Lithium vs. valproate vs. olanzapine vs. quetiapine as maintenance monotherapy for bipolar disorder: a population-based UK cohort study using electronic health records

Joseph F. Hayes[1], Louise Marston[2], Kate Walters[2], John R. Geddes[3], Michael King[1], David P.J. Osborn[1]

[1]Division of Psychiatry, University College London, London W1T 7NF, UK; [2]Department of Primary Care and Population Health, University College London, London, UK; [3]Department of Psychiatry, University of Oxford, Oxford, UK

It is unclear which maintenance treatment for bipolar disorder is superior in clinical practice. Randomized controlled head-to-head trials of available drugs either do not exist or are inconclusive. We aimed to compare rates of monotherapy treatment failure in individuals prescribed lithium, valproate, olanzapine or quetiapine by a population-based cohort study using electronic health records. 5,089 patients with bipolar disorder were prescribed lithium (N=1,505), valproate (N=1,173) olanzapine (N=1,366) or quetiapine (N=1,075) as monotherapy. Treatment failure was defined as time to stopping medication or add-on of another mood stabilizer, antipsychotic, antidepressant or benzodiazepine. In unadjusted analyses, the duration of successful monotherapy was longest in individuals treated with lithium. Treatment failure had occurred in 75% of those prescribed lithium by 2.05 years (95% CI: 1.63-2.51), compared to 0.76 years (95% CI: 0.64-0.84) for those prescribed quetiapine, 0.98 years (95% CI: 0.84-1.18) for those prescribed valproate, and 1.13 years for those prescribed olanzapine (95% CI: 1.00-1.31). Lithium's superiority remained in a propensity score matched analysis; when treatment failure was defined as stopping medication or add-on of a mood stabilizer or antipsychotic; and when treatment failure was restricted to more than three months after commencing the study drug. Lithium appears to be more successful as monotherapy maintenance treatment than valproate, olanzapine or quetiapine. Lithium is often avoided because of its side effect profile, but alternative treatments may reduce the time to being prescribed more than one drug, with potential additive side effects of these treatments.

Key words: Bipolar disorder, maintenance treatment, lithium, valproate, olanzapine, quetiapine, electronic health records, cohort study

Bipolar disorder is a lifelong recurrent illness with high rates of hospitalization, suicide and comorbidity[1]. It is the sixth most common cause of disability in the world, responsible for the loss of more disability-adjusted life years than all forms of cancer or major neurological conditions such as epilepsy and Alzheimer's disease[2]. Long-term drug treatment is often required to prevent relapse or recurrence. Even with treatment, the proportion of people who remain in remission is low[3].

A number of drug treatments are recommended for maintenance in bipolar disorder. In the UK, the most commonly used medications are lithium, valproate, olanzapine and quetiapine[4]. This reflects previous National Institute for Health and Care Excellence (NICE) guidance on first-line monotherapy maintenance treatment, which suggested equivalence of these drugs[5]. Globally, there is a range of prescribing advice, which includes additionally lamotrigine, carbamazepine, oxcarbazepine, aripiprazole and other second generation antipsychotics[6-8]. Recent meta-analyses and network meta-analyses have highlighted the superiority of lithium[9,10], and these results have contributed to the change in NICE guidance in September 2014, where lithium is presented as first line[11]. However, no randomized controlled trial (RCT) has conclusively proved the benefit of lithium over other drugs, and there are no trials that compare valproate vs. olanzapine, valproate vs. quetiapine or olanzapine vs. quetiapine directly.

The applicability of RCT results to people with bipolar disorder in the real world may be limited by the exclusion criteria adopted in those trials, and by diagnostic heterogeneity, diagnosis or treatment rejection, and complex presentations of the illness occurring over the life course[12,13]. These concerns have been raised when considering RCTs in other areas of medicine: applying their results to managing a lifelong illness of unpredictable course is not straightforward[14,15]. Necessary trials are also costly and difficult to run for sufficient periods in relation to the time course of bipolar disorder[16]. Electronic health records offer an opportunity to augment RCT findings with head-to-head comparison studies which include large numbers of patients, representative of real world clinical practice, and long follow-up periods.

Using data from a large UK primary care database, The Health Improvement Network (THIN), we aimed to compare rates of stopping medication or add-on of another psychotropic drug in individuals prescribed lithium, valproate, olanzapine or quetiapine as maintenance monotherapy for bipolar disorder. This outcome represents a combination of both effectiveness and tolerability of the study medication, and is similar to that used in many RCTs of maintenance treatment for bipolar disorder[9,10].

METHODS

Study design and setting

This was a prospective study of primary care data collected between January 1, 1995 and December 31, 2013. The scheme for THIN was approved by the National Health Service South-

East Multicentre Research Ethics Committee, and scientific approval for this study was obtained from Cegedim Strategic Data (CSD) Medical Research's Scientific Review Committee.

THIN is a UK primary care database that contains anonymized patient information from routine clinical consultations[17]. General practitioners (GPs) use Read codes, a hierarchical coding system, to record information in THIN[18]. These codes include diagnoses (which map onto ICD-10 codes), symptoms, examination findings, referrals, test results and information from hospital specialists, creating a longitudinal record for each patient[19]. In the UK, GPs are responsible for issuing all drug prescriptions if treatment is ongoing, following advice from a psychiatrist, and this information is also available[20].

At the time this cohort was extracted, THIN contained records for over 11 million people[17]. Patients in the database have been shown to be broadly representative of the UK population, and GPs contributing data have been shown to be representative in terms of consultation and prescribing statistics[21,22]. Approximately 98% of the UK population is registered with a GP practice[23]. The incidence rate of bipolar disorder in THIN has been shown to be similar to other European cohorts[24], and validity of severe mental illness diagnoses held in primary care has been established[25].

Participants

Patients with a diagnosis of bipolar disorder were included if they had at least one 28-day prescription of lithium, valproate, olanzapine or quetiapine after January 1, 1995, or after the date at which the GP practice met quality assurance criteria for data entry (based on computer usage and mortality recording rates)[26,27]. Patients were excluded if they received a diagnosis of schizophrenia at any time. They were also excluded if they were prescribed another of the study drugs, or any other mood stabilizer, antipsychotic, antidepressant or benzodiazepine at the start of follow-up, or in the month before this. The cohort was therefore one in which the intention was to treat with lithium, valproate, olanzapine or quetiapine monotherapy. Patients were censored at date of death, leaving the GP practice or the end of the study period (December 31, 2013).

Main outcome

Patients were followed up until they stopped the study drug, or had a mood stabilizer, an antipsychotic, an antidepressant or a benzodiazepine added to their treatment regimen. Date of first prescription was taken as the start of exposure time. The end of the prescription was calculated from the prescription length and prescribing instructions coded by the GP.

Patients were considered to have a period of continuous prescribing if another prescription for the same drug was issued within three months of the predicted end date. If this did not occur, the date of stopping the study drug was the end date of the final prescription.

Observed pre-treatment variables for propensity score estimation

Socio-demographic, psychiatric and physical health characteristics at baseline were extracted from each patient's electronic health record. Psychiatric and physical health problems were considered present if referenced in the patient notes. If a patient had multiple entries of the same (or similar) Read codes, the start date of the condition was taken as the earliest date of entry.

A propensity score (PS) for each individual was estimated using variables defined *a priori*, based on existing research[28,29]. The PS attempts to account for all of the covariates that predict receiving a particular study drug[29,30]. The PS was then checked by comparison of covariate balance across treatments, within strata. The included variables were: gender; age at start of treatment with the study drug; year of entry to the cohort; ethnicity (grouped as White, Black, Asian, mixed, other, with missing values coded as White); physical health history at baseline (ischemic heart disease, myocardial infarction, cerebrovascular event, hypertension, renal disease, thyroid disease, liver disease, type 2 diabetes mellitus, epilepsy, history of alcohol dependence, history of illicit drug use); smoking status (grouped as never-smoker, ex-smoker, current smoker); body mass index (BMI) (grouped as healthy weight, overweight (BMI 25 to 30), obese (BMI over 30)); mental health history at baseline (history of anxiety symptoms, hypomania as most proximal diagnosis code, history of depressive symptoms, sleep disturbance, previous treatment with the study drug before baseline, incident diagnosis of bipolar disorder); and clustering by GP practice. These variables were selected because they represent factors influencing prescribing choice (such as risk factors for adverse effects with a particular study medication)[11].

Although PS estimation cannot remove all bias, it has been postulated to also reduce confounding from unmeasured variables, because of their association with measured covariates[31,32]. Therefore in this study, for a given PS, exposure to lithium, valproate, olanzapine or quetiapine is presumed to have been at random[33].

Statistical analysis

Cox regression analyses were conducted comparing the rates of stopping the study drug or add-on of another psychotropic medication in the four treatment groups. Analyses were adjusted for gender, age, ethnicity and calendar year. Time to treatment failure was summarized by Kaplan-Meier curves. The proportional hazards model was tested formally with analysis of Schoenfeld residuals[34].

Table 1 Characteristics of patients with bipolar disorder prescribed lithium, valproate, olanzapine or quetiapine monotherapy

	Lithium	Valproate	Olanzapine	Quetiapine
Patients	1,505	1,173	1,336	1,075
Female, N (%)	860 (57.1)	631 (53.8)	733 (54.9)	735 (68.4)
Age at entry to the cohort, median (IQR)	44.9 (35.4-58.7)	41.6 (31.4-53.7)	40.9 (31.9-52.7)	38.5 (29.3-49.8)
Total years of follow-up, median (IQR)	4.2 (1.5-8.6)	3.0 (1.1-6.3)	3.6 (1.4-6.9)	2.1 (0.9-3.9)
GP practice contacts per year of follow-up, median (IQR)	12.1 (7.1-19.7)	14.8 (8.7-23.7)	14.3 (8.8-24.6)	17.9 (11.8-26.9)
Non-White ethnic background, N (%)	44 (2.9)	50 (4.3)	65 (4.9)	35 (3.3)
Health at baseline, N (%)				
Ischemic heart disease, myocardial infarction, cerebrovascular event history	76 (5.0)	80 (6.8)	58 (4.3)	41 (3.8)
Renal disease history	51 (3.4)	36 (3.1)	33 (2.5)	42 (3.9)
Thyroid disease history	161 (10.7)	89 (7.6)	89 (6.7)	75 (7.0)
Diabetes	77 (5.1)	87 (7.4)	42 (3.1)	71 (6.6)
Epilepsy	29 (1.9)	82 (7.0)	37 (2.8)	34 (3.2)
Obesity (BMI>30)	617 (41.0)	488 (41.6)	482 (36.1)	467 (43.4)
Previous anxiety symptoms	98 (6.5)	102 (8.7)	133 (10.0)	154 (14.3)
Previous alcohol dependence	7 (0.5)	3 (0.3)	12 (0.9)	7 (0.6)
Current smoker	518 (34.4)	462 (39.4)	571 (42.7)	425 (39.5)
Bipolar disorder characteristics at baseline, N (%)				
Incident diagnosis	318 (19.6)	396 (34.0)	543 (41.7)	416 (40.8)
Previous depressive episode	845 (56.1)	701 (59.8)	826 (61.8)	788 (73.3)
Hypomania as most recent diagnosis	234 (15.5)	154 (13.1)	238 (17.8)	125 (11.6)
Previous record of taking study drug	936 (62.2)	507 (43.2)	463 (34.7)	328 (30.5)

GP – general practitioner, IQR – interquartile range, BMI – body mass index

The PS was calculated using multinomial logistic regression using the covariates described as independent variables, with drug treatment as the dependent variable. The PS was then used as a linear term in a Cox regression analysis that also included age and calendar year[35]. This model was shown to be superior to stratifying on PS using Akaike information criterion and Bayesian information criterion[36], and was a more efficient use of data than PS matching, because it uses all patients.

Analysis using PS matching was then completed. Although matched analyses may include a non-representative sample of patients receiving treatment, they may provide a more valid estimate of treatment effect as they compare patients with similar observed characteristics[35,37]. Pairwise matching was performed for each patient in the valproate, olanzapine and quetiapine groups with individuals in the lithium treated group. Patients were matched on a one-to-one basis if their PS was within 0.01 of each other; all other patients were dropped from the analysis.

Supplementary analyses excluding benzodiazepine and antidepressant add-on as a source of treatment failure were carried out. A supplementary analysis excluding patients who stopped the study drug or had psychotropic medication

Table 2 Rates of treatment failure by drug

	N events	Person years at risk	Rate per 100 person years at risk	Treatment failure, hazard ratio (95% CI)			
				Unadjusted	Model 1	Model 2	Model 3
Lithium	1,151	1,570	73.3 (65.5-81.8)	1	1	1	1
Valproate	909	777	116.9 (102.9-132.4)	1.25 (1.14- 1.37)	1.22 (1.11-1.34)	1.19 (1.09-1.31)	1.20 (1.10-1.32)
Olanzapine	977	893	109.4 (96.3-123.7)	1.19 (1.08-1.30)	1.19 (1.08-1.30)	1.16 (1.05-1.28)	1.17 (1.07-1.29)
Quetiapine	814	457	177.9 (157.9-199.8)	1.48 (1.35-1.62)	1.31 (1.19-1.44)	1.30 (1.18-1.44)	1.32 (1.20-1.45)

Model 1: adjusted for clustering by primary general practitioner (GP) practice, age, gender and calendar year

Model 2: adjusted for propensity score, clustering by GP practice, age and calendar year

Model 3: propensity score matched (pairwise matching with lithium) adjusted for clustering by GP practice, age and calendar year
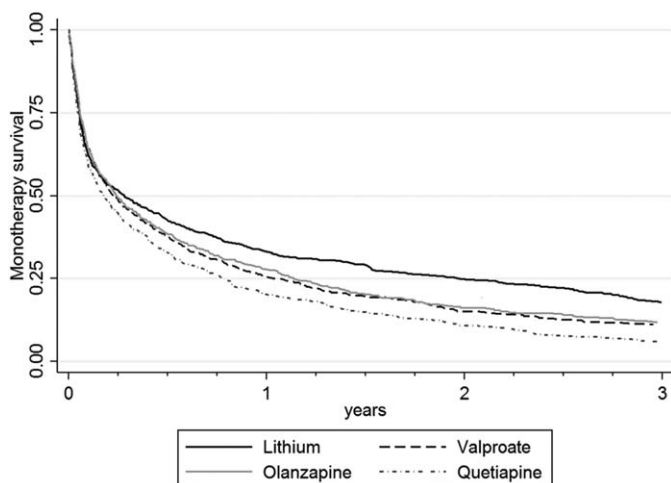
Figure 1 Time to treatment failure (stopped treatment or add-on of mood stabilizer, antipsychotic, antidepressant or benzodiazepine) (unadjusted)

added-on within the first three months of follow-up was also performed.

All analyses were completed using STATA 13[38].

RESULTS

A total of 14,396 individuals had a diagnosis of bipolar disorder. Of these, 5,089 were prescribed monotherapy with one of the study drugs at the start of cohort follow-up: lithium was prescribed to 1,505 people, valproate to 1,173, olanzapine to 1,366 and quetiapine to 1,075 people. Individuals prescribed lithium tended to be older than other groups, with more years of follow-up data and fewer GP practice contacts during this period. They were less likely to have a previous record of depression in their notes and less likely to be an incident case (Table 1).

In unadjusted analyses, the overall rate of treatment failure was increased for valproate, olanzapine and quetiapine when compared to lithium (Table 2). Treatment failure had occurred in 75% of those prescribed lithium by 2.05 years (95% CI: 1.63-2.51), compared to 0.76 years (95% CI: 0.64-0.84) for those prescribed quetiapine, 0.98 years (95% CI: 0.84-1.18) for those prescribed valproate, and 1.13 years for those prescribed olanzapine (95% CI: 1.00-1.31). The median time to treatment fail-

ure in the lithium monotherapy group was 0.28 years (95% CI: 0.23-0.35), compared to 0.17 years (95% CI: 0.14-0.21) in the quetiapine group, 0.22 years (95% CI: 0.19-0.27) in the valproate group, and 0.24 years (95% CI: 0.21-0.28) in the olanzapine group. The differences between treatments became more apparent the longer the duration of treatment (Figure 1).

Lithium's superiority remained after adjustment for clustering by GP practice, age, gender, calendar year, and ethnicity. It also remained after adjusting for PS, age and calendar year, and after matching by PS (Table 2), with olanzapine having the least elevated hazard ratio (HR) (1.16, 95% CI: 1.05-1.28). Compared to olanzapine, quetiapine had an increased rate of monotherapy failure (HR 1.12, 95% CI: 1.02-1.23) in the PS adjusted model. Compared to valproate, olanzapine and quetiapine had similar rates of treatment failure (HR 0.97, 95% CI: 0.89-1.06 and HR 1.09, 95% CI: 0.99-1.19, respectively). The proportional hazards assumption held for all analyses. Before pairwise matching, PS scores were most different for lithium (median 0.45, interquartile range, IQR 0.25-0.61) and quetiapine (median 0.14, IQR 0.08-0.25). After matching, the median PS was 0.21 (IQR 0.13-0.30) for lithium and 0.14 for quetiapine (IQR 0.08-0.25).

Individuals prescribed lithium or valproate were more likely to require antipsychotic add-on (19.53% and 18.41%, respectively) than those prescribed olanzapine or quetiapine monotherapy (10.25% and 9.02%, respectively). Conversely, individuals prescribed olanzapine and quetiapine were more likely to require mood stabilizer add-on (14.07% and 12.56%, respectively) compared to lithium and valproate (6.71% and 5.20%, respectively).

Supplementary analyses produced similar results to the primary analyses. If treatment failure was restricted to stopping the study drug or add-on of a mood stabilizer or antipsychotic medication, PS adjusted HRs were elevated for all drugs compared to lithium (Table 3). The same was true if patients failing in the first three months of follow-up were excluded from the analysis (Table 3, Figure 2).

DISCUSSION

As far as we are aware, this study represents the only head-to-head comparison of the four most common maintenance treatments for bipolar disorder, and has the longest follow-up

Table 3 Supplementary analyses using propensity score adjusted model

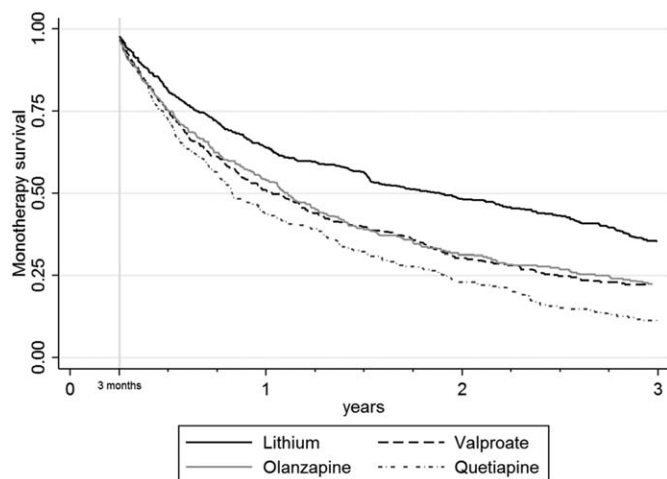|	Treatment failure, hazard ratio (95% CI)		
	Excluding benzodiazepine add-on	Excluding benzodiazepine and antidepressant add-on	Excluding failures in the first three months of treatment
Lithium	1	1	1
Valproate	1.25 (1.14-1.37)	1.18 (1.08-1.29)	1.22 (1.06-1.40)
Olanzapine	1.10 (1.00-1.22)	1.17 (1.07-1.28)	1.26 (1.09-1.45)
Quetiapine	1.25 (1.13-1.38)	1.13 (1.03-1.25)	1.20 (1.04-1.40)

Figure 2 Time to treatment failure (excluding failures in the first three months of treatment) (unadjusted)

and largest cohort of any direct comparison of treatments for bipolar disorder. RCTs making these comparisons do not exist and are unlikely to be conducted.

The overall rate of treatment failure (represented by stopping index medication or requiring add-on of a mood stabilizer, antipsychotic, antidepressant or benzodiazepine) was increased for valproate, olanzapine and quetiapine when compared to lithium. This was also true if failures within the first three months were excluded (i.e., once the patient had been stabilized on the prescribed drug). These results suggest that monotherapy with lithium may be more successful than the other recommended drugs. The rate of treatment failure was also elevated for quetiapine compared to olanzapine, while it was not possible to separate the other drugs from each other.

The use of contemporaneous, representative medical records avoided the risk of potential biases relating to selection into the study. Information bias should partially have been avoided by the use of prescribing data as exposure: in the UK, GPs are responsible for all ongoing prescribing within the national health system[20], which is detailed and well recorded in THIN. However, exposure to the study drug was approximated through prescriptions issued to patients, and may not reflect how the patient used the medication. Poor adherence to prescribed drug regimens is a problem with all medications, and this is particularly true if side effects are unpleasant, as can be the case with all of the study drugs[39,40]. In this study, stopping the drug will be reflected in the outcome, but erratic adherence cannot be detected. It is possible that erratic adherence is more likely for drugs other than lithium (as this is more closely monitored through regular blood tests). This may have contributed to lithium's perceived superiority, but we found that patients prescribed lithium had *fewer* GP contacts, and other longitudinal cohort studies have not shown differential adherence[40].

Treatment failure was defined as stopping the study drug or add-on of any mood stabilizer, antipsychotic, antidepressant or benzodiazepine. It is likely that addition of a mood stabiliz-

er or antipsychotic represents more serious treatment failure than addition of an antidepressant (which would only occur during a depressive relapse) or a benzodiazepine (which may be used short term to avoid a relapse). A supplementary analysis excluding addition of these drugs had similar results. It may be the case that both of these outcomes fail to capture what is important to patients in terms of relapse, recurrence, functioning and quality of life. However, through examining monotherapy treatment failure, we believe we have described a proxy for these important outcomes which captures both tolerability and effectiveness and highlights very common need for adjunctive drug treatments. This outcome has also been used in a number of RCTs of maintenance drug treatment for bipolar disorder and therefore comparison with these results is possible. For example, the largest trial of lithium vs. valproate treatment had a primary outcome of "time to new intervention for an emerging mood episode"[41].This trial found similar results to our study (HR 1.41, 95% CI: 1.00-1.92), but was not powered to directly compare lithium and valproate.

A limitation of interpretation of data from cohort studies is the inability to rule out important confounding effects. We attempted to account for confounding by indication by building a PS model that included important clinical predictors of treatment allocation[30]. This included physical health variables which may lead a clinician to avoid a certain drug because of its side effect profile, e.g. renal disease with lithium or cardiovascular disease with olanzapine. Characteristics such as gender, age and BMI were also included, as valproate is contraindicated (though commonly prescribed) in women of childbearing potential[5], and olanzapine has the potential to cause rapid weight gain[42]. Adjusting for the GP practice should account for physician preference for a particular drug. Once these covariates were adjusted for, there was a similar propensity for patients to be prescribed valproate, olanzapine or quetiapine, with patients prescribed lithium having slightly higher scores. Despite this, we cannot rule out the possibility that these confounders were imperfectly adjusted for, or that other important confounders were not included in the PS model.

Unfortunately, we were unable to separate treatment failure relating to emergent manic (or hypomanic) episodes from depressive episodes, and there is evidence that the study drugs may be differentially effective in preventing a particular polarity of illness[9]. However, an ideal "mood stabilizer" would protect against both polarities of relapse[43], and this is what our study captures. We were also unable to examine the physician's reason for treatment initiation, and it may be that quetiapine's apparent inferiority is because in some patients it is prescribed as maintenance treatment, but for shorter term indications (which we hoped to capture in the supplementary analysis). There were too few patients on monotherapy with other recommended maintenance treatments, such as lamotrigine or aripiprazole, to include these drugs in the analysis.

In conclusion, this study provides necessary supplementary and complementary evidence to RCT findings for maintenance treatments for bipolar disorder. In real world clinical

practice, lithium appears to be the most effective treatment to prevent relapse or recurrence of bipolar disorder and may prolong the time before adjunctive prescribing is necessary. This finding echoes the results of recent meta-analyses that suggest lithium is superior to other drugs in protecting against both manic and depressive relapse[9,10]. This is important as lithium is often avoided because of its side effect profile[44], but monotherapy with valproate, olanzapine or quetiapine is more likely to fail sooner and may result in patients experiencing the additive side effects of multiple psychotropic drugs.

REFERENCES

1. Saunders KE, Goodwin GM. The course of bipolar disorder. Adv Psychiatr Treat 2010;16:318-28.
2. Murray CJ, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. Lancet 1997;349:1436-42.
3. Perlis RH, Ostacher MJ, Patel JK et al. Predictors of recurrence in bipolar disorder: primary outcomes from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). Am J Psychiatry 2006;163: 217-24.
4. Hayes J, Prah P, Nazareth I et al. Prescribing trends in bipolar disorder: cohort study in the United Kingdom THIN primary care database 1995-2009. PLoS One 2011;6:e28725.
5. National Institute for Health and Care Excellence. Bipolar disorder: the management of bipolar disorder in adults, children and adolescents, in primary and secondary care. GC38. London: National Institute for Health and Care Excellence, 2006.
6. American Psychiatric Association. American Psychiatric Association practice guideline for the treatment of patients with bipolar disorder. Washington: American Psychiatric Association, 2002.
7. Mok YM, Chan HN, Chee KS et al. Ministry of Health clinical practice guidelines: bipolar disorder. Singapore Med J 2011;52:914-8.
8. Yatham LN, Kennedy SH, Parikh SV et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) collaborative update of CANMAT guidelines for the management of patients with bipolar disorder: update 2013. Bipolar Disord 2013; 15:1-44.
9. Miura T, Noma H, Furukawa TA et al. Comparative efficacy and tolerability of pharmacological treatments in the maintenance treatment of bipolar disorder: a systematic review and network meta-analysis. Lancet Psychiatry 2014;1:351-9.
10. Severus E, Taylor MJ, Sauer C et al. Lithium for prevention of mood episodes in bipolar disorders: systematic review and meta-analysis. Int J Bipolar Disord 2014;2:15.
11. National Institute for Health and Care Excellence. Bipolar disorder: the management of bipolar disorder in adults, children and adolescents, in primary and secondary care. GC185. London: National Institute for Health and Care Excellence, 2014.
12. Reed C, Novick D, Gonzalez-Pinto A et al. Observational study designs for bipolar disorder – What can they tell us about treatment in acute mania? Prog Neuropsychopharmacol Biol Psychiatry 2009;33:715-21.
13. Baldessarini RJ. Treatment research in bipolar disorder. CNS Drugs 2002; 16:721-9.
14. Lancet Editors. Cross design synthesis: a new strategy for studying medical outcomes? Lancet 1992;340:944-6.
15. Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ 1996;312:1215-8.
16. Baldessarini RJ, Tohen M, Tondo L. Maintenance treatment in bipolar disorder. Arch Gen Psychiatry 2000;57:490-2.
17. IMS Health. Our data. www.epic-uk.org.
18. Chisholm J. The Read clinical classification. BMJ 1990;300:1092.
19. Davé S, Petersen I. Creating medical and drug code lists to identify cases in primary care databases. Pharmacoepidemiol Drug Saf 2009;18:704-7.
20. Health and Social Care Information Centre. Prescriptions dispensed in the community, Statistics for England - 2001-2011. Leeds: Health and Social Care Information Centre, 2012.
21. Blak B, Thompson M, Bourke A. National representativeness and data quality of The Health Improvement Network (THIN) database of primary care information for epidemiological research. Presented at the 9th Annual Conference of the UK Federation of Primary Care Research, Liverpool, November 2006.
22. Blak BT, Thompson M, Dattani H et al. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. Inform Prim Care 2011;19:251-5.
23. Lis Y, Mann RD. The VAMP Research multi-purpose database in the UK. J Clin Epidemiol 1995;48:431-43.
24. Hardoon S, Hayes JF, Blackburn R et al. Recording of severe mental illness in United Kingdom primary care, 2000-2010. PLoS One 2013;8:e82365.
25. Nazareth I, King M, Haines A et al. Accuracy of diagnosis of psychosis on general practice computer system. BMJ 1993;307:32-4.
26. Horsfall L, Walters K, Petersen I. Identifying periods of acceptable computer usage in primary care research databases. Pharmacoepidemiol Drug Saf 2013;22:64-9.
27. Maguire A, Blak BT, Thompson M. The importance of defining periods of complete mortality reporting for research using automated data from primary care. Pharmacoepidemiol Drug Saf 2009;18:76-83.
28. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. J Am Stat Assoc 1984;79:516-24.
29. Holmes WM. Using propensity scores in quasi-experimental designs. Thousand Oaks: Sage Publications, 2013.
30. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41-55.
31. Joffe MM, Rosenbaum PR. Invited commentary: Propensity scores. Am J Epidemiol 1999;150:327-33.
32. Austin PC, Mamdani MM, Stukel TA et al. The use of the propensity score for estimating treatment effects: administrative versus clinical data. Stat Med 2005;24:1563-78.
33. Becker SO, Ichino A. Estimation of average treatment effects based on propensity scores. Stata J 2002;2:358-77.
34. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika 1982;69:239-41.
35. d'Agostino RB. Tutorial in biostatistics: Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med 1998;17:2265-81.
36. Lin TH, Dayton CM. Model selection information criteria for non-nested latent class models. J Educ Behav Stat 1997;22:249-64.
37. Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. Am Stat 1985;39:33-8.
38. Stata. Release 13. Statistical Software, 2013.
39. Sajatovic M, Valenstein M, Blow FC et al. Treatment adherence with antipsychotic medications in bipolar disorder. Bipolar Disord 2006;8:232-41.
40. Sajatovic M, Valenstein M, Blow F et al. Treatment adherence with lithium and anticonvulsant medications among patients with bipolar disorder. Psychiatr Serv 2007;58:855-63.
41. Geddes J, Goodwin G, Rendell J et al. Lithium plus valproate combination therapy versus monotherapy for relapse prevention in bipolar I disorder (BALANCE): a randomised open-label trial. Lancet 2010;375:385-95.
42. Eder U, Mangweth B, Ebenbichler C et al. Association of olanzapine-induced weight gain with an increase in body fat. Am J Psychiatry 2001; 158:1719-22.
43. Bauer MS, Mitchner L. What is a "mood stabilizer"? An evidence-based response. Am J Psychiatry 2004;161:3-18.
44. Shine B, McKnight RF, Leaver L et al. Long-term effects of lithium on renal, thyroid, and parathyroid function: a retrospective analysis of laboratory data. Lancet 2015;386:461-8.