Pharmacological treatment profiles in the FACE-BD cohort: An unsupervised machine learning study, applied to a nationwide bipolar cohort

Sébastien Brodeur, Hugo Terrisse, Arnaud Pouchon, Ophelia Godin, Bruno Aouizerate, Valerie Aubin, Frank Bellivier, Raoul Belzeaux, Thierry Bougerol, Philippe Courtet, et al.

Version of Record: https://www.sciencedirect.com/science/article/pii/S0165032721001543
Manuscript_fa9f176714cb570fd14eb8ebe1714fb9

Pharmacological treatment profiles in the FACE-BD cohort: an unsupervised machine learning study, applied to a nationwide bipolar cohort

Sébastien Brodeur[7] M.D. M.Sc., Hugo Terrisse[11] M.Sc., Arnaud Pouchon[7] M.D. M.Sc., Ophelia Godin[1] PhD, Bruno Aouizerate[3] M.D. PhD, Valerie Aubin[9] M.D. PhD, Frank Bellivier[2] M.D. PhD, Raoul Belzeaux[5] M.D. PhD, Thierry Bougerol[7] M.D. PhD, Philippe Courtet[4] M.D. PhD, Caroline Dubertret[10] M.D. PhD, Sebastien Gard[3] M.D. PhD, Emmanuel Haffen[12] M.D. PhD, Chantal Henry[13] M.D. PhD, Marion Leboyer[1] M.D. PhD, Emilie Olié [4] M.D. PhD, Paul Roux[8] M.D. PhD, Ludovic Samalin[14] M.D. PhD, Raymund Schwan[6] M.D. PhD, The FondaMental Advanced Centers of Expertise in Bipolar Disorders (FACE-BD) Collaborators, Bruno Etain[2] M.D. PhD, Jean-Luc Bosson[11] M.D. PhD, Mircea Polosan[7] M.D. PhD

*List of FondaMental Advanced Center of Expertise (FACE-BD) collaborators:

FACE-BD Clinical Coordinating Center (Fondation FondaMental); B. Etain, C. Henry, E. Olié, M. Leboyer, E. Haffen and PM Llorca;

FACE-BD Data Coordinating Center (Fondation FondaMental); V. Barteau, S. Bensalem, O. Godin, H. Laouamri, and K. Souryis;

FACE-BD Clinical Sites and Principal Collaborators in France;

1. AP-HP, DHU PePSY, Pôle de Psychiatrie et d'Addictologie des Hôpitaux Universitaires H Mondor, Créteil ; S. Hotier, A. Pelletier, N. Drancourt, JP. Sanchez, E. Saliou, C. Hebbache, J. Petrucci, L. Willaume and E. Bourdin ;
2. AP-HP, GH Saint-Louis–Lariboisière–Fernand Widal, Pôle Neurosciences, Paris ; F. Bellivier, M. Carminati, B. Etain, E. Marlinge, M. Meyrel;
3. Hôpital C. Perrens, Centre Expert Trouble Bipolaire, Service de Psychiatrie Adulte, Pôle 3-4-7, Bordeaux ; B. Antoniol, A. Desage, S. Gard, A. Jutant, K. Mbailara, I. Minois, and L. Zanouy;
4. Département d'Urgence et Post Urgence Psychiatrique, CHRU Montpellier, Montpellier ; C. Abettan, L. Bardin, A. Cazals, P. Courtet, B. Deffinis, D. Ducasse, M. Gachet, A. Henrion, E. Martinerie, F. Molière, B. Noisette, E. Olié and G. Tarquini;
5. Pôle de Psychiatrie, addictologie et pédopsychiatrie, Hôpital Sainte Marguerite, Marseille ; R. Belzeaux, N. Correard, F. Groppi, A. Lefrere, L. Lescalier., E. Moreau, J. Pastol, M. Rebattu, B. Roux and N. Viglianese;
6. Service de Psychiatrie et Psychologie Clinique, CHU de Nancy, Hôpitaux de Brabois, Vandœuvre Les Nancy ; R. Cohen, Raymond Schwan, J. P. Kahn, M. Milazzo, and O. Wajsbrot-Elgrabli;
7. Service Universitaire de Psychiatrie, CHU de Grenoble et des Alpes, Grenoble ; T. Bougerol, B. Fredembach, A. Suisse, S. Brodeur, A. Pouchon, and M. Polosan
8. Centre Hospitalier de Versailles, Service Universitaire de Psychiatrie d'adultes, Le Chesnay ; A. M. Galliot, I. Grévin, A. S. Cannavo, N. Kayser, C. Passerieux, and P. Roux; Service de Psychiatrie,
9. Centre Hospitalier Princesse Grace, Monaco ; V. Aubin, I. Cussac, M.A Dupont, J. Loftus, and I. Medecin;

10. AHPH, Departement de Psychiatrie, Hopital Louis Mourier, Colombes, France ; C. Dubertret, N. Mazer, C. Portalier.
11. TIMC-IMAG, University Grenoble Alpes, France
12. Service de psychiatrie, CHU de Besançon, laboratoire de Neurosciences, Université de Franche-Comté ; E. Haffen
13. Université Paris Descartes, Pôle Hospitalo-Universitaire Paris 15ème, GHU, Centre Hospitalier Sainte Anne.
14. CHU de Clermont Ferrand, Pôle de Psychiatrie, Clermont Ferrand : Service de Psychiatrie de l'adulte B, Centre Expert Trouble Bipolaire, CHU de Clermont-Ferrand, Clermont-Ferrand, France : PM. Llorca, L. Samalin, L., C. Moreau, D. Lacelle, S.Pires, C. Doriat, and O. Blanc.

Corresponding author:
Mircea Polosan
Pôle de Psychiatrie et Neurologie, CHU Grenoble, France
Grenoble Institute of Neurosciences
Office : (+33) 0*4 76 765 414
Fax : (+33) 0*4 76 765 225
E-mail : MPolosan@chu-grenoble.fr

AKNOWLEDGEMENT:

PM Llorca, V. Barteau, S. Bensalem, O. Godin, H. Laouamri, K. Souryis, S. Hotier, A. Pelletier, N. Drancourt, JP. Sanchez, E. Saliou, C. Hebbache, J. Petrucci, L. Willaume, E. Bourdin, M. Carminati, E. Marlinge, M. Meyrel, B. Antoniol, A. Desage, A. Jutant, K. Mbailara, I. Minois, and L. Zanouy, C. Abettan, L. Bardin, A. Cazals, B. Deffinis, D. Ducasse, M. Gachet, A. Henrion, E. Martinerie, F. Molière, B. Noisette, G. Tarquini, N. Correard, F. Groppi, A. Lefrere, L. Lescalier., E. Moreau, J. Pastol, M. Rebattu, B. Roux, N. Viglianese, R. Cohen, J. P. Kahn, M. Milazzo, O. Wajsbrot-Elgrabli, B. Fredembach, A. Suisse, A. M. Galliot, I. Grévin, A. S. Cannavo, N. Kayser, I. Cussac, M.A Dupont, J. Loftus, I. Medecin, N. Mazer, C. Portalier contributed to database collection.

ABSTRACT

Background:
Despite thorough and validated clinical guidelines based on bipolar disorders subtypes, large pharmacological treatment heterogeneity remains in these patients. There is limited knowledge about the different treatment combinations used and their influence on patient outcomes. We attempted to determine profiles of patients based on their treatments and to understand the clinical characteristics associated with these treatment profiles.

Methods:
This multicentre longitudinal study was performed on a French nationwide bipolar cohort database. We performed hierarchical agglomerative clustering to search for clusters of individuals based on their treatments during the first year following inclusion. We then compared patient clinical characteristics according to these clusters.

Results:
Four groups were identified among the 1795 included patients: group 1 ("heterogeneous" n=1099), group 2 ("lithium" n= 265), group 3 ("valproate" n=268), and group 4 ("lamotrigine" n=163). Proportion of bipolar 1 disorder, in groups 1 to 4 were: 48.2%, 57.0%, 48.9% and 32.5%. Groups 1 and 4 had greater functional impact at baseline and a less favorable clinical and functioning evolution at one-year follow-up, especially on GAF and FAST scales.

Limitations:
The one-year period used for the analysis of mood stabilizing treatments remains short in the evolution of bipolar disorder.

Conclusions:
Treatment profiles are associated with functional evolution of patients and were not clearly determined by bipolar subtypes. These profiles seem to group together common patient phenotypes. These findings do not seem to be influenced by the duration of disease prior to inclusion and neither by the number of treatments used during the follow-up period.

KEYWORDS:
Bipolar disorder, maintenance treatment, mood stabilizers, unsupervised machine learning, functioning.

INTRODUCTION

Bipolar disorders (BD) are classified according to a clinical spectrum predominantly represented by bipolar 1 disorder (BD1) and bipolar 2 disorder (BD2). Fluctuations in energy and mood state are core features of this chronic disease **(Grande et al., 2016; Vieta et al., 2018)**. BD affect daily functioning and quality of life of people who suffer from it, mainly because of their early onset, the severity of the disease and the recurrence of episodes.

Acute treatment of mood relapses is essential, but because of the chronic nature of bipolar disorder, attention should also be emphasized on maintenance treatment **(Carvalho et al., 2020)**. Most patients with BD will require mood stabilizers as the basis of the maintenance treatment **(Yatham et al., 2018)**. Despite the emergence and development of new mood stabilizers since the 1960s, lithium remains the treatment of choice for maintenance therapy because of its efficacy in reducing the risk of relapses and recurrences (especially manic episodes) and its ability in maintaining sustained periods of remission **(Bauer, 2020; Hayes et al., 2016; Post, 2018; Tondo et al., 2019)**. However, the use of second generation antipsychotics, as mood stabilizers, remains very common in clinical practice **(Rhee et al., 2020)**.

Usually, the clinician will choose among several maintenance treatments proposed in the 1st and 2nd line options according to practice guidelines **(Yatham et al., 2018)**. Even when the clinician knows the bipolar subtypes accurately, they will proceed by a trial-and-error approach until they find a treatment that is effective and well tolerated by the patient (Alda and Manchia, 2018). This period can be very long and may bring consequences for the patient and substantial costs for the society **(Grande et al., 2016)**. Also, it is surprising that, despite established treatment guidelines and known epidemiological factors predisposing to a good clinical response, it remains very difficult to individualize treatment. In this regard, there is great heterogeneity **(López-Villarreal et al., 2019; Sarrazin et al., 2018; Wolfers et al., 2018)** among patients in terms of their clinical evolution, beyond the subtypes of bipolarity, but also in terms of their treatments **(Bauer et al., 2018; Bauer et al., 2013).** Usually, for one given patient, different treatments are tried and more often, in combination **(Bauer et al., 2013; Fornaro et al., 2016)**. In this context, it is mandatory to understand the treatment choices made by clinicians, in real clinical practice, and to figure out the rationale behind these choices. To date, few studies have succeeded in determining the prescribing patterns of clinicians treating bipolar patients.

In order to further understand the patient's treatment heterogeneity, some statistical methods are better suited to explore this variability **(Passos et al., 2019a)**. More precisely, data mining methods render possible understanding of complex data from large datasets, through computer and statistical methods **(Iniesta et al., 2016)**. Rather than proposing a hypothesis and verifying it with traditional statistical tests, data mining let the data speak by themselves **(Passos et al., 2019b)**. Conventional methods are hypothesis-based and not designed to manage different variables from diverse modalities simultaneously **(Menke, 2018)**. Machine learning has been used, in the last decade, for almost all psychiatric disorders. However, to date no study has attempted to find, in bipolar disorder, the different patterns of treatments, in a naturalistic way.

By using a large cohort of patients suffering from bipolar disorder in France, we tried to determine if groups of patients' profiles could be formed based on their treatments or combination of treatments, within a one-year period. Following this, the objective was to compare this new classification with the current bipolar subtypes classification.

MATERIAL AND METHODS

Study design

This multicentre longitudinal study was based on the patient database from FondaMental Advanced Centres of Expertise for Bipolar Disorders (FACE-BD) cohort. This cohort has been fully described elsewhere **(Godin et al., 2020; Henry et al., 2015; Henry et al., 2011; Henry et al., 2017).** We searched for clusters of individuals based on their treatments during the first year following inclusion. We then compared the patient's baseline characteristics and one-year follow-up according to these clusters. We similarly compared the patients according to the usual bipolar subtypes.

Setting

The FACE-BD cohort recruited patients from a national network of 12 expert centres in France, set up by the French FondaMental Foundation (www.fondation-fondamental.org). General practitioners and psychiatrists referred outpatients with bipolar disorders to these Expert Centres. Recruitment started in June 2008. We extracted the data on 02/04/2019. The FACE-BD cohort includes a follow-up at 12, 24 and 36 months (respectively V12, V24 and V36).

The assessment protocol, including a letter of information for patients, was approved by the institutional review board (CPP-Ile de France IX, January 18th, 2010), in accordance with French laws for non-interventional studies. Anonymized data are stored in a national database that was approved by the French body overseeing the safety of computerized databases (Commission Nationale de l'Informatique et des Libertés, DR-2011-069).

Participants

Inclusion criteria for our study included an evaluation in one of the Expert center at baseline and at least one follow-up visit, a diagnosis of bipolar disorder and ambulatory patients between 18 and 65 years old. The criteria for non-inclusion were having a schizophrenia spectrum and other psychotic disorder diagnosis.

Variables

The structured clinical interview for DSM-IV Axis I disorders (SCID) **(First, 1996)** was used for the diagnosis of bipolar disorders, as well as bipolar subtypes. We calculated days on treatment during the first year for each active ingredient. Hence, each patient can have up to 365 days on treatment for each active ingredient. Evaluation of patients at inclusion and V12 used clinical evaluation and biological measurements. Clinical evaluation at inclusion and V12 included characterisation of mood episodes within last year and questionnaires. Clinical evaluation at inclusion also included characterisation of mood episodes within lifetime. We evaluated the mood using the Montgomery and Asberg Depression Rating Scale (MADRS) **(Montgomery and Asberg, 1979)**, the Young Mania Rating Scale (YMRS) **(Young et al., 1978),** the Quick Inventory of Depressive Symptoms 16 items (QIDS-SR16) **(Rush et al., 2003)** and the Altman Mania Rating Scale (AMRS) **(Altman et al., 1997)**. We evaluated the global functioning using the Global Assessment of Functioning (GAF) **(Endicott et al., 1976)** and the Functioning Assessment Short Test (FAST) **(Rosa et al., 2007)**. We assessed the severity of the disease using the Clinical Global Impression—Severity scale (CGI-S) **(Guy, 1976)**. We assessed the quality of sleep using the Pittsburgh Sleep Quality Index (PSQI) **(Buysse et al., 1989)**. At inclusion, we evaluated childhood trauma using the Childhood

Trauma Questionnaire (CTQ) **(Bernstein, 1998)** and childhood symptomatology of ADHD using the Wender Utah Rating Scale (WURS) **(Ward et al., 1993)**. We evaluated adherence to pharmacological treatment using the Medication Adherence Rating Scale (MARS) **(Misdrahi et al., 2004)** and side effects of treatments using the Patient Rated Inventory of Side Effects inventory (PRISE-M) **(Rush, 1999)**. Biological measurements at inclusion and V12 were creatinine, HDL, LDL, total cholesterol, triglycerides, ASAT, ALAT, fasting blood sugar, platelets, TSH, T3 and T4.

Data sources/measurement

Expert Centres filled up the data through a web application, e-bipolar©. Clinicians reported all treatments received. We then computed the days on treatment for each active ingredient during the first year following inclusion. The QIDS-SR16, AMRS, PSQI, PRISE-M, MARS, WURS and CTQ were self-administered. The clinician administered the MADRS, YMRS, FAST, CGI-S and GAF.

Statistical methods

The available data determined the sample size.

We analysed the quantitative variables without modification. We performed hierarchical agglomerative clustering to understand patterns of treatments for bipolar patients. The features we used to classify the patients were the days on treatment for each active ingredient present in the dataset during the first year following the inclusion. The distance metric was the squared Euclidean distance. We used Ward's agglomerative methods. We chose the number of clusters using the visual analysis of the dendrogram associated to the hierarchical agglomerative clustering and the interpretability of the groups. We then compared the clinical outcomes between the identified clusters. We used Chi-squared tests for categorical outcomes if all expected frequencies of the contingency table were higher than 5, else we used Fisher's exact test. We used ANOVA for continuous outcomes. For continuous outcomes at one-year follow-up, if the value at inclusion was available, we used ANCOVA instead of the ANOVA, to account for the inclusion value. We performed the same analysis for bipolar subtypes comparisons. All tests were bilateral, with a first-type error level of 0.05. As the purpose of this paper is to describe patient profiles, we did not correct for multiple testing. However, we provide all calculated p-values for transparency. We did not impute missing data and performed the analysis for complete cases only. We performed all analysis using R software version 3.6.3. As a post-hoc analysis, we compared the ANCOVA for the FAST scale using either the bipolar subtypes or the clusters using Akaike Information criterion (AIC). The lowest the AIC, the better the model explains the data.

RESULTS

3302 patients have been included in the FACE-BD cohort since the implementation of expert centers in France. 1494 patients were lost to follow-up after one year. 1808 patients were retained. Thirteen patients were excluded because they had inconsistent treatment dates. A total of 1795 patients were analyzed.

Patients within the study population had a median age of 41 years and consisted of 60.5% females and 39.5% males. Diagnosis of bipolar 1 was found in 865 patients, bipolar 2 in 764 patients and bipolar NOS in 166 patients. A proportion of 81.7% of the patients were referred to expert centers by a psychiatrist. Median age at first psychotropic treatment was 25-year-old and 21 at first mood episode. At baseline visit, 29.7% of patients had a current episode. In the last year before inclusion, 39% had been hospitalized at least once. At baseline, 69.9% of patients declared having complete remission between episodes. At inclusion, 39.5% of

patients had attempted suicide in their lives. Results for the total included population can be found in Table 1 and supplementary material.

We found 117 different treatments in the database. Among these, 107 treatments were used by less than 5% of the study cohort during the first-year follow-up. The 10 others were: lithium (36.2%), valproate (30.1%), lamotrigine (20.7%), quetiapine (15.1%), aripiprazole (13.5%), venlafaxine (11.5%), escitalopram (7.9%), olanzapine (6.2%), alprazolam (6.2%), cyamemazine (6%). When these treatments were used, this was for more than 300 days throughout the study period (*see complete data of treatment in supplementary material*) for more information). Information regarding the other treatments can be found in supplementary material.

Main results

Hierarchical ascendant clustering

Hierarchical ascendant clustering (HAC) was performed with 117 treatments. The resulting dendrogram is represented in Figure 1.

At the dendrogram level, a cluster stood out from the classification immediately. This group had a significantly different treatment profile compared with the other groups established, even after several divisions in the number of patients. This is followed by other divisions in the dataset. Overall, four groups were selected for their size and clinical relevance and by using the Calinski function on RStudio. That method initially suggested that we should choose between 2 clusters. However, this choice of cluster was purely statistical and had no clinical relevance as we would end up with a group of 1530 patients and 265 patients. Subsequently, in order to be as comprehensive as possible but also keeping in mind the clinical perspective, Calinski function suggested that the optimal division choice was 4 clusters. This division in four groups is represented by the red line on the Dendrogram Cluster (Fig 1).

The first group in the classification represented 1099 out of 1795 included patients. The treatments received during one year in this group were mainly: lithium (27.8%), valproate (24.3%), quetiapine (18.1%) and aripiprazole (15.7%). Group 2 was the one that stood out to the right during the first division of the dendrogram and consisted of 265 patients. The treatments received during one year in this group were mainly: lithium (99.2%), lamotrigine (23.4%) and venlafaxine (17%). Group 3 consisted of 268 patients and the main treatments were mainly valproate (88.4%), venlafaxine (33.2%) and lithium (18.3%). Group 4 consisted of 163 patients. The treatments received during one year in this group were mainly: lamotrigine (93.9%), quetiapine (27.6%) lithium (19.6%) and aripiprazole (17.2%) (see Figure 2). In the rest of the paper, we will refer to group 1 as “heterogeneous”, group 2 as “lithium”, group 3 as “valproate” and group 4 as “lamotrigine”. *(Figure 2 and Table 2.1)*

Clinical description according to clusters *(Tables 2.2 and 2.4)*
Group “heterogeneous” had a higher proportion of BD1 (48.2%) versus BD2 (41.8%). Group “lithium” was formed by a higher proportion of BD1 (57.0%) while there was a higher proportion of BD2 (57.1%) in the group “lamotrigine” compared with the other groups. The median age was higher (43 [34.8; 53]) in group “valproate”. There was no statistically significant difference in the type of referring doctors, as well as in the occupational status of the different patients from the individualized groups. The first three groups had received a similar number of treatments in the year of follow-up (median =2). Group 4 had used more

treatments during the year of follow-up (median=3). Duration of disease prior to inclusion was comparable in the first three groups (median=15 years) except in the "lamotrigine" group with a median of 14 years.

Groups "heterogeneous" and "lithium" started psychotropic treatment at a younger age than other groups. At first episode, patients were older in group "valproate" (23 [18; 31]). At the lifetime assessment level, patients in all groups had at least one MDE in greater proportion compared to other types of mood episodes. Group "lithium" had a higher proportion of patients who had at least one-lifetime manic episodes (53.1%) as well as manic episodes with psychotic characteristics (39.8%). There was a higher proportion of rapid cycles in the group "lamotrigine" (28.5%). There was no statistically significant difference between groups in the number of patients hospitalized at V12. At the initial visit, patients had mainly MDE in the past year compared with other types of episodes. Also, the group "lithium" had more manic episodes with (15.2%) and without (22.1%) psychotic characteristics in the last year, at the initial visit. There was no statistically significant difference in patients who already attempted suicide between the different clusters established at the initial visit and at V12. The proportion of patients with a current episode, at V12, decreased in all groups compared to the initial visit.

Self-administered questionnaire *(Table 2.3)*
On the QIDS-SR16 questionnaire at the initial visit, group "lamotrigine" had a higher score (11 [6; 16]) and was in the moderate depression category (11–15). The decrease in depressive symptoms was greater in this same group at V12 (6 [4; 12]). On the ALTMAN scale, the four clusters had a score of less than 5 and indicated a low probability of a manic or hypomanic episode, at both initial and follow-up visits. On the PSQI scale, the scores of each cluster at the initial visit exceeded the threshold of more than 5 indicating sleep problems for most patients. On the WURS scale, the median score of patients in all subtypes was less than 46, suggesting the absence of ADHD in childhood. Side effects were assessed using the PRISE-M scale on a range of 0 to 62 (62 corresponding to a maximum of side effects). Groups "heterogeneous" (12 [6; 19]) and "lamotrigine" (13 [6; 19.8]) had more side effects from their treatment than groups "lithium" and "valproate". Adherence to treatment was similar across all clusters according to MARS scale.

Clinician administered scale *(Table 2.3)*
On the MADRS scale, patients in groups "lithium" and "valproate" had a lower score, while patients in groups "heterogeneous" and "lamotrigine" had greater depressive symptoms, both at baseline and V12. All groups showed an improvement between baseline and V12. Based on YMRS scale, patients had no hypomanic/manic symptoms at either V0 or V12.

All groups had functional impairment on the FAST scale at baseline. A score of 12 or more indicated problems with functioning in daily life. Group "lamotrigine" had a higher median score (22 [12.8; 29]) than other clusters. At V12, group "lithium" had the lowest scores on the FAST scale (10 [3; 22]), and more than 50% of them did not have functioning impairment. As well, group "valproate" had a median score of (12 [3; 23]) in the lower portion of the diagnostic threshold for functional problems on this scale. For the treatment profiles classification, the AIC was 9253 for the ANOVA at V0 and 6957 for the ANCOVA at V12. For bipolar subtypes classification, the AIC was 9258 for the ANOVA at V0 and 6960 for the ANCOVA at V12. This indicates that treatment profiles better explain the FAST at baseline and its evolution at V12 than bipolar subtypes.

On the CGI scale, most patients were in the mildly-ill category, both at baseline and V12, in all clusters. At V12 only group "lithium" had changed category with a median score of 2 [1; 4] indicating the category "borderline mentally ill". On the Global Assessment of Functioning Scale [GAF] at V0, there were more patients in the 100-71 category [absent or transient symptoms] and 70-51 [mild symptoms of moderate intensity] with no statistically significant difference between groups established. At V12, Group "lithium" patients had a higher proportion in category 71–100 GAF [59.2%], and a lower proportion in the category 50-0.

Psychiatric and somatic comorbidities *(Table 2.5)*
There was a higher proportion of comorbid anxiety disorders in cluster "lamotrigine" (51.6%), as well as feeding and eating disorders [30.7%]. At the somatic level, there was a higher proportion of hypertension (11.6%) and hair loss (20.8%) in group "valproate". However, there was a higher proportion of thyroid disease (17.4%) in the group "lamotrigine". There was no clinically relevant difference between clusters according to the BMI level.

Biological parameters
There was no clinically relevant difference in blood parameters between the different groups.

DISCUSSION

Key results
Our study demonstrated that hierarchical ascending classification categorized patients according to different pharmacological treatment profile. In fact, four groups were identified, from the treatments they received, over a period of one year: group "heterogeneous", group "lithium", group "valproate" and group "lamotrigine". Groups "heterogeneous" and "lamotrigine" had a more pronounced depressive symptomatology and greater functional impact, at baseline, compared with groups "lithium" and "valproate". Despite this difference already existing at baseline between the 4 groups, group "lithium" and group "valproate" had a more favourable evolution based on different clinical scales, especially the ones measuring functioning (GAF, FAST). Moreover, these findings do not seem to be influenced by the duration of disease prior to inclusion and neither by the number of treatments used during the follow-up period.

This new classification seems to demonstrate clinical relevance. In fact, patients in group "lithium" were functioning better. Furthermore, these treatment clusters were formed beyond classical bipolarity subtypes. Treatments received during the one-year period, either in mono or polytherapy, seemed to further explain the functional evolution of the patients during this one-year follow-up.

Strengths
We obtained data from all patients (1795 patients) followed in FondaMental Academic Centres of Expertise for Bipolar Disorders, in France. This brings us results, with an interesting scope, since they reflect a population of specialized clinics, across a whole country. This is even more relevant, since randomized controlled clinical trials sometimes struggle to include population that reflects the clinical reality **(Vieta et al., 2018)**. Moreover, data obtained for each patient were exhaustive (self-administered scales, clinician-administered scales, blood tests, etc.). Also, the methods used (data mining) made possible the exploitation of a large data set and allowed to find relevant links between these same data.

Finally, using individualized drugs instead of drug classes permitted to exploit the data set without any boundary.

Interpretation

Based on results obtained, treatment used over one-year period seems to better predict functional prognosis than bipolar subtypes. Previous literature highlighted the differences in terms of functional evolution for bipolar disorder subtypes 1 and 2 **(Yatham et al., 2018)**. Beyond bipolarity subtypes, patients predominantly treated with lithium in our cohort during one-year follow-up, were doing better than all other patients in other clusters. Beyond clinical scores, which were all improved in all clusters, functioning in group "lithium" was improved in a clinically relevant way and also in a statistically significant manner. This finding is consistent with what is stated by some authors about the place that lithium should deserve in BD1 and BD2 patients **(Tondo et al., 2019)**; this adding up to lithium's potential neuroprotective effects **(Hozer et al., 2020)**. This is even more important since functioning is a therapeutic target which has been lately studied, even in euthymic patients with BD **(Léda-Rêgo et al., 2020; Sole et al., 2018)**.

Treatment of bipolar disorders is heterogenous in clinical practice **(Bauer et al., 2013)** as shown by the use of several mood-stabilizing treatments. In the same way, for more than half of the patients it was difficult, according to our methodology, to group them solely based on treatment. These patients were clustered in group "heterogeneous" (n=1099). This strong heterogeneity is in itself a reality, but could it be a different phenotype of the disease or a sign of the non-systematic use of first-line drugs by clinicians?

Side effects associated with lithium and valproic acid intake are often a hindrance for clinicians in their daily practice. However, our study shows fewer side effects in groups that had many patients treated with these drugs (groups "lithium" and "valproate"), compared with groups "heterogeneous" and "lamotrigine". In addition, renal, metabolic and thyroid blood tests were comparable in all four groups. Given our current results, concerns about side effects associated with lithium do not seem to be that justified compared to other treatments although based on the short maintenance period.

The profiles formed according to the treatments seem to group together common patient phenotypes. Could this additional classification explain common pathophysiological mechanisms? This way of conceptualizing bipolar disorders, which is different from the current classification is being studied by other groups **(Fuente-Tomas et al., 2019; Tamminga et al., 2017).**

Limitations

One of the major limitations of the study remains the duration of the analysis. Indeed, the one-year period under mood stabilization remains short in the course of bipolar disease. In addition, drug dosage was not taken into account in our study in order to form different clusters. Previous treatment section prior to inclusion had a high rate of missing data potentially due to recall bias, which limited this data analysis. Also, patients could be recruited at any stage of the illness in our study. Nevertheless, the duration of illness prior to inclusion didn't have an impact on our results. In fact, the duration of illness prior to inclusion in our new clusters was very similar. Also, non-pharmacological treatment section of the database was not thoroughly documented, such as psychotherapy. Moreover, combination of pharmacological and non-pharmacological treatments would have been very useful to better understand the impact of combining these two approaches. In addition, due to missing data

regarding the total number of thymic episodes, patient polarity index was not calculated. Based on the chosen methodology and our database, all the reasons behind clinician's choice for maintenance treatment were not exhaustively available for analysis. Unsupervised machine learning does not have the same aim as a randomized controlled trial (RCT). In fact, it can help revealing new findings among clinical data from patients we encounter in daily practice. Moreover, it can generate a hypothesis which can be further explored through RCTs. In that sense, these results remain exploratory and should therefore be interpreted with caution. Despite these limitations, observational studies are well known for their nonrestrictive inclusion criteria and to be more representative of real-life patients followed up in specialized clinics.

In conclusion, treatment profiles were associated with functional evolution of patients and were not clearly determined by bipolar subtypes. These findings do not seem to be influenced by the duration of disease prior to inclusion and neither by the number of treatments used during the follow-up period. This sheds light on future research to better understand these specific phenotypes beyond bipolar subtypes defined by the DSM-5. An additional question would be to determine what factors predict accurately the odds to be part of one group or another at baseline visit. Our findings encourage the use of supervised machine learning approaches to better predict the clinical response, functional evolution and side effects to different treatments.

DISCLOSURE

Dr. Brodeur, H. Terrisse, Dr. Pouchon, O. Godin, Dr. Aubin, Dr. Bellivier, Dr. Belzeaux, Dr. Bougerol, Dr. Courtet, Dr. Dubertret, Dr. Gard, Dr. Henry, Dr. Leboyer, Dr. Roux, Dr. Schwan, Dr. Bosson have nothing to disclose. Dr. Aouizerate reports personal fees from Janssen-Cilag, personal fees from Lilly, personal fees from Sanofi, during the conduct of the study. Dr. Haffen reports personal fees and non-financial support from Janssen, personal fees and non-financial support from Lundbeck, personal fees and non-financial support from Otsuka, outside the submitted work. Dr. Olié reports personal fees from Janssen-Cilag, outside the submitted work. Dr. Samalin reports personal fees and non-financial support from Janssen, personal fees and non-financial support from Lundbeck, personal fees and non-financial support from Otsuka, outside the submitted work.. Dr. Etain reports grants from INSERM, grants from Assistance Publique des Hôpitaux de Paris, grants from Agence Nationale pour la Recherche, grants from Fondation de France, grants from Research Council of Norway, personal fees from SANOFI, outside the submitted work. Dr. Polosan reports personal fees from Lundbeck, personal fees from Janssen-Cilag, outside the submitted work.

The FondaMental Advanced Centers of Expertise in Bipolar Disorders (FACE-BD) Collaborators have no conflicts of interest to disclose except PM Llorca. PM Llorca has received grants, honoraria, or consulting fees from ESAI, Gedeon Richeter, Janssen, Lundbeck, Otsuka and Sanofi.

FUNDING

This work was supported (in part) by the Fondation FondaMental, Créteil, France and by the Investissements d'Avenir programs managed by the ANR under references ANR-11-IDEX-0004-02 and ANR-10-COHO-10-01. This funding source had no role in the study design, data collection, analysis, preparation of the manuscript, or decision to submit the manuscript for publication. M Polosan had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

AUTHORS CONTRIBUTION:

Sébastien Brodeur worked on literature search, tables, study design, data collection, data analysis, data interpretation, tables and writing. Hugo Terrisse worked on study design, data collection, data analysis, data interpretation, figures, tables and writing. Sébastien Brodeur and Hugo Terrisse contributed equally to this manuscript. Jean-Luc Bosson worked on study design.

Mircea Polosan worked on literature search, study design, data collection, data analysis, data interpretation and writing. Arnaud Pouchon worked on data interpretation and data analysis. Bruno Etain worked on data collection, data analysis, data interpretation and writing. Ophelia Godin, Bruno Aouizerate, Valerie Aubin, Frank Bellivier, Raoul Belzeaux, Thierry Bougerol, Philippe Courtet, Caroline Dubertret, Sebastien Gard, Emmanuel Haffen, Chantal Henry, Marion Leboyer, Emilie Olié, Christine Passerieux, Ludovic Samalin, Raymund Schwan worked on data collection and reviewed the article.

REFERENCES

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.URL http://www.R-project.org/.

Alda, M., Manchia, M., 2018. Personalized management of bipolar disorder. Neuroscience letters 669, 3–9.

Altman, E. G., Hedeker, D., Peterson, J. L., Davis, J. M., 1997. The Altman Self-Rating Mania Scale. Biol Psychiatry 42, 948–955.

Bauer, M., 2020. Lithium: about discrepancies between efficacy and clinical use. Acta Psychiatrica Scandinavica 142, 159-160.

Bauer, M., Andreassen, O.A., Geddes, J.R., Vedel Kessing, L., Lewitzka, U., Schulze, T.G., Vieta, E., 2018. Areas of uncertainties and unmet needs in bipolar disorders: clinical and research perspectives. The lancet. Psychiatry 5, 930–939.

Bauer, M., Glenn, T., Alda, M., Sagduyu, K., Marsh, W., Grof, P., Munoz, R., Severus, E., Ritter, P., Whybrow, P.C., 2013. Drug treatment patterns in bipolar disorder: analysis of long-term self-reported data. Int J Bipolar Disord 1, 5.

Bernstein, D. F. L., 1998. Childhood trauma Questionnaire, A retrospective Self-Report. Manual. The Psychological Corporation, San Antonio. .

Buysse, D. J., Reynolds, C. F., 3rd, Monk, T.H., Berman, S. R., Kupfer, D. J., 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 28, 193–213.

Carvalho, A. F., Firth, J., Vieta, E., 2020. Bipolar Disorder. N Engl J Med 383, 58–66.

Endicott, J., Spitzer, R. L., Fleiss, J. L., Cohen, J., 1976. The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. Arch Gen Psychiatry 33, 766–771.

First, M.B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W., 1996. Structured clinical interview for DSM-IV axis I disorders, clinician version (SCID-CV). American Psychiatric Press., Washington, DC.

Fornaro, M., De Berardis, D., Koshy, A.S., Perna, G., Valchera, A., Vancampfort, D., Stubbs, B., 2016. Prevalence and clinical features associated with bipolar disorder polypharmacy: a systematic review. Neuropsychiatr Dis Treat 12, 719–735.

Fuente-Tomas, L., Arranz, B., Safont, G., Sierra, P., Sanchez-Autet, M., Garcia-Blanco, A., Garcia-Portilla, M.P., 2019. Classification of patients with bipolar disorder using k-means clustering. PloS one 14, e0210314.
Godin, O., Leboyer, M., Mazroui, Y., Aouizerate, B., Azorin, J.M., Raoul, B., Bellivier, F., Polosan, M., Courtet, P., Dubertret, C., Henry, C., Kahn, J.P., Loftus, J., Olié, E., Passerieux, C., Costagliola, D., Etain, B., 2020. Trajectories of functioning in bipolar disorders: A longitudinal study in the FondaMental Advanced Centers of Expertise in Bipolar Disorders cohort. Aust N Z J Psychiatry, 4,867,420,945,796.
Grande, I., Berk, M., Birmaher, B., Vieta, E., 2016. Bipolar disorder. Lancet (London, England) 387, 1561–1572.
Guy, W., 1976. ECDEU assessment manual for psychopharmacology, Rev. 1976 ed. Rockville, Md.: U.S. Dept. of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute of Mental Health, Psychopharmacology Research Branch, Division of Extramural Research Programs.
Hayes, J. F., Marston, L., Walters, K., Geddes, J. R., King, M., Osborn, D. P., 2016. Lithium vs. valproate vs. olanzapine vs. quetiapine as maintenance monotherapy for bipolar disorder: a population-based UK cohort study using electronic health records. World psychiatry: official journal of the World Psychiatric Association (WPA) 15, 53–58.
Henry, C., Etain, B., Godin, O., Dargel, A.A., Azorin, J.M., Gard, S., Bellivier, F., Bougerol, T., Kahn, J.P., Passerieux, C., Aubin, V., Courtet, P., Leboyer, M., 2015. Bipolar patients referred to specialized services of care: Not resistant but impaired by sub-syndromal symptoms. Results from the FACE-BD cohort. Aust N Z J Psychiatry 49, 898–905.
Henry, C., Etain, B., Mathieu, F., Raust, A., Vibert, J.F., Scott, J., Leboyer, M., 2011. A French network of bipolar expert centres: a model to close the gap between evidence-based medicine and routine practice. Journal of affective disorders 131, 358–363.
Henry, C., Godin, O., Courtet, P., Azorin, J.M., Gard, S., Bellivier, F., Polosan, M., Kahn, J.P., Roux, P., Aubin, V., Costagliola, D., Leboyer, M., Etain, B., Laouamri, H., Souyris, K., Barteau, V., Geoffroy, P.A., Raust, A., Belzeaux, R., Loftus, J., Sportiche, S., Aouizerate, B., Olié, E., Ducasse, D., Viglianese, N., Lescalier, L., Cohen, R.F., Wajsbrot-Elgrabli, O., Garçon, S., M'Bailara, K., Hardy-Bayle, M.C., Grevin, I., 2017. Outcomes for bipolar patients assessed in the French expert center network: A 2-year follow-up observational study (FondaMental Advanced Centers of Expertise for Bipolar Disorder [FACE-BD]). Bipolar disorders.
Hozer, F., Sarrazin, S., Laidi, C., Favre, P., Pauling, M., Cannon, D., McDonald, C., Emsell, L., Mangin, J. F., Duchesnay, E., Bellani, M., Brambilla, P., Wessa, M., Linke, J., Polosan, M., Versace, A., Phillips, M. L., Delavest, M., Bellivier, F., Hamdani, N., d'Albis, M.A., Leboyer, M., Houenou, J., 2020. Lithium prevents grey matter atrophy in patients with bipolar disorder: an international multicenter study. Psychological medicine, 1–10.
Iniesta, R., Stahl, D., McGuffin, P., 2016. Machine learning, statistical learning and the future of biological research in psychiatry. Psychological medicine 46, 2455–2465.
Léda-Rêgo, G., Bezerra-Filho, S., Miranda-Scippa, Â., 2020. Functioning in euthymic patients with bipolar disorder: a systematic review and meta-analysis using the Functioning Assessment Short Test. Bipolar disorders.
López-Villarreal, A., Sánchez-Morla, E. M., Jiménez-López, E., Martínez-Vizcaíno, V., Aparicio, A. I., Mateo-Sotos, J., Rodriguez-Jimenez, R., Vieta, E., Santos, J. L., 2019. Progression of the functional deficit in a group of patients with bipolar disorder: a cluster analysis based on longitudinal data. European archives of psychiatry and clinical neuroscience.

Menke, A., 2018. Precision pharmacotherapy: psychiatry's future direction in preventing, diagnosing, and treating mental disorders. Pharmacogenomics and personalized medicine 11, 211–222.
Misdrahi, D., Verdoux, H., Llorca, P.M., Baylé, F.J., 2004. [Therapeutic adherence and schizophrenia: the interest of the validation of the French translation of Medication Adherence Rating Scale (MARS)]. L'Encephale 30, 409-410.
Montgomery, S. A., Asberg, M., 1979. A new depression scale designed to be sensitive to change. The British journal of psychiatry: the journal of mental science 134, 382–389.
Passos, I.C., Ballester, P.L., Barros, R.C., Librenza-Garcia, D., Mwangi, B., Birmaher, B., Brietzke, E., Hajek, T., Lopez Jaramillo, C., Mansur, R.B., Alda, M., Haarman, B.C.M., Isometsa, E., Lam, R.W., McIntyre, R.S., Minuzzi, L., Kessing, L.V., Yatham, L.N., Duffy, A., Kapczinski, F., 2019a. Machine learning and big data analytics in bipolar disorder: A position paper from the International Society for Bipolar Disorders Big Data Task Force. Bipolar disorders 21, 582–594.
Passos, I. C., Mwangi, B., Kapczinski, F., 2019b. Personalized Psychiatry: Big Data Analytics in Mental Health. Springer International Publishing, Cham.
Post, R.M., 2018. The New News about Lithium: An Underutilized Treatment in the United States. Neuropsychopharmacology 43, 1174–1179.
Rhee, T. G., Olfson, M., Nierenberg, A. A., Wilkinson, S. T., 2020. 20-Year Trends in the Pharmacologic Treatment of Bipolar Disorder by Psychiatrists in Outpatient Care Settings. The American journal of psychiatry, appiajp202019091000.
Rosa, A.R., Sánchez-Moreno, J., Martínez-Aran, A., Salamero, M., Torrent, C., Reinares, M., Comes, M., Colom, F., Van Riel, W., Ayuso-Mateos, J.L., Kapczinski, F., Vieta, E., 2007. Validity and reliability of the Functioning Assessment Short Test (FAST) in bipolar disorder. Clin Pract Epidemiol Ment Health 3, 5.
Rush, A., O'Neal, BL. , 1999. Patient Rated Inventory of Side Effects (PRISE): unpublished rating scale., University of Texas Southwestern Medical Center: Dallas.
Rush, A.J., Trivedi, M.H., Ibrahim, H.M., Carmody, T.J., Arnow, B., Klein, D.N., Markowitz, J.C., Ninan, P.T., Kornstein, S., Manber, R., Thase, M.E., Kocsis, J.H., Keller, M.B., 2003. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 54, 573–583.
Sarrazin, S., Cachia, A., Hozer, F., McDonald, C., Emsell, L., Cannon, D.M., Wessa, M., Linke, J., Versace, A., Hamdani, N., D'Albis, M.A., Delavest, M., Phillips, M.L., Brambilla, P., Bellani, M., Polosan, M., Favre, P., Leboyer, M., Mangin, J.F., Houenou, J., 2018. Neurodevelopmental subtypes of bipolar disorder are related to cortical folding patterns: An international multicenter study. Bipolar disorders 20, 721–732.
Sole, B., Bonnin, C.M., Jimenez, E., Torrent, C., Torres, I., Varo, C., Valls, E., Montejo, L., Gomez-Ocana, C., Tomioka, Y., Vieta, E., Martinez-Aran, A., Reinares, M., 2018. Heterogeneity of functional outcomes in patients with bipolar disorder: a cluster-analytic approach. Acta Psychiatr Scand 137, 516-527.
Tamminga, C.A., Pearlson, G.D., Stan, A.D., Gibbons, R.D., Padmanabhan, J., Keshavan, M., Clementz, B.A., 2017. Strategies for Advancing Disease Definition Using Biomarkers and Genetics: The Bipolar and Schizophrenia Network for Intermediate Phenotypes. Biological psychiatry. Cognitive neuroscience and neuroimaging 2, 20–27.
Tondo, L., Alda, M., Bauer, M., Bergink, V., Grof, P., Hajek, T., Lewitka, U., Licht, R.W., Manchia, M., Müller-Oerlinghausen, B., Nielsen, R.E., Selo, M., Simhandl, C., Baldessarini,

R.J., 2019. Clinical use of lithium salts: guide for users and prescribers. Int J Bipolar Disord 7, 16.
Vieta, E., Berk, M., Schulze, T. G., Carvalho, A. F., Suppes, T., Calabrese, J. R., Gao, K., Miskowiak, K. W., Grande, I., 2018. Bipolar disorders. Nat Rev Dis Primers 4, 18,008.
Ward, M.F., Wender, P. H., Reimherr, F. W., 1993. The Wender Utah Rating Scale: an aid in the retrospective diagnosis of childhood attention deficit hyperactivity disorder. The American journal of psychiatry 150, 885–890.
Wolfers, T., Doan, N.T., Kaufmann, T., Alnæs, D., Moberget, T., Agartz, I., Buitelaar, J.K., Ueland, T., Melle, I., Franke, B., Andreassen, O.A., Beckmann, C.F., Westlye, L.T., Marquand, A.F., 2018. Mapping the Heterogeneous Phenotype of Schizophrenia and Bipolar Disorder Using Normative Models. JAMA Psychiatry.
Yatham, L.N., Kennedy, S.H., Parikh, S.V., Schaffer, A., Bond, D.J., Frey, B.N., Sharma, V., Goldstein, B.I., Rej, S., Beaulieu, S., Alda, M., MacQueen, G., Milev, R.V., Ravindran, A., O'Donovan, C., McIntosh, D., Lam, R.W., Vazquez, G., Kapczinski, F., McIntyre, R.S., Kozicky, J., Kanba, S., Lafer, B., Suppes, T., Calabrese, J.R., Vieta, E., Malhi, G., Post, R.M., Berk, M., 2018. 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 disorders.
Young, R.C., Biggs, J. T., Ziegler, V. E., Meyer, D. A., 1978. A rating scale for mania: reliability, validity and sensitivity. The British journal of psychiatry: the journal of mental science 133, 429–435.

Table 1: Subtypes of bipolar disorder

Name of variable	Type 1	Type 2	NOS	Total population	p value
			n (%) or median [Q1; Q3]		
N	**865**	**764**	**166**	**1795**	
Age	39 [31; 48]	43 [33; 52]	44.5 [33.2; 54]	41 [32; 50]	8.6×10^{-7}
Gender					8.9×10^{-3}
Woman	492 (56.9)	490 (64.4)	102 (61.4)	1084 (60.5)	
Men	372 (43.1)	271 (35.6)	64 (38.6)	707 (39.5)	
Referring doctor					2.5×10^{-4}
Psychiatrist	676 (86.1)	551 (77.8)	119 (76.8)	1346 (81.7)	
General practionner	70 (8.9)	113 (16.0)	25 (16.1)	208 (12.6)	
Other	39 (5.0)	44 (6.2)	11 (7.1)	94 (5.7)	
Education level (years)	15 [12; 17]	14 [12; 16]	14 [11; 16]	14 [12; 17]	1.1×10^{-3}
Occupational status					4.7×10^{-1}
Active	429 (57.3)	413 (60.7)	88 (62.0)	930 (59.2)	
Unemployed	152 (20.3)	124 (18.2)	21 (14.8)	344 (21.9)	
Other	168 (22.4)	143 (21.0)	33 (23.2)	297 (18.9)	
Age at first psychotropic treatment V0	24 [19; 31]	27 [20; 37]	25.5 [20; 35.8]	25 [20; 34]	6.5×10^{-10}
Age at first episode	21 [17; 27]	22 [17; 30]	22 [17; 31.5]	21 [17; 29]	1.1×10^{-3}
Duration of illness before V0	15 [8; 23]	15 [9; 25]	16 [8; 27]	15 [8; 24]	3.37×10^{-2}
Number of treatments received during one-year follow-up	2 [1; 3]	2 [1; 3]	2 [1; 3]	2 [1; 3]	1.27×10^{-1}
Actual mood episode					
V0	216 (25.6)	249 (33.5)	53 (33.3)	518 (29.7)	1.5×10^{-3}
At least one hospitalization					
V0	303 (50.1)	157 (28.4)	36 (31.6)	496 (39.0)	9.2×10^{-14}
Complete remission between episodes V0	532 (73.1)	429 (66.2)	91 (70.0)	1052 (69.9)	2.1×10^{-2}
Number of patient who already attempts suicide					
Lifetime V0	303 (36.4)	305 (40.9)	79 (48.8)	687 (39.5)	7.1×10^{-3}

V0 = baseline visit
V12= one-year follow-up visit

Table 2: Clusters

2.1: Treatment received for clusters

Name of variable	Group 1 "heterogeneous"	Group 2 "lithium"	Group 3 "valproate"	Group 4 "lamotrigine"	p value
			n (%) or median [Q1; Q3]		
N	**1099**	**265**	**268**	**163**	
Number of treatments received during one year follow-up	2 [1; 3]	2 [1; 3]	2 [1; 3]	3 [2; 4]	1.41×10^{-5}
Treatment received, n (%)					
lithium	305 (27.8)	263 (99.2)	49 (18.3)	32 (19.6)	
valproate	267 (24.3)	22 (8.3)	237 (88.4)	15 (9.2)	
lamotrigine	142 (12.9)	62 (23.4)	14 (5.2)	153 (93.9)	
quetiapine	199 (18.1)	11 (4.2)	16 (6)	45 (27.6)	
aripiprazole	173 (15.7)	13 (4.9)	28 (10.4)	28 (17.2)	
venlafaxine	55 (5)	45 (17)	89 (33.2)	18 (11)	
escitalopram	115 (10,5)	10 (3,8)	7 (2,6)	9 (5,5)	
olanzapine	76 (6.9)	13 (4.9)	13 (4.9)	10 (6.1)	
alprazolam	95 (8.6)	6 (2.3)	6 (2.2)	4 (2.5)	
cyamemazine	75 (6.8)	16 (6)	10 (3.7)	6 (3.7)	
Days on treatment if received, median [Q1, Q3]					
lithium	306 [183; 365]	365 [352.5; 365]	365 [230; 365]	302 [44.75; 365]	
valproate	339 [182.5; 365]	61.5 [50; 115.75]	365 [365; 365]	302 [156.5; 365]	
lamotrigine	195 [89.75; 365]	365 [259.25; 365]	157.5 [89; 252.75]	365 [318; 365]	
quetiapine	331 [182.5; 365]	64 [23.5; 85.5]	246.5 [120; 330.5]	356 [217; 365]	
aripiprazole	352 [152; 365]	78 [56; 113]	349.5 [260.25; 365]	362 [150; 365]	
venlafaxine	188 [99.5 ; 365]	365 [261 ; 365]	365 [324 ; 365]	166 [43 ; 357.5]	
escitalopram	365 [192,5 ; 365]	97.5 [68,5 ; 240,25]	90 [57 ; 123,5]	62 [48 ; 217]	
olanzapine	365 [134.5 ; 365]	183 [31 ; 365]	204 [38 ; 365]	268 [204 ; 365]	
alprazolam	365 [204; 365]	129 [44.25; 165]	115.5 [78.5; 173.5]	269.5 [192.5; 317.75]	
cyamemazin	319 [178; 365]	365 [149; 365]	111.5 [19; 349.25]	179 [71.5; 320.25]	

2.2: Descriptive data for clusters

Name of variable	Group 1 "heterogeneous"	Group 2 "lithium"	Group 3 "valproate"	Group 4 "lamotrigine"	p value
			n (%) or median [Q1; Q3]		
N	**1099**	**265**	**268**	**163**	
Bipolar subtype:					
BP type 1	530 (48.2)	151 (57.0)	131 (48.9)	53 (32.5)	1.7×10^{-4}
BP type 2	459 (41.8)	96 (36.2)	116 (43.3)	93 (57.1)	
BP NOS	110 (10,0)	18 (6,8)	21 (7,8)	17 (10,4)	
Age	40 [31; 50]	40 [32; 51]	43 [34.8; 53]	41 [32; 50.5]	8.6×10^{-4}
Gender					1.4×10^{-2}
Woman	659 (60.2)	150 (56.6)	158 (59.0)	117 (71.8)	
Men	436 (39.8)	115 (43.4)	110 (41.0)	46 (28.2)	
Referring doctor					4.5×10^{-1}
Psychiatrist	814 (80.7)	203 (83.5)	202 (80.5)	127 (87.6)	
General practionner	137 (13.6)	25 (10.3)	34 (13.5)	12 (8.3)	
Other	58 (5.7)	15 (6.2)	15 (6.0)	6 (4.1)	
Education level (years)	14 [12; 16]	15 [12; 17]	15 [12; 17]	15 [12; 17]	2.6×10^{-3}
Occupational status					9.2×10^{-1}
Active	573 (59.6)	131 (57.7)	150 (60.0)	76 (57.1)	
Unemployed	182 (18.9)	46 (20.3)	41 (16.4)	28 (21.1)	
Other	206 (21.4)	50 (22.0)	59 (23.6)	29 (21.8)	

2.3: Clinical scales for clusters

Name of variable	Group 1 "heterogeneous"	Group 2 "lithium"	Group 3 "valproate"	Group 4 "lamotrigine"	p value
	n (%) or median [Q1; Q3]				
N	**1099**	**265**	**268**	**163**	
Self-administered					
QIDS-SR16					
V0	9 [5; 15]	8 [4; 13]	8.5 [4; 14]	11 [6; 16]	5.7×10^{-3}
V12†	7 [3 ; 11,8]	6 [3 ; 9]	6 [2,5 ; 11]	6 [4 ; 12]	5.7×10^{-3}
ALTMAN					
V0	2 [0; 4]	1 [0; 4]	2 [0; 5]	2 [0; 5]	1.8×10^{-2}
V12†	2 [0 ; 4,2]	2 [0 ; 4]	2 [0 ; 4]	2 [0 ; 4]	9.1×10^{-1}
PSQI (sleep disturbance)					
V0	6 [4; 10]	6 [4; 8]	6 [4; 9]	7.5 [5; 11]	1.7×10^{-5}
V12†	6 [4 ; 9]	5 [3 ; 8]	6 [3,5 ; 8]	6 [4 ; 9]	3.5×10^{-4}
PRISEM					
V0	12 [6; 19]	10 [5; 16]	10 [4.2; 18]	13 [6; 19.8]	3.5×10^{-2}
V12†	11 [6 ; 18]	9 [4 ; 15]	9 [3 ; 16]	10 [6 ; 16,8]	5.3×10^{-5}
MARS (Adherence to treatment)					
V0	7 [6; 8]	7 [6; 8]	7 [6; 8]	7 [6; 9]	1.7×10^{-1}
V12†	8 [7 ; 9]	8 [7 ; 9]	8 [7 ; 9]	8 [7 ; 9]	1.3×10^{-1}
WURS V0	29 [18; 45]	24 [12; 38]	33 [19.2; 46]	29 [16.8; 43.2]	8.2×10^{-6}
CTQ V0					
Score Physical abuse	5 [5; 7]	5 [5; 6]	5 [5; 7]	5 [5; 7]	1.4×10^{-1}
Score Emotional neglect	12 [9; 16]	11 [8; 14]	11 [9; 16]	12 [8; 16]	9.7×10^{-3}
Score Physical neglect	6 [5; 9]	6 [5; 8]	6 [5; 8]	6 [5; 8]	5.3×10^{-2}
Score Sexual abuse	5 [5; 6]	5 [5; 6]	5 [5; 6]	5 [5; 6]	8.2×10^{-1}
Score Emotional Abuse	9 [6; 14]	7 [5; 11]	9 [5; 14]	9 [6; 13.2]	5.4×10^{-4}
Clinician-administered					
MADRS					
V0	8 [3; 15]	6 [2; 15]	6 [2.8; 14]	9 [3; 19]	2.2×10^{-2}
V12†	6 [2 ; 13]	4 [1 ; 9]	4 [0,2 ; 10]	6 [2 ; 14]	3.6×10^{-3}
YMRS					
V0	0 [0; 3.8]	0 [0; 2]	1 [0; 4]	1 [0; 3]	6.3×10^{-1}
V12†	0 [0 ; 2]	0 [0 ; 2]	0 [0 ; 3]	0 [0 ; 2]	3.2×10^{-1}
FAST					
V0	19 [9; 32]	16 [7; 28]	16 [7; 29]	22 [12.8; 29]	2.8×10^{-2}
V12†	15 [6 ; 27]	10 [3 ; 22]	12 [3 ; 23]	15 [5 ; 25]	1.3×10^{-5}
CGI-Severity					
V0	3 [2; 5]	3 [1; 4]	3 [2; 4]	3 [2; 4]	7.7×10^{-3}
V12†	3 [2 ; 4]	2 [1 ; 4]	3 [1 ; 4]	3 [1 ; 4]	1.3×10^{-4}
GAF:					
V0					6.5×10^{-2}
100-71	331 (32.6)	100 (41.7)	97 (39.4)	44 (30.3)	
70-51	525 (51.8)	107 (44.6)	118 (48.0)	74 (51.0)	
50-0	158 (15.6)	33 (13.8)	31 (12.6)	27 (18.6)	
V12					1.7×10^{-2}
100-71	381 (45.2)	126 (59.2)	110 (50.5)	69 (53.9)	
70-51	383 (45.4)	73 (34.3)	92 (42.2)	50 (39.1)	
50-0	79 (9.4)	14 (6.6)	16 (7.3)	9 (7.0)	

V0 = baseline visit, V12= one-year follow-up visit, MADRS= Montgomery and Asberg Depression Rating Scale, YMRS= Young Mania Rating Scale, QIDS-SR16= Quick Inventory of Depressive Symptoms 16 items, AMRS= Altman Mania Rating Scale, GAF= Global Assessment of Functioning, FAST= Functioning Assessment Short Test, CGI-S= Clinical Global Impression—Severity scale, PSQI= Pittsburg Sleep Quality Index, CTQ= Childhood Trauma Questionnaire, WURS= Wender Utah Rating Scale, MARS= Medication Adherence Rating Scale, PRISE-M= Patient Rated Inventory of Side Effects inventory
†= adjusted for V0
‡,=Fisher test

2.4: Clinical evaluation for clusters

Name of variable	Group 1 "heterogeneous"	Group 2 "lithium"	Group 3 "valproate"	Group 4 "lamotrigine"	p value
	n (%) or median [Q1; Q3]				
N	**1099**	**265**	**268**	**163**	
First episode V0					
Age at first episode	21 [17; 28]	22 [18; 30]	23 [18; 31]	22 [18; 29]	1.1×10^{-3}
Age at first psychotropic treatment V0	25 [19; 33]	25 [20; 33.8]	27 [21; 36]	27 [20; 37]	1.2×10^{-2}
Cyclothymia before first episode	131 (13.9)	23 (10.0)	43 (18.4)	24 (16.6)	5.8×10^{-2}
Duration of illness before V0	15 [8; 24]	15 [9; 23]	15 [8; 25]	14 [8; 22]	3.17×10^{-1}
Actual mood episode					
V0	322 (30.0)	68 (26.5)	66 (25.5)	62 (39.2)	1.5×10^{-2}
V12	187 (21,8)	31 (14,4)	45 (20,1)	33 (24,6)	6.9×10^{-2}
Last Year					
At least one MDE					
V0	636 (63.8)	146 (61.6)	140 (59.1)	111 (76.6)	4.6×10^{-3}
V12	274 (33,3)	50 (24,5)	58 (27,4)	42 (33,1)	5.3×10^{-2}
At least one MDE with psychotic features					
V0 ‡	34 (3.4)	9 (3.8)	4 (1.7)	0 (0.0)	4.2×10^{-2}
V12 ‡	13 (1,6)	6 (3,0)	2 (1,0)	1 (0,8)	3.9×10^{-1}
At least one hypomanic episode					
V0	329 (34.5)	58 (25.8)	70 (31.2)	50 (34.5)	8.4×10^{-2}
V12	123 (15,5)	37 (18,6)	35 (16,5)	15 (11,9)	4.4×10^{-1}
At least one manic episode					
V0	133 (13.5)	53 (22.1)	33 (14.6)	10 (6.7)	2.2×10^{-4}
V12 ‡	36 (4,5)	10 (5,0)	4 (1,9)	2 (1,6)	1.5×10^{-1}
At least one manic episode with psychotic features					
V0	75 (7.4)	37 (15.2)	22 (9.4)	8 (5.3)	5.2×10^{-4}
V12 ‡	22 (2,7)	6 (3,0)	1 (0,5)	0 (0,0)	4.2×10^{-2}
At least one mixed episode					
V0	80 (8.4)	15 (6.4)	19 (8.7)	8 (5.6)	5.3×10^{-1}
V12	45 (5,6)	4 (2,0)	11 (5,3)	4 (3,2)	1.5×10^{-1}
At least one mixed episode with psychotic features					
V0 ‡	15 (1.5)	5 (2.1)	4 (1.8)	2 (1.4)	8.8×10^{-1}
V12 ‡	8 (1,0)	0 (0,0)	3 (1,4)	0 (0,0)	3.5×10^{-1}
At least one hospitalization					
V0	305 (38.7)	89 (47.1)	61 (35.1)	41 (34.2)	5.6×10^{-2}
V12	109 (14,4)	21 (11,6)	18 (9,4)	18 (15,9)	2.2×10^{-1}
Lifetime V0					
At least one MDE	863 (95.9)	210 (94.6)	208 (96.7)	131 (99.2)	1.6×10^{-1}
At least one MDE with psychotic features	97 (9.7)	24 (9.9)	25 (10.4)	5 (3.4)	7.9×10^{-2}
At least one hypomanic episode	609 (76.3)	149 (78.8)	136 (75.1)	98 (83.1)	3.4×10^{-1}
At least one manic episode	460 (46.0)	128 (53.1)	104 (43.3)	41 (26.5)	2.7×10^{-6}
At least one manic episode with psychotic features	322 (31.6)	97 (39.8)	58 (23.7)	18 (11.7)	4.7×10^{-9}
At least one mixed episode	177 (19.2)	45 (19.8)	43 (20.0)	23 (16.8)	8.8×10^{-1}
At least one mixed episode with psychotic features	66 (6.5)	16 (6.7)	15 (6.3)	4 (2.7)	3.2×10^{-1}
Rapid cycling V0	130 (13.6)	38 (16.8)	29 (12.8)	41 (28.5)	4.1×10^{-5}
Complete remission between episodes V0	638 (69.3)	164 (71.9)	162 (73.6)	88 (64.2)	2.4×10^{-1}
Seasonal pattern V0	314 (32.8)	77 (33.3)	80 (35.4)	50 (36.0)	9.5×10^{-1}
Number of patient who already attempts suicide Lifetime V0	439 (41.4)	93 (36.6)	89 (33.5)	66 (41.2)	7.9×10^{-2}

Last year V12 ‡	22 (2.5)	2 (0.9)	4 (1.7)	5 (3.8)	3.1×10^{-1}

V0 = baseline visit, V12= one-year follow-up visit
†= adjusted for V0
‡,=Fisher test

2.5: Comorbidities for clusters

Name of variable	**Group 1** "heterogeneous"	**Group 2** "lithium"	**Group 3** "valproate"	**Group 4** "lamotrigine"	**p value**
N	**n (%) or median [Q1; Q3]** **1099**	**265**	**268**	**163**	
Psychiatric comorbidities V0					
Current smoking	473 (44.6)	107 (42.1)	111 (42.7)	62 (39.0)	5.5×10^{-1}
Anxiety disorders	459 (45.8)	85 (35.4)	109 (44.5)	81 (51.6)	7.8×10^{-3}
Substance use disorders:	349 (34.5)	72 (30.0)	74 (29.7)	49 (31.6)	3.5×10^{-1}
Alcool use disorder	243 (24.0)	42 (17.5)	52 (20.9)	40 (25.8)	1.1×10^{-1}
Cannabis use disorder	186 (18.4)	46 (19.2)	36 (14.5)	17 (11.0)	6.7×10^{-2}
Feeding and Eating Disorder	189 (18.7)	32 (13.2)	41 (16.6)	47 (30.7)	1.7×10^{-4}
Somatic Symptom and Related Disorder	39 (3.9)	5 (2.1)	13 (5.2)	8 (5.2)	2.8×10^{-1}
Somatic comorbidities V0					
Neurologic disease:					
Migraine	208 (20.0)	50 (19.7)	51 (20.3)	40 (25.8)	4.0×10^{-1}
Epilepsy ‡	17 (1.6)	6 (2.4)	3 (1.2)	2 (1.3)	8.0×10^{-1}
Head trauma	132 (12.8)	38 (15.2)	33 (13.3)	23 (15.0)	7.1×10^{-1}
Cardiovascular disease:					
Arrythmia	33 (3.2)	10 (4.0)	9 (3.6)	8 (5.3)	6.2×10^{-1}
Hypertension	62 (6.0)	23 (9.1)	29 (11.6)	14 (9.2)	1.2×10^{-2}
Metabolic and endocrine disorders:					
Diabetes ‡	29 (2.8)	4 (1.6)	12 (4.9)	5 (3.2)	1.8×10^{-1}
Thyroid disease	99 (9.6)	36 (14.3)	33 (13.2)	27 (17.4)	9.6×10^{-3}
Dyslipidemia	185 (18.0)	30 (12.0)	43 (17.3)	21 (13.6)	9.1×10^{-2}
Skin disease:					
Acne	134 (13.0)	35 (13.9)	36 (14.4)	28 (18.2)	3.8×10^{-1}
Cutaneous adverse drug reactions	86 (8.3)	16 (6.4)	11 (4.4)	13 (8.4)	1.6×10^{-1}
Hair loss	148 (14.3)	29 (11.5)	52 (20.8)	20 (13.0)	1.9×10^{-2}
Urinary tract disorder:					
Nephropathy ‡	33 (3.2)	3 (1.2)	4 (1.6)	1 (0.6)	9.3×10^{-2}
Gastro-intestinal disorder:					
Drug-induced hepatitis ‡	14 (1.4)	3 (1.2)	3 (1.2)	2 (1.3)	1.0
BMI					
V0	25.2 [22.5; 28.7]	24.9 [22.2; 27.9]	24.5 [22.4; 28.4]	24.7 [22.4; 29.0]	4.8×10^{-1}
V12	25.4 [22,6 ; 29,1]	24.5 [21,9 ; 28,2]	24.8 [22,6 ; 28,4]	24.7 [22,1 ; 28,5]	4.1×10^{-5}

V0 = baseline visit, V12= one-year follow-up visit
†= adjusted for V0
‡,=Fisher test

Figure 1 : cluster dendrogram :

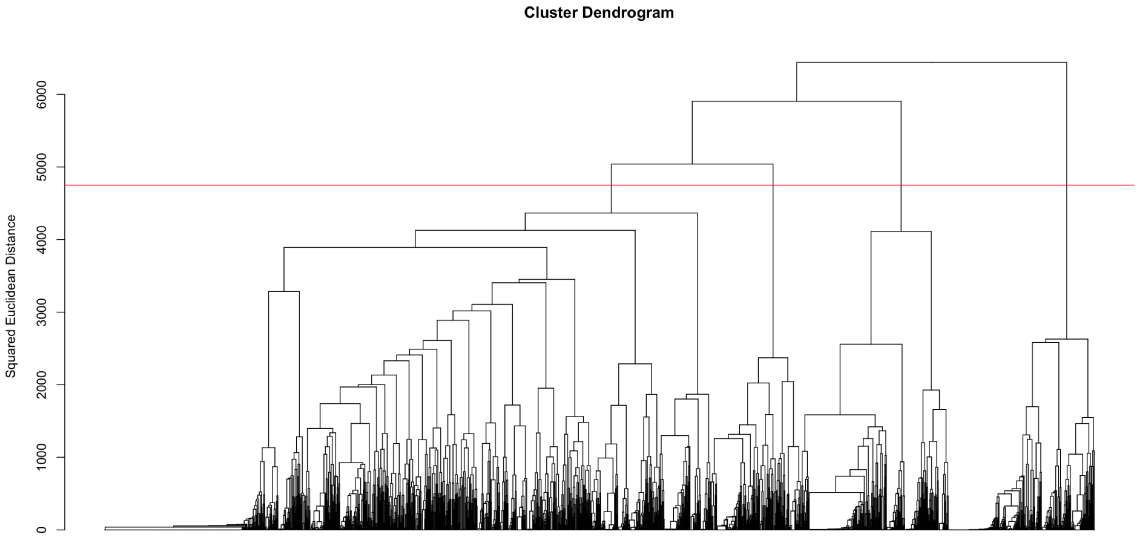


Figure 1 : Days on treatment for each active ingredient present in the dataset, during the first year following the inclusion, were used to classify patients. The distance metric was the squared Euclidean distance and Ward's agglomerative method was used. The number of clusters was determined using the visual analysis of the dendrogram associated to the hierarchical agglomerative clustering and the interpretability of the groups.

Figure 1: Days on treatment for each active ingredient present in the dataset, during the first year following the inclusion, were used to classify patient. The division in four groups was suggested by Calinski function on RStudio and is represented by the red line on the dendrogram cluster.

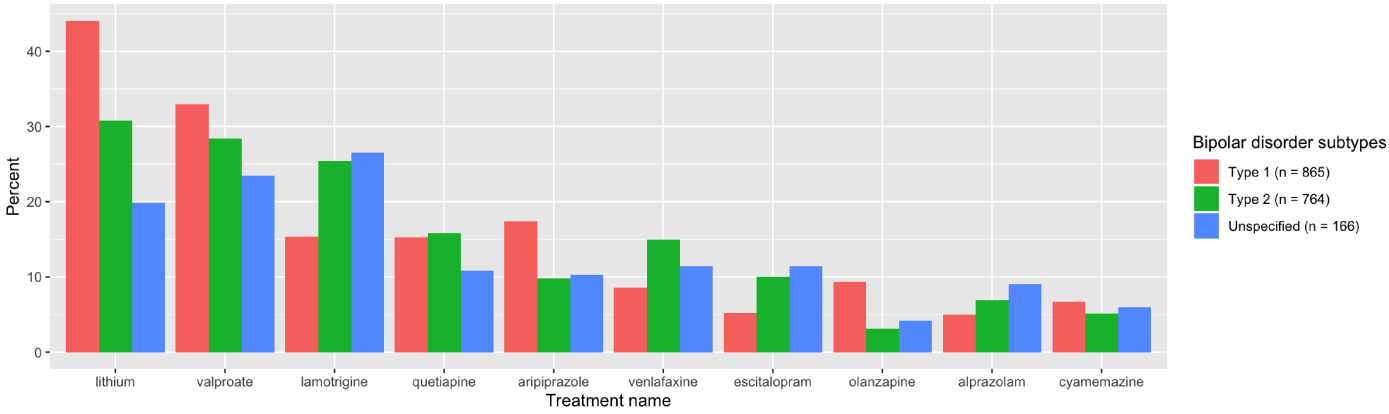
Figure 1 :

Treatment received during one-year follow-up period (bipolar disorder subtypes): The treatments received during one year in bipolar disorder 1 were mainly: lithium (44 %), valproate (32.9%), lamotrigine (15.4), quetiapine (15.3%) and aripiprazole (17.3%). As for bipolar disorder 2, they received lithium (30.8%), valproate (28.4%), lamotrigine (25.4), quetiapine (15.8%) and aripiprazole (9.8%). Finally regarding, bipolar disorder NOS, patients received mainly lithium (19.9 %), valproate (23.5%), lamotrigine (26.5), quetiapine (10.8%) and aripiprazole (10.2%).

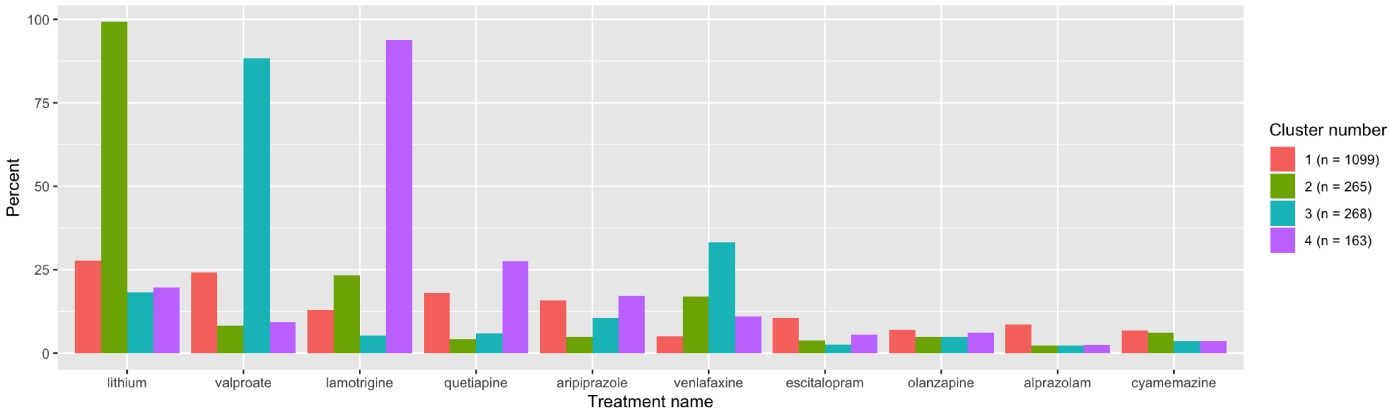
Treatment received during one-year follow-up period (treatment clusters): The treatments received during one year in group 1 were mainly: lithium (27.8%), valproate (24.3%), quetiapine (18.1%) and aripiprazole (15.7%). The treatments received during one year in group 2 were mainly: lithium (99.2%), lamotrigine (23.4%) and venlafaxine (17%). The main treatments in group 3 were mainly valproate (88.4%), venlafaxine (33.2%) and lithium (18.3%). The treatments received during one year in group 4 were mainly: lamotrigine (93.9%), quetiapine (27.6%) lithium (19.6%) and aripiprazole (17.2%).

Figure 2: Treatment received during one-year follow-up period:

2.1: Treatment received during one-year follow-up period : bipolar disorder subtypes



2.2: Treatment received during one-year follow-up period: treatment clusters



2.1: The treatments received during one year in bipolar disorder 1 were mainly: lithium (44 %), valproate (32·9%), lamotrigine (15·4), quetiapine (15·3%) and aripiprazole (17·3%). As for bipolar disorder 2, they received lithium (30·8%), valproate (28·4%), lamotrigine (25·4), quetiapine (15·8%) and aripiprazole (9·8%). Finally regarding, bipolar disorder NOS, patients received mainly lithium (19·9 %), valproate (23·5%), lamotrigine (26·5), quetiapine (10·8%) and aripiprazole (10·2%).

2.2: The treatments received during one year in group 1 were mainly: lithium (27·8%), valproate (24·3%), quetiapine (18·1%) and aripiprazole (15·7%). The treatments received during one year in group 2 were mainly: lithium (99·2%), lamotrigine (23·4%) and venlafaxine (17%). The main treatments in group 3 were mainly valproate (88·4%), venlafaxine (33·2%) and lithium (18·3%). The treatments received during one year in group 4 were mainly: lamotrigine (93·9%), quetiapine (27·6%) lithium (19·6%) and aripiprazole (17·2%).