Foundation, Myriad, National Institute of Mental Health, NovoNordisk, Orexigen Therapeutics, Inc., Shire, Sunovian,Takeda Pharmaceutical Company Ltd., and Transcept Pharmaceutical, Inc. She is also an inventor on United States Patent No. 6,323,236 B2, Use of Sulfamate Derivatives for Treating Impulse Control Disorders, and along with the patient's assignee, University of Cincinnati, Cincinnati, Ohio, has received payments from Johnson & Johnson, which has exclusive rights under the patent.

Dr. Reilly-Harrington receives royalties from Oxford University Press, the American Psychological Association, and New Harbinger. She serves as a consultant for United Biosource Corporation and was a shareholder in Concordant Rater Systems.

Dr. Nierenberg is a consultant for Abbott Laboratories, Alkermes, American Psychiatric Association, Appliance Computing Inc. (Mindsite), Basliea, Brain Cells, Inc., Brandeis University, Bristol Myers Squibb, Clintara, Corcept, Dey Pharmaceuticals, Dainippon Sumitomo (now Sunovion), Eli Lilly and Company, EpiQ, L.P./Mylan Inc., Forest, Genaissance, Genentech, GlaxoSmithKline, Healthcare Global Village, Hoffman LaRoche, Infomedic, Intra-Cellular Therapies, Lundbeck, Janssen Pharmaceutica, Jazz Pharmaceuticals, Medavante, Merck,

Methylation Sciences, NeuroRx, Naurex, Novartis, PamLabs, Parexel, Pfizer, PGx Health, Otsuka, Ridge Diagnostics Shire, Schering-Plough, Somerset, Sunovion, Takeda Pharmaceuticals, Targacept, and Teva; consulted through the MGH Clinical Trials Network and Institute (CTNI) for Astra Zeneca, Brain Cells, Inc, Dianippon Sumitomo/Sepracor, Johnson and Johnson, Labopharm, Merck, Methylation Science, Novartis, PGx Health, Shire, Schering-Plough, Targacept and Takeda/Lundbeck Pharmaceuticals. He is a stakeholder in Appliance Computing, Inc. (MindSite), Brain Cells, Inc., and Medavante. He receives research support from American Foundation for Suicide Prevention, AHRQ, Brain and Behavior Research Foundation, Bristol-Myers Squibb, Cederroth, Cephalon, Cyberonics, Elan, Eli Lilly, Forest, GlaxoSmithKline, Intra-Cellular Therapies, Janssen Pharmaceutica, Lichtwer Pharma, Marriott Foundation, Mylan, NIMH, PamLabs, PCORI, Pfizer Pharmaceuticals, Shire, Stanley Foundation, Takeda, and Wyeth-Ayerst. Honoraria include Belvoir Publishing, University of Texas Southwestern Dallas, Brandeis University, Bristol-Myers Squibb, Hillside Hospital, American Drug Utilization Review, American Society for Clinical Psychopharmacology, Baystate Medical Center, Columbia University, CRICO, Dartmouth Medical School, Health New England, Harold Grinspoon Charitable Foundation, IMEDEX, International Society for Bipolar Disorder, Israel Society for Biological Psychiatry, Johns Hopkins University, MJ Consulting, New York State, Medscape, MBL Publishing, MGH Psychiatry Academy, National Association of Continuing Education, Physicians Postgraduate Press, SUNY Buffalo, University of Wisconsin, University of Pisa, University of Michigan, University of Miami, University of Wisconsin at Madison, APSARD, ISBD, SciMed, Slack Publishing and Wolters Klower Publishing, ASCP, NCDEU, Rush Medical College, Yale University School of Medicine, NNDC, Nova Southeastern University, NAMI, Institute of Medicine, CME Institute, ISCTM, World Congress on Brain Behavior and Emotion, Congress of the Hellenic Society for Basic and Clinical Pharmacology, and ADAA. He has copyright joint ownership with MGH for Structured Clinical Interview for MADRS and Clinical Positive Affect Scale.

Abstract

Introduction

Depressive episodes are often prevalent among patients with bipolar disorder, but little is known regarding the differential patterns of development over time. We aimed to determine and

characterize trajectories of depressive symptoms among adults with bipolar disorder during 6 months of systematic treatment.

Method

The pragmatic clinical trial, *Bipolar Clinical Health Outcomes Initiative in Comparative Effectiveness* (CHOICE) randomized 482 outpatients with bipolar disorder to lithium or quetiapine. Depressive symptoms were rated at up to 9 visits using the *Montgomery-Asberg Depression Rating Scale* (MADRS). *Growth Mixture Modelling* was utilized to identify trajectories and multinomial regression analysis estimated associations with potential predictors.

Results

Four distinct trajectories of depressive symptoms were identified. The *Responding* class (60.3 %) with a rapid reduction and subsequent low level; the *Partial-responding* class (18.4%) with an initial reduction followed by an increase during the remaining weeks; the *Fluctuating* class (11.6 %) with a fluctuation in depressive symptoms; the *Non-responding* class (9.7%) with sustained moderate-severe depressive symptoms. Bipolar type I predicted membership of the *Non-responding* class and randomization to quetiapine predicted membership of either the *Responding* or the *Non-responding* class.

Conclusion

Approximately 30% experienced a partial or fluctuating course and almost 10% had a chronic course with moderate-severe depression during 6 months. Patients diagnosed with bipolar type 1 had higher risk of being categorized into a class with a worse outcome. While no differences in average overall outcomes occurred between the lithium and quetiapine groups, trajectory analysis revealed that the lithium group had more variable courses.

Keywords

Bipolar disorder, depressive symptoms, trajectories, Growth Mixture Modelling

Introduction

Bipolar disorder has a negative impact on the patient's quality of life, with depressive symptoms having a particularly strong effect on their well-being. The unpredictable nature of the illness makes

it challenging to treat efficiently and despite state-of-the-art treatment, many patients do not respond sufficiently and have symptoms approximately 50% of the time[1]. Furthermore, the medical treatment with mood stabilizers and antipsychotics is associated with severe side-effects and antidepressants can cause manic switches [1,2]. Only little is known regarding symptom development in patients with bipolar disorder and it is therefore highly relevant to investigate this aspect to gain more knowledge of the illness and whether similarities exist between this patient group and patients with unipolar depression. Furthermore, the medical treatment of bipolar depression differs from the treatment of unipolar depression and it is therefore important to investigate treatment the response of the medical treatment used in bipolar depression.

Longitudinal follow-up studies have found that bipolar disorder is chronic and dominated by depressive episodes rather than manic or hypomanic episodes [1,3,4]. Therefore, better treatment of depressive symptoms represents one of the main challenges in the treatment of patients with bipolar disorder [5,6]. Further research is of great clinical importance and could potentially contribute to a more personalized and improved treatment.

Recently, group based trajectory models have gained much traction because of their usefulness when studying heterogeneity in symptom development. In these models, underlying subgroups within a population are identified and a growth curve for each subgroup is estimated [7,8]. Several studies have examined trajectories of depressive symptoms and found symptom development to be highly heterogeneous [9–19]. However, these studies were either conducted among the general population [9–12,14,15] or among patients suffering from unipolar depressive disorder [16–19]. Only few clinical studies have performed trajectory analyses to explore symptom development among patients with bipolar disorder and none of these have investigated how depressive symptoms improve over time [20–22]. Hence, the primary purpose of this study was to estimate trajectories of depressive symptoms among outpatients with diagnosed bipolar disorder to examine, if patients could be classified into subgroups, where they shared similar patterns of depressive symptom improvement. Our secondary aim was to investigate if specific covariates predicted membership of the identified trajectory classes.

Method

Data source

Data was obtained from the *Bipolar Clinical Health Outcomes Initiative in Comparative Effectiveness* (CHOICE) study [23]. Bipolar CHOICE was a pragmatic six-month randomized controlled multisite trial comparing treatment with lithium to treatment with quetiapine among outpatients diagnosed with bipolar disorder type I (68.3%) or type II (31.7%) according to DSM-IV-TR criteria [24]. Participants had to be 18 years or older and at least mildly symptomatic at inclusion time with a Clinical Global Impression scale for bipolar disorder (CGI-BP) score ≥ 3 [25]. Participants were excluded if they had any contraindication to lithium or quetiapine, were in a crisis (e.g. inpatient hospitalization), or if they were currently treated with lithium or quetiapine. 692 outpatients with bipolar disorder were screened and 482 patients met the inclusion criteria and were randomized to either lithium (240 patients) or quetiapine (242 patients) along with adjunctive personalized treatment (APT) (except lithium and quetiapine). Demographics and clinical features were monitored among participants at baseline and found to be similar between the two randomized groups [26].

The study took place at 11 sites in the United States and was conducted from September 2010 to September 2013. At study entry, trained clinical research coordinators collected sociodemographic and clinical information and the symptomatology was closely monitored with several rating scales both at baseline and during eight follow-up visits at week: 2, 4, 6, 8, 12, 16, 20, and 24. The study was approved by the Institutional Review Boards at all sites and all participants signed approved informed consent forms prior to initiation of the trial. Further details about the study design is described elsewhere [23].

Outcome measure

Depressive symptoms were assessed using the *Montgomery Asberg Depression Rating Scale* (MADRS), a well-established ten item rating scale measuring the overall severity of depressive symptoms ranging from 0 to 60 [27]. We chose this scale due it previously having shown a high reliability when measuring depressive symptoms among patients with unipolar depression[28]. The following cut-off values where used to grade depressive symptoms severity: 0-6: Symptoms absent; 7-19: Mild depression; 20-34: Moderate depression; $\geq$35: Severe depression[29]. Criteria for treatment remission was defined as a MADRS score $\leq$ 12 at the last follow up visit and treatment response was defined as a MADRS score reduction $\geq$ 50% measured from baseline to the last follow up visit.

Predictors

We included the following socio-demographic and clinical covariates assessed at study entry in the predictor analysis: sex, age, age at first manic episode (if appropriate), age at first depressive episode, ethnicity (Caucasian/other), education (high school or less/college or more), employment (employed or student/not employed), type of bipolar disorder (type 1/type 2), CGI mania score at baseline, history of psychiatric hospitalization, history of suicide attempt, any current anxiety disorder, current diagnosis of post-traumatic stress disorder (PTSD), level of psychotic symptoms at baseline, family history of mental illness (schizophrenia, bipolar disorder or unipolar depression among first line relatives) comorbidities (diabetes, hypertension, hyperlipidaemia), history of suicide attempts in the family (parents, siblings or children), treatment arm (lithium/quetiapine), alcohol abuse within the last 12 months, lifetime alcohol abuse, any substance abuse within the last 12 months, any lifetime substance abuse.

Statistical analysis

All models were estimated in Mplus editor version 7.2.

We used *Growth Mixture Modelling* (GMM) to estimate trajectories of depressive symptoms [7,8]. This is a data driven person-centred approach, where subgroups within the population are identified based on prototypical patterns in slope and intercepts [8]. In that way a growth curve for each subgroup is estimated and variation within each subgroup is allowed [7,8,30].

We first estimated *Latent Class Analysis* (LCA) and *Latent Growth Curve Models* (LGCM) to check if it appeared that subgroups within the population actually existed. We then estimated *Longitudinal Class Growth Analysis* (LCGA) and GMM models with different growth functions (i.e., linear, quadratic, or cubic) and with increasing number of classes. All models with an entropy below 0.7 and models that could not converge were excluded hence doubting the validity of the estimated models.

When comparing the different models, we examined fit-estimates (*Baysian Information Criterion* (BIC); adjusted BIC and *Akaike Information Criterion* (AIC)) where lower fit estimates indicated a better fit of data. Furthermore, we looked at class sizes, entropy, posterior probabilities for classification accuracy, clinical utility and the bootstrap log-likelihood test which tests whether the model with n number of classes is a significantly better fit of data compared to a model with n-1 number of classes [7].

After deciding on the final model we included all covariates to test their association with the identified trajectories. We used an explorative approach when performing the predictor analysis. We applied a three-step approach which takes into account classification error [31]. When using this approach, the pre-defined covariates did not influence the formation of the trajectories, but their association with the identified classes could be tested after trajectory class identification [32]. We first performed a univariable analysis studying the association between the identified trajectories and each co-variate separately. Afterwards, we included all covariates with significant interclass differences (p-value<0.05) in a multinomial logistic regression analyses. Missing data was handled by using the *Full Information Maximum Likelihood* (FIML) approach[33]. To test for multiple testing, we also performed bonferroni type adjustment on our results. The results were presented as odds ratios (OR) with 95% confidence intervals (95%-CI) and corresponding p-values. Further details regarding the statistical analysis are presented in the supplementary material.

Results

The LGCM model showed poor fit estimates and significant variance in growth factors which suggested the appearance of multiple classes. The cubic GMM model showed the lowest fit estimates in all class-models compared to the quadratic and linear models and compared to all the LCGA models. Furthermore, the mean value of the cubic term was significant in at least one of the classes in all the cubic-term-models. We therefore decided on a cubic model with the variance for the cubic term fixed to zero.

Goodness of fit statistics for the cubic model with one to five classes are presented in table 1. Fit estimates decreased at progressing number of classes and the p-value for the bootstrap log-likelihood test stayed significant.

We decided on the four class model based on a high drop in fit estimate values from the three to the four class model and a significant p-value in the bootstrap likelihood ratio test. Furthermore, the four class model revealed important clinical information concerning two classes with a fluctuating course of depressive symptoms which were not shown in the two or three class models. We did not choose the five class model, since the extra class did not contribute with additional clinical important information and one of the classes was rather small (5.68%).

The four identified trajectories are presented in figure 1 and the observed individual values within each of the identified classes are presented in figure 2A. The *Responding* class (60.3%) was characterized by a mean decrease in depressive symptoms during the first eight weeks and then stabilized at a low level throughout the rest of the study period (figure 2a). This group had a mean 78% reduction in depressive symptoms and ended up with an average MADRS score of 4.5 at the end of follow-up. We found that 86% of the patients categorized into this group had a treatment response (defined as a MADRS reduction≥ 50%) and 95.9% experienced remission (defined as a MADRS score≤ 12) at the end of the follow up period (figure 2b). The *Partial-responding* class (18.4%) was characterized by an average rapid reduction in depressive symptoms during the first two weeks followed by a slower decrease throughout 2-16 weeks and an increase between the 16-24 weeks. The mean MADRS level remained within the level of mild depression at the end of the follow-up (figure 2a). The *Partial-responding* class had in average a depressive symptom reduction of 33% from baseline to the end of the follow-up period with 27.5% having experienced treatment response and 15.9% experienced remission (figure 2b). The *Fluctuating* class (11.6 %) was characterized by a fluctuating course of depressive symptoms. The trajectory initial declined, but rebounded to an average level of moderate depression after 12 weeks, followed by a drop to a level of mild depressive symptoms during the 16-24 weeks. At the end of the follow up, the group experienced a mean 58% decrease in depressive symptoms compared to baseline. 66.7% of the patients had a treatment response and 52.5% experienced remission (figure 2B). The *Non-responding* class (9.7%) had a mean MADRS score of 29.4 at baseline which increasing even further to 31.4 at week 24 (i.e., a 7% increase). None of the patients in this class experienced remission or response during the study period.

Predictors of class membership

Table 2 presents results from the univariable predictor analysis and table 3 presents the results from the multivariable predictor analysis, both using the *Responding* class as the reference. Results from the univariable analysis and multivariable analysis using the three other classes as references are represented in supplementary tables 1 to 3 and supplementary tables 4 to 6 respectively. Race, employment, randomization, current diagnosis of anxiety, history of suicide in family, PTSD, psychotic symptoms at baseline and type of bipolar disorder showed significantly inter class differences in the univariable predictor analysis (all $p<0.05$) and these variables were all included in the multivariable analysis. Here we found that patients diagnosed with bipolar disorder type I (versus type II) had more than threefold higher odds of membership in the *non-responding* class

compared to the *Responding* class (p=0.04); People randomized to quetiapine treatment, compared to patients randomized to lithium, had lower odds of being in the *Fluctuating* (p=0.006) and *Partial-responding* (p=0.023) class compared to the *Responding* class and lower odds of being in the *Fluctuating* class compared to the *Non-responding* class (p=0.016); Finally, the presence of psychotic symptoms was significantly associated with lower odds of being in the *Responding* class compared to the *Partial responding* class (p=0.046). However, when using the bonferroni corrected p-value the new critical p-value was found by dividing the original α-value (0.05) with number of variables included in the multivariable analysis (8). This means that the p value had to be $p<0.00625$ to achieve significance. Therefore, only randomizing to quetiapine treatment was still significantly associated with lower odds of being in the *Fluctuating* class compared to the *Responding* class when using the bonferroni corrected p-value.

Discussion

This study is the first to explore the differential trajectories of depressive symptoms among outpatients with bipolar disorder during 6 months of pharmacotherapy. We used a data driven person-centred approach and identified a four class GMM model, which is consistent with what previous trajectory studies have found in studies of patients with unipolar depression [16,18,19,34]. We found that 60.3% of the patients were classified into the *responding* class where 96% experienced remission and the average MADRS score was 4.5 at the end of the follow-up. Conversely, 9.7% of the patients were classified into the *non-responding* class with moderate to severe depression throughout the entire 6 months despite medical treatment. No patients in this group responded to treatment or had remission. A previous study among patients with an acute episode of bipolar disorder used overall mood symptoms to estimate trajectories and found that 10.2% of the population belonged to a group with persistent depressive symptoms throughout the 4 weeks study duration[21]. However, the present study is the first to indicate that approximately one in ten outpatients with bipolar disorder have a persistent high level of depressive symptoms during 6 months despite state-of-the-art pharmacotherapy.

Regarding potential predictors, we found that a higher level of psychotic symptoms at baseline was significantly associated with lower odds of being in the *Responding* class and that a diagnose with bipolar type I predicted membership of the *Non-responding* class compared to the *Responding* class. Finally, we found that subjects randomized to quetiapine had higher odds of being either in the

Responding class or the *Non-responding* class compared to the *Fluctuating* class and higher odds of being in the *Non-responding* class compared to the *Partial-responding* class. However, after correction with bonferroni only the randomization variable showed significant interclass differences between the *Responding* class and the *Fluctuating* class. These findings may help clinicians to identify those patients with additional need for help against persistent depressive symptoms. The differences between bipolar type I and type II support prior studies indicating that type I represents a more severe illness course, although other studies have found that the burden of depressive symptoms were similar in outpatients with bipolar type I and type II [1,5,35]. The finding that a higher level of psychotic symptoms at baseline was associated with membership of a trajectory with a worse outcome correspond well with the DSM-IV criteria that assigns psychotic symptoms as a marker for illness severity of bipolar disorder. However, the significance of psychotic symptoms in relation to clinical outcome in bipolar disorder is not yet fully understood. Although some studies have found the presence of psychotic symptoms to be associated with a higher burden of disease and higher morbidity among patients with bipolar disorder type 2[36] other studies did not find any difference in clinical or functional outcome between patients with and without psychotic symptoms among patients with bipolar disorder[37,38]. Thus, future large studies should elucidate the influence of psychotic symptoms on illness development among patients with bipolar disorder.

Finally, the found difference between the two treatment arms was surprising since the original bipolar CHOICE trial showed that treatment with lithium and quetiapine were not significantly different overall and the two groups had a similar treatment effect of their depression during the 6 months[23]. Future studies should in more detail and a priori hypotheses investigate whether there may exist differences in treatment outcome between quetiapine and lithium in specific subgroups and if subgroups exist that respond better to quetiapine compared to lithium including potential predictors. Such a finding would represent clinical important information and could eventually help with more personalized medicine in this patient group.

Strengths and Limitations

The study has several strengths including a large study population and a high inter-rater reliability[23]. Additionally, the frequent assessments of up to 9 visits during 6 months follow-up allowed us to apply a detailed model to identify trajectory groups.

Furthermore, the pragmatic study design with broad inclusion- and few exclusion criteria results in a high degree of generalizability to outpatients with bipolar disorder seen in everyday clinical practice [23].

It is important to evaluate the study within its limitations. Selection-bias may be present since only outpatients who sought treatment at academic medical centres were included. Furthermore, hospitalization of the patient was an exclusion criterion which could potentially exclude those with very severe psychopathology. Moreover, the study only had 6 months' follow-up, hence, future trials should explore the extent to which persistent depressive symptoms tend to remit in the longer time perspective and the influence of the predictors of the course in the longer run. In this study we did not know for how long the depressive symptoms had been present prior to inclusion time. This is an important limitation since depressive symptoms can remit spontaneously on their own over time. Another limitation concerns the method of GMM. Our choice of the model was based on careful statistical and clinical considerations. However, the choice of a model in GMM analysis is also a subjective assessment and it is possible that others would have found another model to be the best fit of the data. Finally, despite the large study population, some predictors may not have achieved statistical significance due to small trajectory classes[39]. This could also be part of the reason that many of the found significant predictors were no longer significant after the bonferroni correction.

Conclusion

Among 482 adult outpatients with bipolar disorder treated pharmacologically for 6 months, we identified four distinct mood trajectory classes. Almost one in ten had persistent high depressive symptoms with moderate-severe depression despite 6 months of pharmacotherapy. In the predictor analysis, it was found that psychotic symptoms at baseline and bipolar type I predicted membership of trajectories with a worse course of depressive symptoms. Finally, the trajectory analyses also revealed differences between the two randomized treatment groups which were not detected by conventional statistical methods.

Future studies with an equal size or larger group of outpatients and a longer follow-up period is needed to explore the heterogeneous course of depressive symptoms in patients with bipolar disorder in more detail. Also it is needed to investigate whether patients from the *Non-responding* class may respond better to additional or other treatment regimens.

Results from this and future studies could be essential to obtain a better and more personalized treatment of patients, as previous studies show that depressive symptoms remain the greatest challenge in the treatment of patients with bipolar disorder.

References

1. Kupka RW, Altshuler LL, Nolen WA, et al. Three times more days depressed than manic or hypomanic in both bipolar I and bipolar II disorder. *Bipolar Disord*. 2007;9(5):531-535. doi:10.1111/j.1399-5618.2007.00467.x.
2. Serretti A, Chiesa A, Calati R, et al. Side effects associated with psychotropic medications in patients with bipolar disorder: evidence from two independent samples. *J Psychopharmacol*. 2013;27(7):616-628. doi:10.1177/0269881113485143.
3. Judd LL, Akiskal HS, Schettler PJ, et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Arch Gen Psychiatry*. 2002;59(6):530-537. doi:10.1001/archpsyc.59.6.530.
4. Judd LL, Akiskal HS, Schettler PJ, et al. A prospective investigation of the natural history of the long-term weekly symptomatic status of bipolar ii disorder. *Arch Gen Psychiatry*. 2003;60:261-269. doi:10.1001/archpsyc.60.3.261.
5. Yatham LN, Kennedy SH, Bond DJ, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) 2018 guidelines for the management of patients with bipolar disorder. *Bipolar Disord*. 2018;(20):97-170. doi:10.1111/bdi.12609.
6. Joffe RT, MacQueen GM, Marriott M, Young LT. One-year outcome with antidepressant--treatment of bipolar depression. *Acta Psychiatr Scand*. 2005;112(2):105-109. doi:10.1111/j.1600-0447.2005.00583.x.
7. Jung T, Wickrama A. An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling. *Soc Personal Psychol Compass*. 2008;2(1):302-317.
8. Muthén BO, Muthén LK. Integrating person centered and variable centered Analyses: Growth mixture modeling with latent trajectory classes. *Alcohol Clin Exp Res*. 2000;24(6):882-891. doi:10.1111/j.1530-0277.2000.tb02070.x.

9. Stoolmiller M, Kim HK CD. The Course of Depressive Symptoms in Men from Early Adolescence to Young Adulthood: Identifying Latent Trajectories and Early Predictors. *J Abnorm Psychol*. 2005;114(3):331–345. doi:10.1037/0021-843X.114.3.331.
10. Lincoln KD, Takeuchi DT. Variation in the Trajectories of Depressive Symptoms: Results from the Americans' Changing Lives Study. *Biodemography Soc Biol*. 2010;56(1):24-41. doi:10.1080/19485561003709180.Variation.
11. Smolenski, D. J., Pruitt, L. D., Vuletic, S., Luxton, D. D., & Gahm G (2017, Pruitt LD, Vuletic S, et al. Unobserved Heterogeneity in Response to Treatment for Depression Through Videoconference. *Psychiatr Rehabil Journal*. 2017. http://dx.doi.org/10.1037/prj0000273.
12. Colman I, Ploubidis GB, Wadsworth MEJ, Jones PB, Croudace TJ. A Longitudinal Typology of Symptoms of Depression and Anxiety Over the Life Course. *Biol Psychiatry*. 2007;62(11):1265-1271. doi:10.1016/j.biopsych.2007.05.012.
13. Musliner KL, Munk-Olsen T, Eaton WW, Zandi PP. Heterogeneity in long-term trajectories of depressive symptoms: Patterns, predictors and outcomes. *J Affect Disord*. 2016;192:199-211. doi:10.1016/j.jad.2015.12.030.
14. Costello D, Swendsen J, Rose J, Dierker L. Risk and Protective Factors Associated with Trajectories of Depressed Mood from Adolescence to Early Adulthood. *J Clin P*. 2008;76(2):173-183. doi:10.1037/0022-006X.76.2.173.Risk.
15. Holmes SE, Esterlis I, Mazure CM, et al. Trajectories of depressive and anxiety symptoms in older adults: a 6-year prospective cohort study. *Int J Geriatr Psychiatry*. 2017;(March). doi:10.1002/gps.4761.
16. Wardenaar KJ, Conradi H-J, de Jonge P. Data-Driven Course Trajectories in Primary Care Patients With Major Depressive Disorder. *Depress Anxiety*. 2014;31(9):778-786. doi:10.1002/da.22228.
17. Rhebergen D, Lamers F, Spijker J, de Graaf R, Beekman ATF, Penninx BWJH. Course trajectories of unipolar depressive disorders identified by latent class growth analysis. *Psychol Med*. 2012;42(7):1383-1396. doi:10.1017/S0033291711002509.
18. Wardenaar KJ, Monden R, Conradi HJ, De Jonge P. Symptom-specific course trajectories and their determinants in primary care patients with Major Depressive Disorder: Evidence for two etiologically distinct prototypes. *J Affect Disord*. 2015;179:38-46. doi:10.1016/j.jad.2015.03.029.

19. Thibodeau MA, Quilty LC, De Fruyt F, De Bolle M, Rouillon F, Bagby RM. Latent classes of nonresponders, rapid responders, and gradual responders in depressed outpatients receiving antidepressant medication and psychotherapy. *Depress Anxiety*. 2015;32(3):213-220. doi:10.1002/da.22293.
20. Birmaher B, Gill MK, Axelson DA, et al. Longitudinal trajectories and associated baseline predictors in youths with bipolar spectrum disorders. *Am J Psychiatry*. 2014;171(9):990-999. doi:10.1176/appi.ajp.2014.13121577.
21. M'Bailara K, Cosnefroy O, Vieta E, Scott J, Henry C. Group-based trajectory modeling: A novel approach to examining symptom trajectories in acute bipolar episodes. *J Affect Disord*. 2013;145(1):36-41. doi:10.1016/j.jad.2012.07.007.
22. Köhler-Forsberg O, Madsen T, Behrendt-Møller I, et al. Trajectories of suicidal ideation over 6 months among 482 outpatients with bipolar disorder. *J Affect Disord*. 2017;223. doi:10.1016/j.jad.2017.07.038.
23. Nierenberg AA, McElroy SL, Friedman ES, et al. Bipolar CHOICE (clinical health outcomes initiative in comparative effectiveness): A pragmatic 6-month trial of lithium versus quetiapine for Bipolar disorder. *J Clin Psychiatry*. 2016;77(1):90-99. doi:10.4088/JCP.14m09349.
24. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, (DSM-IV) (4th Edn)*. 4th ed. (American Psychiatric Association Press, Inc., Washing- ton, DC 1994., ed.).; 1994.
25. Spearing MK, Post RM, Leverich GS, Brandt D NW. Modification of the Clinical Global Impression (CGI) Scale for use in bipolar illness: the CGI-BP. *Psychiatry Res*. 1997;73(3):159-171.
26. Nierenberg AA, Sylvia LG, Leon AC, et al. Clinical and Health Outcomes Initiative in Comparative Effectiveness for Bipolar Disorder (Bipolar CHOICE): a pragmatic trial of complex treatment for a complex disorder. *Clin Trials*. 2014;11(1):114-127. doi:10.1177/1740774513512184.
27. Montgomery, S., & Åsberg M. A New Depression Scale Designed to be Sensitive to Change. *Brit J Psychiat*. 1979;134(4):382-389. doi:10.1192/bjp.134.4.382.
28. Carneiro AM, Fernandes F, Moreno RA. Hamilton depression rating scale and montgomery-asberg depression rating scale in depressed and bipolar I patients: Psychometric properties in a Brazilian sample. *Health Qual Life Outcomes*. 2015;13(1):1-8. doi:10.1186/s12955-015-

0235-3.

29. Snaith, R., Harrop, F., Newby, D., & Teale C. Grade scores of the Montgomery-Asberg Depression and the Clinical Anxiety Scales. *Br J Psychiatry*. 1986;148(5):599-601. doi:10.1192/bjp.148.5.599.
30. Muthen BO. The potential of growth mixture modeling. *Infant Child Dev*. 2006;3(December 2007):3-5. doi:10.1002/icd.
31. Asparouhov T, Muthén B. Auxiliary Variables in Mixture Modeling: Three-Step Approaches Using M plus. *Struct Equ Model A Multidiscip J*. 2014. doi:10.1080/10705511.2014.915181.
32. Asparouhov T, Muthen B. Auxiliary Variables in Mixture Modeling : 3-Step Approaches Using Mplus. *Mplus Web Notes No 15*. 2013;(15):1-48. http://www.statmodel.com/examples/webnotes/webnote15.pdf.
33. Muthén LK, Muthén BO. Mplus Statistical Analysis With Latent Variables User ' S Guide. 2017;Version 8.
34. Tada M, Uchida H, Mizushima J, Suzuki T, Mimura M, Nio S. Antidepressant dose and treatment response in bipolar depression: Reanalysis of the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) data. *J Psychiatr Res*. 2015;68:151-156. doi:10.1016/j.jpsychires.2015.06.015.
35. Datto C, Pottorf WJ, Feeley L, LaPorte S, Liss C. Bipolar II compared with bipolar I disorder: baseline characteristics and treatment response to quetiapine in a pooled analysis of five placebo-controlled clinical trials of acute bipolar depression. *Ann Gen Psychiatry*. 2016;15(1):9. doi:10.1186/s12991-016-0096-0.
36. Mazzarini L, Colom F, Pacchiarotti I, et al. Psychotic versus non-psychotic bipolar II disorder. *J Affect Disord*. 2010;126(1-2):55-60. doi:10.1016/j.jad.2010.03.028.
37. Jr PEK, Mcelroy SL, Havens JR, et al. Psychosis in Bipolar Disorder: Phenomenology and Impact on Morbidity and Course of Illness. *Compr Psychiatry*. 2003;44(4):263-269. doi:10.1016/S0010-440X(03)00089-0.
38. Burton CZ, Ryan KA, Kamali M, et al. Psychosis in bipolar disorder : Does it represent a more " severe " illness ? *Bipolar Disord*. 2018;20:18-26. doi:10.1111/bdi.12527.
39. Muthén LK, Muthén BO. How to Use a Monte Carlo Study to Decide on Sample Size and Determine How to Use a Monte Carlo Study to Decide on Sample Size and Determine Power. *Struct Equ Model A Multidiscip J*. 2009;9(4):599-620. doi:10.1207/S15328007SEM0904.

Table 1 Goodness of fit statistic for one to five class solutions for the Cubic GMM analysis.

No. of classes	Fit estimates [a]			P-values [b]	Classification accuracy		
	AIC [c]	*BIC* [d]	*adjBIC* [e]	*Bootstrap* [f]	*Entropy* [g]	*Class accuracy* [h]	*Class size (%)* [i]
1 class	25359	25438	25378	-	-	-	-
2 class	25254	25354	25278	< 0.001	0.805	0.96 0.89	79.83 20.17
3 class	25224	25345	25253	< 0.001	0.802	0.94 0.75 0.85	74.47 6.54 18.98
4 class	25181	25324	25216	< 0.001	0.779	0.92 0.91 0.79 0.80	9.74 60.29 18.36 11.60
5 class	25152	25315	25191	< 0.001	0.791	0.87 0.91 0.78 0.90 0.81	12.59 58.45 13.41 9.88 5.68

[a] *Statistical information criteria. A lower value indicates better fit.*
[b] *A significant P-value indicates that the model with the number of classes is better fit of the data compared to the model with the k-1 number of classes*
[c] *Akaike Information Criteria.*
[d] *Bayesian Information Criteria.*
[e] *Sample size adjusted BIC.*
[f] *Bootstrap Likelihood Ratio Test.*
[g] *Entropy is estimated based on the average posterior probability and ranges from 0 to 1, where higher estimates represent greater classification accuracy for the model.*
[h] *Expresses the average accuracy of membership classification in each class. As the entropy measure it ranges from 0 to 1, where higher estimates represent greater classification accuracy.*
[i] *Distribution of the total sample into identified classes based on the posterior probability.*

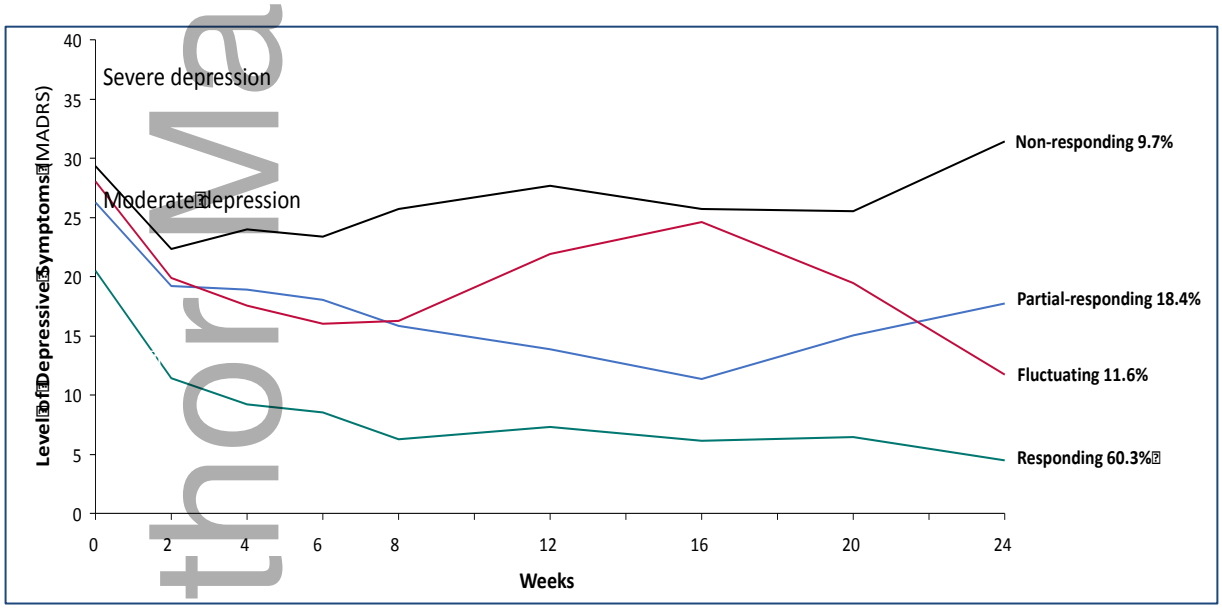
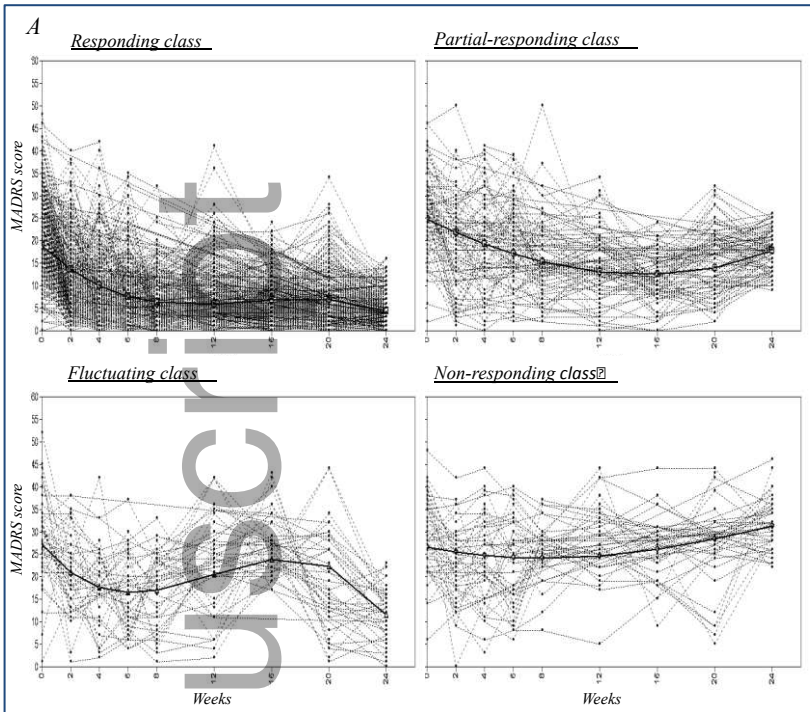


Figure 1 Trajectories of depressive symptoms measured by the MADRS score during 6 months of mood-stabilizing treatment among 482 outpatients with bipolar disorder.



B

Classes [a]	Treatment response [b] n (%)	Remission [c] n (%)
Partial-responding	22 (27.5)	13 (15.9)
Fluctuating	26 (66.7)	21 (52.5)
Responding	227 (86.0)	255 (95.9)
Non-responding	0 (0)	0 (0)

[a] Classes are based on the most likely class membership
[b] Treatment response is defined as a 50% or higher decrease of the MADRS score. Number of missing values in each classes (%): Partial-responding 6 (7.0); Fluctuating 5 (11.4); Responding 90 (29.2); Non-responding 4 (9.1);
[c] Remission is defined as a MADRS score lower or equal to 12. Number of missing values in each class: Partial-responding 4 (4.7); Fluctuating 4 (9.1); Responding 42 (13.6); Non-responding 4 (9.1)

Figure 2 Presentation of the four identified trajectory classes

A) Estimated mean trajectories of depressive symptoms measured by the MADRS score (bold lines) and the observed individual values within each of the identified classes. The trajectory classes are based *on "most likely class membership".*

B) Treatment response and remission in the different trajectory classes. All results are presented as valid percent, meaning percentage of valid (non missing) observations.

Table 2 Predictors for membership of the identified trajectory classes based on the univariable regression analysis. The Responding class is the reference. Significant findings are bold.

	Partial-responding class		Fluctuating class		Non-responding class	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Age at baseline	1.01 (0.99-1.04)	0.320	1 (0.97-1.04)	0.833	1.01 (0.98-1.04)	0.444
Gender						
-Female	1.05 (0.55-2)	0.878	2.25 (0.87-5.8)	0.093	1.57 (0.76-3.26)	0.221
-Male	1		1		1	
Race						
-All others	1.19 (0.58-2.46)	0.641	**2.43 (1.05-5.64)**	**0.038**	1.11 (0.5-2.44)	0.803
-Caucasian	1		1		1	
Education						
-College or more.	0.68 (0.33-1.37)	0.277	0.77 (0.3-1.97)	0.589	0.61 (0.29-1.3)	0.201
-High school or less	1		1		1	
-Employment						
-Not employed	**2.08 (1.06-4.08)**	**0.034**	1.73 (0.74-4.03)	0.203	1.67 (0.82-3.39)	0.155
-Employed/student	1		1		1	
Randomizarion						
-Quetiapine	**0.46 (0.24-0.89)**	**0.020**	**0.25 (0.09-0.66)**	**0.005**	1.06 (0.52-2.15)	0.871
-Lithium	1		1		1	
Previous psychiatric hospitalization						
-Yes	1.49 (0.79-2.8)	0.217	1.15 (0.51-2.57)	0.743	1.55 (0.78-3.07)	0.208
-No	1		1		1	
Psychotic symptoms	**1.27 (1.05-1.54)**	**0.013**	1.27 (0.99-1.61)	0.054	**1.23 (1.02-1.46)**	**0.026**
Anxiety						
-Yes	1.64 (0.84-3.17)	0.145	1.69 (0.72-3.97)	0.231	**2.59 (1.19-5.65)**	**0.016**
-No	1		1		1	
Psychiatris illness in family						
-Yes	1.41 (0.65-3.05)	0.380	2.27 (0.68-7.5)	0.181	1.43 (0.62-3.32)	0.401
-No	1		1		1	
Suicide in family						
-Yes	0.79 (0.22-2.81)	0.713	**2.75 (1-7.59)**	**0.050**	1.14 (0.37-3.52)	0.817
-No	1		1		1	
Previous suicide attempts						
-Yes	1.23 (0.64-2.37)	0.531	1.41 (0.61-3.23)	0.420	1.85 (0.92-3.71)	0.083
-No	1		1		1	
PTSD current.						
-Yes	1.47 (0.53-4.08)	0.463	2.06 (0.65-6.52)	0.221	**3.62 (1.53-8.6)**	**0.004**
-No	1		1		1	
Comorbidities						
-Yes	1.61 (0.84-3.09)	0.155	0.99 (0.41-2.43)	0.989	0.91 (0.42-1.95)	0.802
-No	1		1		1	
Alcohol abuse within the last 12 months						
-Yes	0.38 (0.05-2.66)	0.332	1.18 (0.3-4.68)	0.810	1.85 (0.67-5.11)	0.235
-No	1		1		1	
Alcohol abuse lifetime						
-Yes	0.97 (0.48-1.93)	0.923	1.4 (0.6-3.24)	0.437	1.83 (0.91-3.68)	0.092
-No	1		1		1	
Any substance abuse within the last 12 months						
-Yes	1.31 (0.52-3.29)	0.566	1.76 (0.61-5.11)	0.300	1.67 (0.66-4.19)	0.277
-No	1		1		1	
Any substance abuse lifetime						
-Yes	1.44 (0.76-2.73)	0.266	1.34 (0.58-3.05)	0.492	1.78 (0.89-3.58)	0.105
-No	1		1		1	
Type of -Bipolar disorder						
-Bipolar type I	1.54 (0.75-3.14)	0.238	1.16 (0.49-2.77)	0.732	**4.04 (1.4-11.65)**	**0.010**
-Bipolar type II	1		1		1	
CGI-mani at baseline	1.19 (0.9-1.56)	0.218	1.07 (0.74-1.54)	0.728	1.28 (0.99-1.65)	0.061
Age at first manic episode	1.02 (0.98-1.05)	0.273	1.02 (0.96-1.07)	0.496	1 (0.96-1.03)	0.836
Age at first depressive episode	1 (0.95-1.05)	0.920	0.99 (0.94-1.05)	0.780	0.99 (0.95-1.03)	0.492

Table 3 Predictors for membership of the different classes based on the multivariable regression analysis. The responding class is the reference. Significant findings are bold.

	Partial-responding class 1		**Fluctuating class (11.6)**		**Non-responding class (9.7**	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Race						
-All others	0.97 (0.43-2.18)	0.942	2.34 (0.89-6.17)	0.085	0.9 (0.4-2.04)	0.801
-Caucasian	1		1		1	
-Employment						
-Not employed	1.72 (0.83-3.53)	0.142	1.4 (0.52-3.72)	0.502	1.32 (0.61-2.82)	0.481
-Employed/student	1		1		1	
Randomizarion						
-Quetiapine	**0.44 (0.22-0.89)**	**0.023**	**0.23 (0.08-0.65)**	**0.006**	1.02 (0.49-2.14)	0.949
-Lithium	1		1		1	
Anxiety						
-Yes	1.56 (0.76-3.18)	0.222	1.63 (0.6-4.42)	0.341	2.04 (0.9-4.62)	0.087
-No	1		1		1	
Suicide in family						
-Yes	0.86 (0.25-3.02)	0.817	2.92 (0.84-10.12)	0.091	0.93 (0.27-3.23)	0.907
-No	1		1		1	
PTSD current.						
-Yes	0.95 (0.26-3.44)	0.943	1.38 (0.4-4.74)	0.612	2.25 (0.88-5.72)	0.090
-No	1		1		1	
Type of -Bipolar disorder						
-Bipolar type I	0.86 (0.41-1.83)	0.699	1.29 (0.46-3.63)	0.624	**0.32 (0.11-0.95)**	**0.040**
-Bipolar type II	1		1		1	
Psychotic symptoms	**1.23 (1-1.5)**	**0.046**	1.19 (0.92-1.54)	0.197	1.08 (0.9-1.3)	0.387

Table 1 Goodness of fit statistic for one to five class solutions for the Cubic GMM analysis.

No. of classes	Fit estimates [a]			P-values [b]	Classification accuracy		
	AIC [c]	*BIC* [d]	*adjBIC* [e]	*Bootstrap* [f]	*Entropy* [g]	*Class accuracy* [h]	*Class size (%)* [i]
1 class	25359	25438	25378	-	-	-	-
2 class	25254	25354	25278	< 0.001	0.805	0.96 0.89	79.83 20.17
3 class	25224	25345	25253	< 0.001	0.802	0.94 0.75 0.85	74.47 6.54 18.98
4 class	25181	25324	25216	< 0.001	0.779	0.92 0.91 0.79 0.80	9.74 60.29 18.36 11.60
5 class	25152	25315	25191	< 0.001	0.791	0.87 0.91 0.78 0.90 0.81	12.59 58.45 13.41 9.88 5.68

[a] *Statistical information criteria. A lower value indicates a better fit.*
[b] *A significant P-value indicates that the model with n number of classes is a better t of the data compared to a model with n – 1 number of classes*
[c] *Akaike Information Criteria.*
[d] *Bayesian Information Criteria.*
[e] *Sample size adjusted BIC.*
[f] *Bootstrap Likelihood Ratio Test.*
[g] *Entropy is estimated based on the average posterior probability and ranges from 0 to 1 where higher estimates represent greater classification accuracy for the model.*
[h] *Expresses the average accuracy of membership classification in each class. As the entropy measure it ranges from 0 to 1 where higher estimates represent greater classification accuracy.*
[i] *Distribution of the total sample into identified classes based on the posterior probability.*

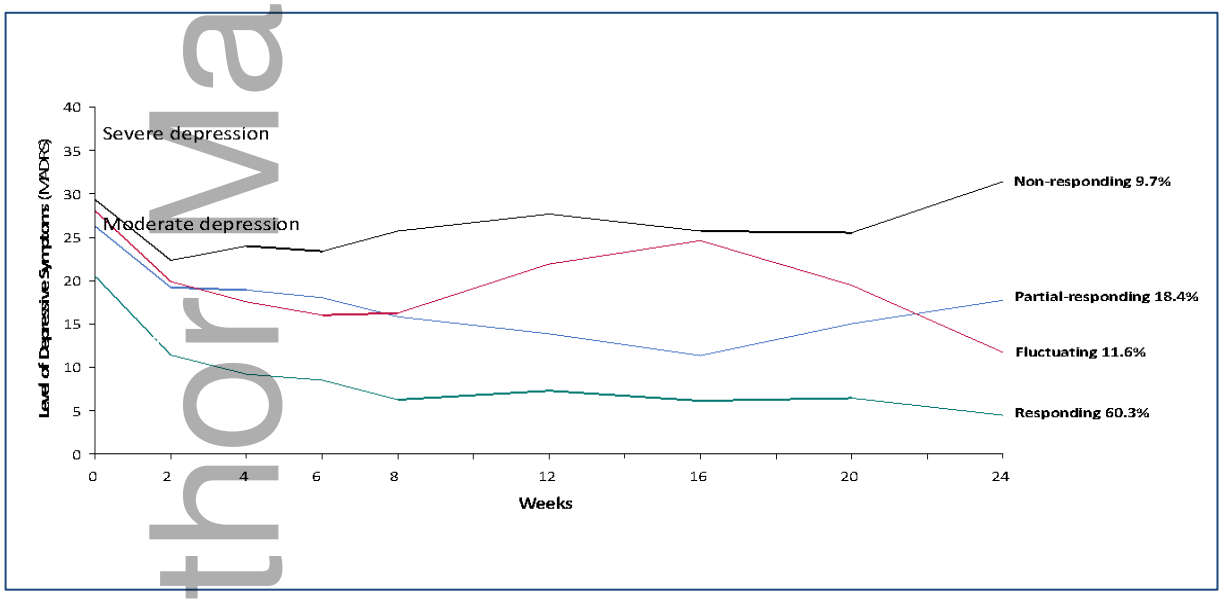
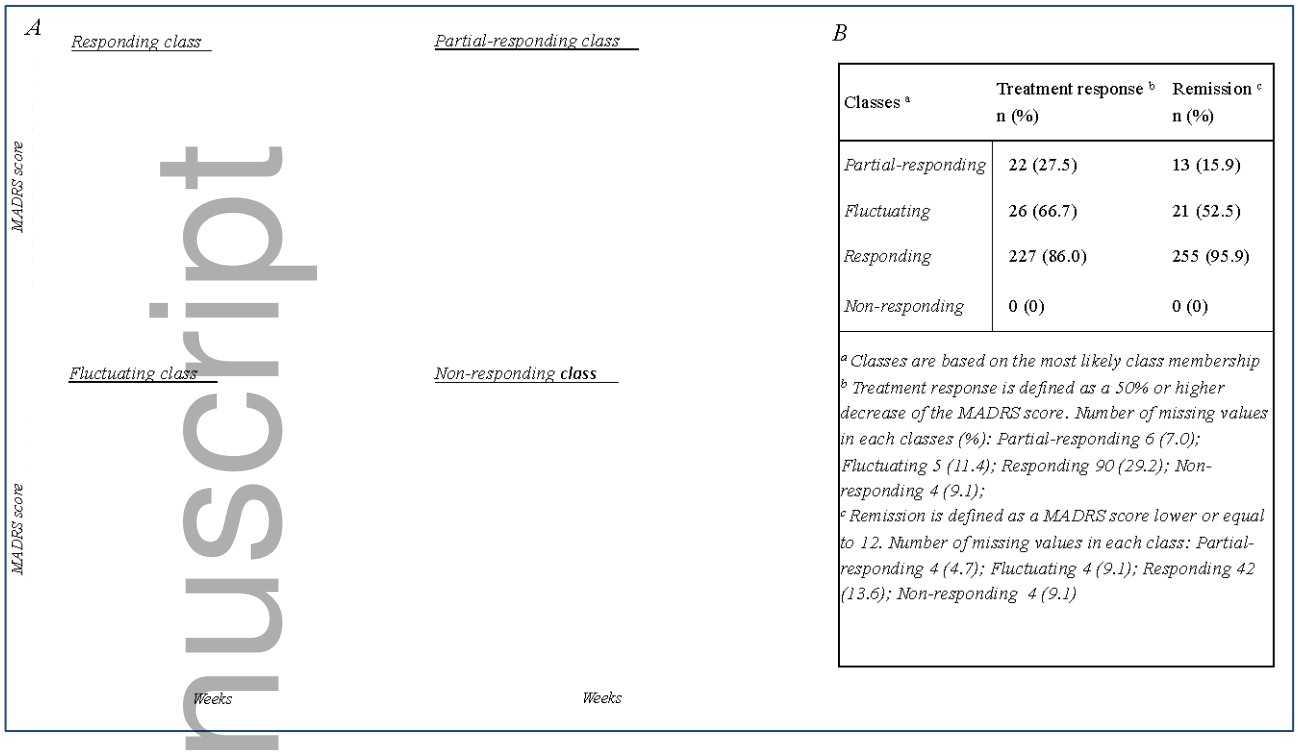


Figure 1 Trajectories of depressive symptoms measured by the MADRS score during 6 months of mood-stabilizing treatment among 482 outpatients with bipolar disorder.



B

Classes [a]	Treatment response [b] n (%)	Remission [c] n (%)
Partial-responding	22 (27.5)	13 (15.9)
Fluctuating	26 (66.7)	21 (52.5)
Responding	227 (86.0)	255 (95.9)
Non-responding	0 (0)	0 (0)

[a] Classes are based on the most likely class membership
[b] Treatment response is defined as a 50% or higher decrease of the MADRS score. Number of missing values in each classes (%): Partial-responding 6 (7.0); Fluctuating 5 (11.4); Responding 90 (29.2); Non-responding 4 (9.1);
[c] Remission is defined as a MADRS score lower or equal to 12. Number of missing values in each class: Partial-responding 4 (4.7); Fluctuating 4 (9.1); Responding 42 (13.6); Non-responding 4 (9.1)

Figure 2 Presentation of the four identified trajectory classes
A) Estimated mean trajectories of depressive symptoms measured by the MADRS score (bold lines) and the observed individual values within each of the identified classes. The trajectory classes are based on "most likely class membership".
B) Treatment response and remission in the different trajectory classes. All results are presented as valid percent, meaning percentage of valid (non missing) observations.

Table 2 Predictors for membership of the identified trajectory classes based on the univariable regression analysis. The Responding class is the reference. Significant findings are bold.

	Partial-responding class		Fluctuating class		Non-responding class	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Age at baseline	1.01 (0.99-1.04)	0.320	1 (0.97-1.04)	0.833	1.01 (0.98-1.04)	0.444
Gender						
-Female	1.05 (0.55-2)	0.878	2.25 (0.87-5.8)	0.093	1.57 (0.76-3.26)	0.221
-Male	1		1		1	
Race						
-All others	1.19 (0.58-2.46)	0.641	**2.43 (1.05-5.64)**	**0.038**	1.11 (0.5-2.44)	0.803
-Caucasian	1		1		1	
Education						
-College or more.	0.68 (0.33-1.37)	0.277	0.77 (0.3-1.97)	0.589	0.61 (0.29-1.3)	0.201
-High school or less	1		1		1	
-Employment						
-Not employed	**2.08 (1.06-4.08)**	**0.034**	1.73 (0.74-4.03)	0.203	1.67 (0.82-3.39)	0.155
-Employed/student	1		1		1	
Randomizarion						
-Quetiapine	**0.46 (0.24-0.89)**	**0.020**	**0.25 (0.09-0.66)**	**0.005**	1.06 (0.52-2.15)	0.871
-Lithium	1		1		1	
Previous psychiatric hospitalization						
-Yes	1.49 (0.79-2.8)	0.217	1.15 (0.51-2.57)	0.743	1.55 (0.78-3.07)	0.208
-No	1		1		1	
Psychotic symptoms	**1.27 (1.05-1.54)**	**0.013**	1.27 (0.99-1.61)	0.054	**1.23 (1.02-1.46)**	**0.026**
Anxiety						
-Yes	1.64 (0.84-3.17)	0.145	1.69 (0.72-3.97)	0.231	**2.59 (1.19-5.65)**	**0.016**
-No	1		1		1	
Psychiatris illness in family						
-Yes	1.41 (0.65-3.05)	0.380	2.27 (0.68-7.5)	0.181	1.43 (0.62-3.32)	0.401
-No	1		1		1	
Suicide in family						
-Yes	0.79 (0.22-2.81)	0.713	**2.75 (1-7.59)**	**0.050**	1.14 (0.37-3.52)	0.817
-No	1		1		1	
Previous suicide attempts						
-Yes	1.23 (0.64-2.37)	0.531	1.41 (0.61-3.23)	0.420	1.85 (0.92-3.71)	0.083
-No	1		1		1	
PTSD current.						
-Yes	1.47 (0.53-4.08)	0.463	2.06 (0.65-6.52)	0.221	**3.62 (1.53-8.6)**	**0.004**
-No	1		1		1	
Comorbidities						
-Yes	1.61 (0.84-3.09)	0.155	0.99 (0.41-2.43)	0.989	0.91 (0.42-1.95)	0.802
-No	1		1		1	
Alcohol abuse within the last 12 months						
-Yes	0.38 (0.05-2.66)	0.332	1.18 (0.3-4.68)	0.810	1.85 (0.67-5.11)	0.235
-No	1		1		1	
Alcohol abuse lifetime						
-Yes	0.97 (0.48-1.93)	0.923	1.4 (0.6-3.24)	0.437	1.83 (0.91-3.68)	0.092
-No	1		1		1	
Any substance abuse within the last 12 months						
-Yes	1.31 (0.52-3.29)	0.566	1.76 (0.61-5.11)	0.300	1.67 (0.66-4.19)	0.277
-No	1		1		1	
Any substance abuse lifetime						
-Yes	1.44 (0.76-2.73)	0.266	1.34 (0.58-3.05)	0.492	1.78 (0.89-3.58)	0.105
-No	1		1		1	
Type of -Bipolar disorder						
-Bipolar type I	1.54 (0.75-3.14)	0.238	1.16 (0.49-2.77)	0.732	**4.04 (1.4-11.65)**	**0.010**
-Bipolar type II	1		1		1	
CGI-mani at baseline	1.19 (0.9-1.56)	0.218	1.07 (0.74-1.54)	0.728	1.28 (0.99-1.65)	0.061
Age at first manic episode	1.02 (0.98-1.05)	0.273	1.02 (0.96-1.07)	0.496	1 (0.96-1.03)	0.836
Age at first depressive episode	1 (0.95-1.05)	0.920	0.99 (0.94-1.05)	0.780	0.99 (0.95-1.03)	0.492

Table 3 Predictors for membership of the different classes based on the multivariable regression analysis. The responding class is the reference. Significant findings are bold.

	Partial-responding class 1		**Fluctuating class (11.6)**		**Non-responding class (9.7**	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Race						
-All others	0.97 (0.43-2.18)	0.942	2.34 (0.89-6.17)	0.085	0.9 (0.4-2.04)	0.801
-Caucasian	1		1		1	
-Employment						
-Not employed	1.72 (0.83-3.53)	0.142	1.4 (0.52-3.72)	0.502	1.32 (0.61-2.82)	0.481
-Employed/student	1		1		1	
Randomizarion						
-Quetiapine	**0.44 (0.22-0.89)**	**0.023**	**0.23 (0.08-0.65)**	**0.006**	1.02 (0.49-2.14)	0.949
-Lithium	1		1		1	
Anxiety						
-Yes	1.56 (0.76-3.18)	0.222	1.63 (0.6-4.42)	0.341	2.04 (0.9-4.62)	0.087
-No	1		1		1	
Suicide in family						
-Yes	0.86 (0.25-3.02)	0.817	2.92 (0.84-10.12)	0.091	0.93 (0.27-3.23)	0.907
-No	1		1		1	
PTSD current.						
-Yes	0.95 (0.26-3.44)	0.943	1.38 (0.4-4.74)	0.612	2.25 (0.88-5.72)	0.090
-No	1		1		1	
Type of -Bipolar disorder						
-Bipolar type I	0.86 (0.41-1.83)	0.699	1.29 (0.46-3.63)	0.624	**0.32 (0.11-0.95)**	**0.040**
-Bipolar type II	1		1		1	
Psychotic symptoms	**1.23 (1-1.5)**	**0.046**	1.19 (0.92-1.54)	0.197	1.08 (0.9-1.3)	0.387