

# Comparative effectiveness of valproic acid in different serum concentrations for maintenance treatment of bipolar disorder: A retrospective cohort study using target trial emulation framework

Yang-Chieh Brian Chen,<sup>a</sup> Chih-Sung Liang,<sup>b,c</sup> Liang-Jen Wang,<sup>d</sup> Kuo-Chuan Hung,<sup>e,f</sup> Andre F. Carvalho,<sup>g</sup> Marco Solmi,<sup>h,i,j,k,l</sup> Eduard Vieta,<sup>m</sup> Ping-Tao Tseng,<sup>n,o,p</sup> Pao-Yen Lin,<sup>a,q</sup> Yu-Kang Tu,<sup>r,s</sup> Chih-Wei Hsu,<sup>a,t,\*</sup> and Edward Chia-Cheng Lai<sup>u</sup>

<sup>a</sup>Department of Psychiatry, Kaohsiung Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Kaohsiung, Taiwan

<sup>b</sup>Department of Psychiatry, Beitou branch, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan

<sup>c</sup>Department of Psychiatry, National Defense Medical Center, Taipei, Taiwan

<sup>d</sup>Department of Child and Adolescent Psychiatry, Kaohsiung Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Kaohsiung, Taiwan

<sup>e</sup>Department of Anesthesiology, Chi Mei Medical Center, Tainan city, Taiwan

<sup>f</sup>Department of Hospital and Health Care Administration, College of Recreation and Health Management, Chia Nan University of Pharmacy and Science, Tainan city, Taiwan

<sup>g</sup>IMPACT (Innovation in Mental and Physical Health and Clinical Treatment) Strategic Research Centre, School of Medicine, Barwon Health, Deakin University, Geelong, VIC, Australia

<sup>h</sup>Department of Psychiatry, University of Ottawa, Ottawa, Canada

<sup>i</sup>Department of Mental Health, The Ottawa Hospital, Ottawa, Canada

<sup>j</sup>Ottawa Hospital Research Institute (OHRI), Clinical Epidemiology Program, University of Ottawa, Ottawa, Canada

<sup>k</sup>School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, Canada

<sup>l</sup>Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin, Berlin, Germany

<sup>m</sup>Bipolar and Depressive Disorders Unit, Hospital Clinic, IDIBAPS, CIBERSAM, University of Barcelona, Barcelona, Catalonia, Spain

<sup>n</sup>Prospect Clinic for Otorhinolaryngology & Neurology, Kaohsiung, Taiwan

<sup>o</sup>Institute of Biomedical Sciences, National Sun Yat-sen University, Kaohsiung, Taiwan

<sup>p</sup>Department of Psychology, College of Medical and Health Science, Asia University, Taichung, Taiwan

<sup>q</sup>Institute for Translational Research in Biomedical Sciences, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan

<sup>r</sup>Institute of Epidemiology and Preventive Medicine, National Taiwan University College of Public Health, Taipei, Taiwan

<sup>s</sup>Department of Dentistry, National Taiwan University Hospital, Taipei, Taiwan

<sup>t</sup>Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan, Taiwan

<sup>u</sup>School of Pharmacy, Institute of Clinical Pharmacy and Pharmaceutical Sciences, College of Medicine, National Cheng Kung University, Tainan, Taiwan

## Summary

**Background** There is a lack of consensus on the optimal serum valproic acid (VPA) concentration for maintenance therapy in bipolar disorder (BD). We aimed to investigate the association between serum VPA levels and risk of mood episode recurrence.

**Methods** We enrolled patients with BD from multiple medical institutions in Taiwan between January 1, 2001 and December 31, 2019. Patients were divided into three groups according to their serum VPA concentrations (< 50 µg/ml, 50–74 µg/ml, and 75–104 µg/ml). Adjusted hazard ratios (aHR) with 95% confidence intervals (CIs) compared times to mood episode recurrence using the < 50 µg/ml group as reference. A systematic review found relevant articles published before February 2022 (PROSPERO: CRD42022309661), and corresponding results were compared.

**Findings** This cohort included 896 patients for an intention-to-treat analysis. Compared with the < 50 µg/ml group, a non-significantly lower risk of mood episode recurrence was found in the 50–74 µg/ml (aHR, 95% CI: 0.86, 0.71–1.05) and 75–104 µg/ml (0.91, 0.71–1.18) groups. A per-protocol analysis of 481 patients found a significant risk

eClinicalMedicine

2022;54: 101678

Published online xxx

<https://doi.org/10.1016/j.eclinm.2022.101678>

eclinm.2022.101678

\*Corresponding author at: Department of Psychiatry, Kaohsiung Chang Gung Memorial Hospital, No. 123, Dapi Road, Niasong District, Kaohsiung City 833, Taiwan.

E-mail address: [harwicacademia@gmail.com](mailto:harwicacademia@gmail.com) (C.-W. Hsu).

reduction in the 50–74 µg/ml group (0.76, 0.60–0.97), with inconclusive results in the ≥ 75 µg/ml group (1.03, 0.73–1.46). A meta-analysis including two studies (254 patients) and our cohort found a similar significantly lower risk of mood episode recurrence in the 50–74 µg/ml group (HR, 95% CI: 0.83, 0.69–0.99), while risk reduction in the 75–99 µg/ml group (0.62, 0.26–1.48) did not differ significantly from that in the < 50 µg/ml group or the 50–74 µg/ml group.

**Interpretation** The combined results of this cohort and previous studies suggest that VPA between 50–74 µg/ml may be a more effective concentration to prevent acute mood episodes during maintenance therapy in patients with BD compared with VPA < 50 µg/ml.

**Funding** Ministry of Science and Technology, Taiwan, and Chang Gung Medical Research Project.

**Copyright** © 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

**Keywords:** Bipolar disorder; Effectiveness; Prophylaxis; Serum level; Valproate

### Research in context

#### *Evidence before this study*

Valproic acid is a common drug of choice for the maintenance treatment of bipolar disorder for which routine serum concentration monitoring is required. We searched PubMed/MEDLINE, Embase, and Cochrane CENTRAL using the terms ("valproate sodium" OR "depakine" OR "valproic acid" OR "divalproex") and ("serum" OR "blood" OR "plasma") and ("level" OR "concentration") and ("bipolar") from the inception of each database up to February 1, 2022. Current treatment guidelines still lack recommendations for serum valproic acid concentrations to prevent mood episode recurrence during the maintenance treatment of bipolar disorder.

#### *Added value of this study*

This large-scale retrospective cohort and systematic review and meta-analysis provides a comprehensive survey of different serum valproic acid concentrations and the risk of mood episode recurrence. Compared with the < 50 µg/ml group, our cohort found a consistently lower risk of recurrent mood episodes in the 50–74 µg/ml group across all analyses. Results remained robust throughout the current meta-analysis. Our cohort showed inconclusive and non-significant results for risk of recurrent mood episodes in the 75–99 µg/ml group, and the meta-analysis found a non-significant risk reduction.

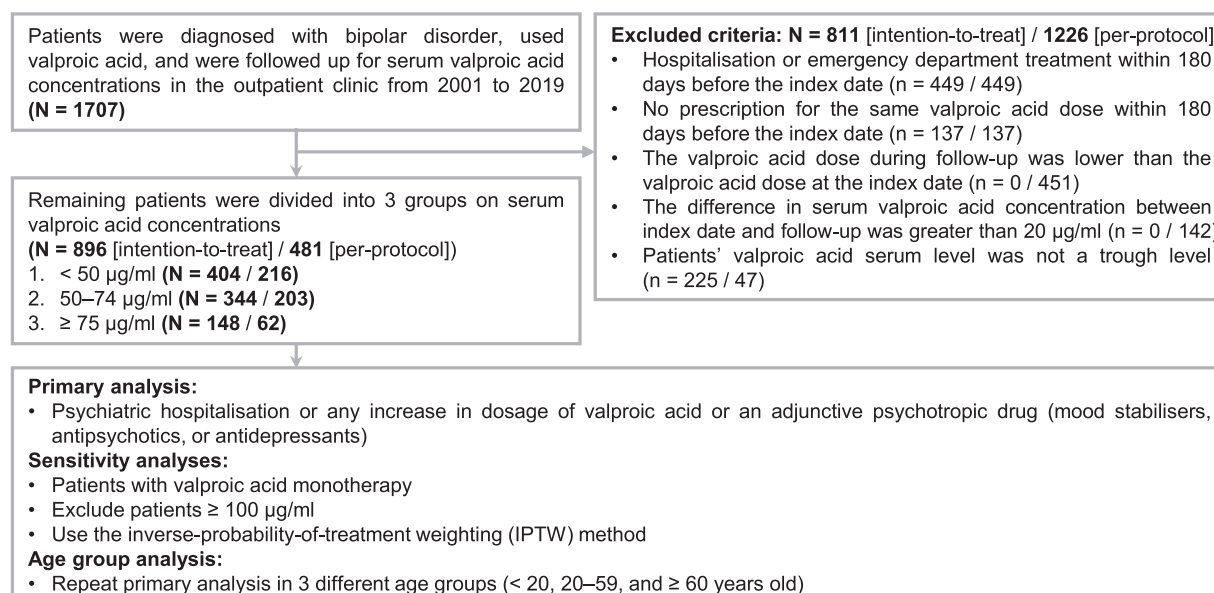
#### *Implications of all the available evidence*

Our findings suggest that the trough levels of 50–74 µg/ml may be a more effective serum valproic acid concentration for preventing mood episode recurrences compared with < 50 µg/ml.

### Introduction

Bipolar disorder (BD) is a serious mental disorder associated with high rates of disability, suicide, and mortality.<sup>1</sup> Prevention of major depressive or manic episode recurrence is a major challenge for the long-term management of BD.<sup>1</sup> Therefore, maintenance treatment after management of acute mood episodes is a crucial step for the prevention of the recurrence of major affective episodes.<sup>1</sup> Mood stabilisers such as valproic acid (VPA) remain the mainstay treatment of choice during the maintenance phase.<sup>1,2</sup> Multiple studies have shown VPA's consistent effectiveness to prevent patients' mood recurrence and improve patients' quality of life when used as maintenance treatment for BD, making it a common drug choice for stable patients with BD.<sup>3</sup>

Although VPA may be effective in preventing the recurrence of any mood episode during maintenance treatment,<sup>4</sup> treatment with VPA requires routine monitoring of serum concentration due to its narrow therapeutic index and potential drug interaction.<sup>5,6</sup> VPA has been known to increase the sedative effects of neuroleptics and benzodiazepines, and also to inhibit the metabolism of lamotrigine.<sup>7</sup> Furthermore, VPA exhibits a complex pharmacokinetic profile due to concentration-dependent protein binding and clearance characteristics that have been shown to vary among patients of different age groups.<sup>8</sup> In the past, few studies have assessed the optimal (i.e., the most efficacious) range of serum VPA concentrations for the maintenance treatment of patients with BD. Current major guidelines (e.g., the British Association for Psychopharmacology (BAP) guideline,<sup>9</sup> 2018 Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) guidelines,<sup>10</sup> and the National Institute for Health and Care Excellence (NICE) guideline<sup>11</sup>) either lack suggestions for an optimal serum VPA concentration or simply suggest



**Figure 1.** Flowchart of patient selection process for the study and outline of analytic procedures.

maintaining a serum VPA concentration within accepted laboratory values, such as 50–100 µg/ml, highlighting the need for further research.

This study aimed to evaluate the therapeutic effectiveness of various serum VPA concentrations for maintenance treatment and the association with the risk of recurrence of major mood episodes in patients with BD. We conducted a multicentre retrospective cohort study to examine the association between different serum VPA concentrations and the risk of mood episode recurrence during maintenance treatment in a large sample of patients with BD. Given previous studies on optimal serum concentrations of lithium, young and older adults may be special populations in clinical practice whose optimal serum levels of VPA may differ from that for adults.<sup>12</sup> Therefore, we also investigated this relationship among different age groups. Second, we carried out a systematic review and meta-analysis to integrate our results with those from previous studies and to summarise the best evidence for this controversial topic.

## Method

### Data source

This cohort study was conducted in line with the Declaration of Helsinki and the REporting of studies Conducted using Observational Routinely-collected Data guidelines (Appendix 1).<sup>13</sup> Informed consent was not required, as all data in the Chang Gung Research Database (CGRD) available for research and analytic purposes were retrospective with patient and provider

information encrypted and de-identified by irreversibly replacing direct identifiers for each individual with unique numbers. The study protocol was approved by the institutional review board of Chang Gung Memorial Hospital (CGMH, No: 202100131Bo). Data used in this study were retrieved from CGRD from January 1, 2001 to December 31, 2019. The database contains multi-institutional de-identified data, including patient demographics (age and sex), disease diagnoses, medications (type, dosage, and duration), hospital visits (inpatient and outpatient), laboratory data (serum VPA concentrations), and examination reports, from the electronic medical records of seven medical institutes of the CGMH in Taiwan. From 1997–2010, the CGRD covered 12.4% and 21.2% of all inpatient and outpatient data, respectively, in Taiwan.

### Participants

Figure 1 shows the patient selection process for this retrospective cohort study and the inclusion and exclusion processes. We designed this observational study to pursue a target trial emulation of a randomised controlled trial (i.e., a hypothetical pragmatic trial) that would have compared the effectiveness of different serum VPA concentrations for maintenance treatment of patients with BD.<sup>14</sup> The specifics of trial emulation are provided in Appendix 2 (Table S2). We included patients diagnosed with BD by using the diagnosis code of International Classification of Diseases 9th and 10th edition: [ICD-9] 296.0–296.8 (except 296.2 and 296.3) or [ICD-10] F30 and F31, and those who received VPA with concurrent laboratory records of serum VPA concentrations. The index date was defined as the first date when the

patient's serum VPA concentration was available. We excluded patients by the following criteria to define the intent-to-treat population: 1) patients with either inpatient or emergency department treatment within 180 days before the index date; 2) patients who had not taken the same VPA dosage within 180 days before the index date; 3) patients' data of serum VPA concentrations were not a trough level.<sup>15</sup> Then, patients were divided into three groups according to serum VPA concentrations: < 50 µg/ml, 50–74 µg/ml, and ≥ 75 µg/ml. For the last group, five patients had serum VPA concentrations of 100–104 µg/ml and they were merged into the ≥ 75 µg/ml group to accommodate for the small sample sizes. These groups were chosen *a priori*, based on a previous study design.<sup>16,17</sup>

Additionally, we added two exclusion criteria to further investigate the per-protocol population constituted by patients who were able to adhere to the study protocol and receive a stable dose of VPA during the study: 1) compared to the VPA dose taken at the index date, the patients' VPA dose was reduced during the follow-up; 2) the difference in the serum concentration between the index date and any follow-up was > 20 µg/ml. In the per-protocol population, two patients in the ≥ 75 µg/ml group had serum VPA concentrations of 100–102 µg/ml.

#### Follow up and outcome measures

The effectiveness outcome was the recurrence of any major mood episode, defined as the event of psychiatric hospitalisation or any increase in the dose of VPA or adjunctive psychotropic medications (mood stabilisers, antipsychotics, and antidepressants). These indicators were based on previous studies.<sup>16,18</sup> We followed up with patients from the index date for two years or until the occurrence of an outcome event.

#### Covariates

We selected the following patient characteristics based on literature review and experts' opinions,<sup>16–18</sup> including age, sex, number of psychiatric hospitalisations within the two years before the index date, medical comorbidities (hypertension, diabetes mellitus, hyperlipidaemia, rheumatic diseases, any cancer, heart diseases, peripheral vascular diseases, cerebral vascular diseases, pulmonary diseases, liver diseases, peptic ulcer diseases, renal diseases, and any bleeding), and concomitant medications (mood stabilisers, antipsychotics, antidepressants, and benzodiazepines). Details regarding the medical comorbidities and concomitant medications are provided in Appendix 3 (Table S3).

#### Statistical analysis

In this study, we analysed the intention-to-treat population as the primary outcome and examined its robustness using the per-protocol population. Continuous variables were described by means with standard deviations, and categorical variables by numbers and

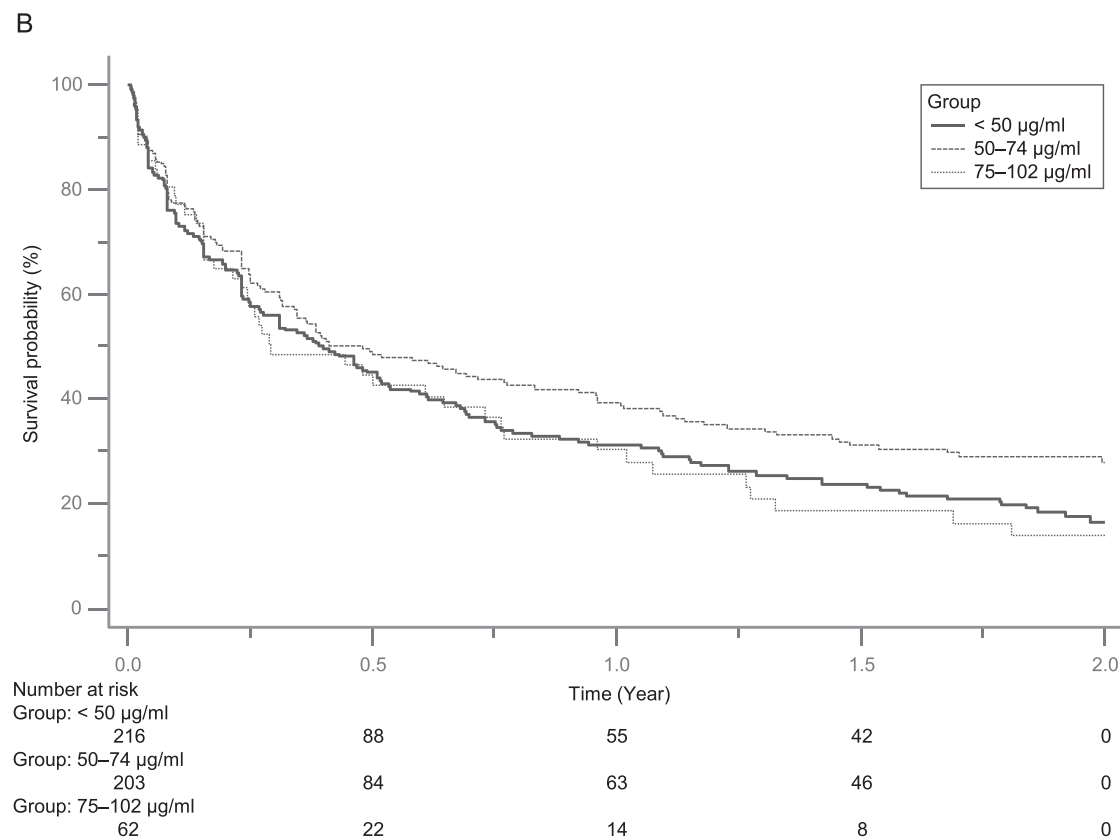
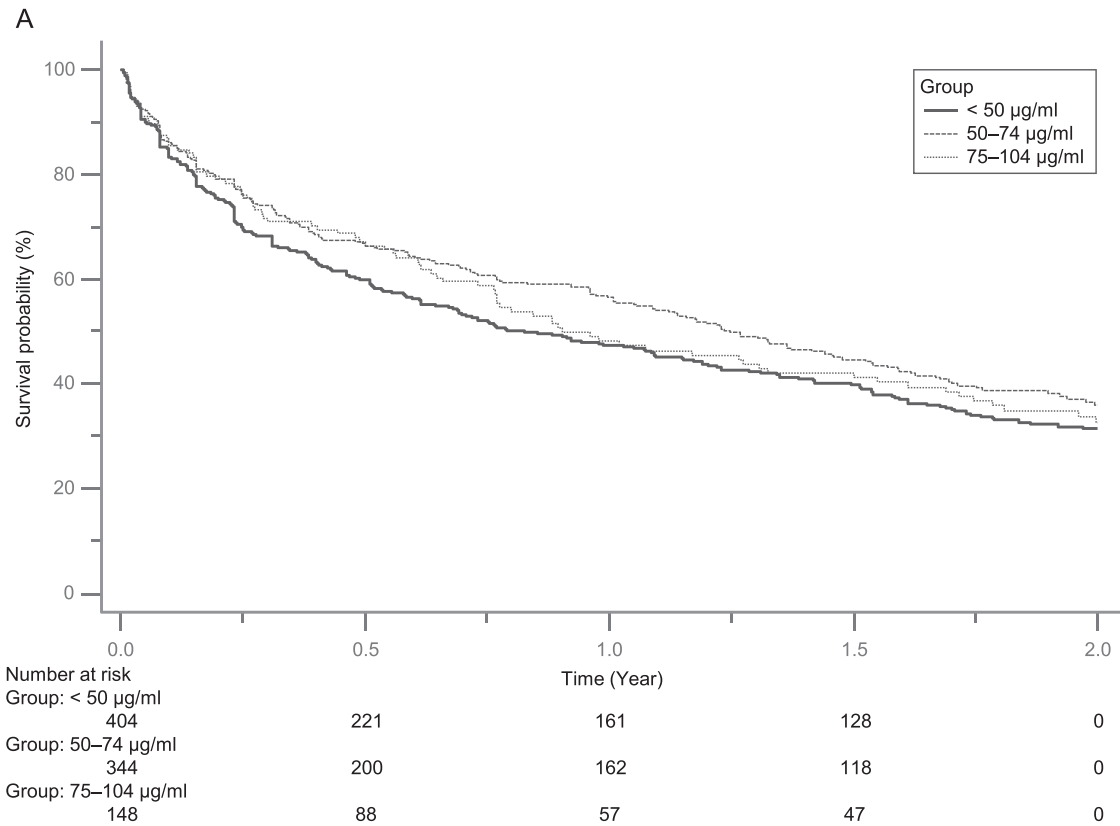
proportions. Kaplan-Meier curves were used for survival analysis (Figures 2 and S1). When applicable, we performed a multivariable Cox regression model to generate adjusted hazard ratios (aHRs) with 95% confidence intervals (CIs) to compare the risk of mood episode recurrence among different groups by the levels of serum VPA concentrations, with the adjustment of all baseline covariates listed in Table 1 (intent-to-treat population) or Appendix 4 (Table S4, per-protocol population). The group with < 50 µg/ml serum concentration was considered the reference group. There were no missing values in this study, but patients were excluded from both analyses, as indicated in Figure 1.

#### Sensitivity analyses

To determine the consistency of our findings, we performed three sensitivity analyses by selecting patients under different scenarios. The first scenario assessed the outcomes of the patients treated with VPA monotherapy. Second, we used a new serum range of 75–99 µg/ml for the third group of patients, excluding the five patients in intention-to-treat analysis (100–104 µg/ml) or two patients in per-protocol analysis (100–102 µg/ml) with serum concentrations ≥ 100 µg/ml. Most treatment guidelines do not recommend serum VPA concentrations greater than 100 µg/ml for the treatment of acute mood episodes,<sup>7,10</sup> leading to the inference that the same should be true for serum VPA concentrations during maintenance treatment. Lastly, we reanalysed the original data using the inverse probability of treatment weighting (IPTW) method with propensity score. The propensity score was generated using logistic regression conditioning on all covariates listed in Table 1. The details of the IPTW approach have been described elsewhere.<sup>19</sup> Briefly, the IPTW approach provided the average treatment effects of drugs based on the entire study population. After IPTW adjustment, the probability of being assigned to each group would be the same, similar to a clinical trial in which all groups have the same probability of treatment assignment through randomisation (Appendix 5–6, Table S5–S6). Moreover, we conducted a stratification analysis by age group (< 20, 20–59, and ≥ 60 years) to examine the association between serum VPA concentrations and prevention of mood episodes among the different age groups.<sup>20</sup> We used the SAS Software version 9.4 (SAS Institute Inc., Cary, NC, USA) for all statistical analyses. Statistical significance was defined as a two-tailed *p*-value of < 0.05.

#### Systematic review and meta-analysis

We used a systematic review of past studies to investigate the effectiveness of different serum VPA concentrations in the maintenance treatment of patients with BD and contextualised our results with those from previous studies. The review was outlined by an *a priori* registered protocol (PROSPERO: CRD42022309661), which



**Figure 2.** Survival analysis using Kaplan-Meier curves. A) Intention-to-treat population; B) Per-protocol population.

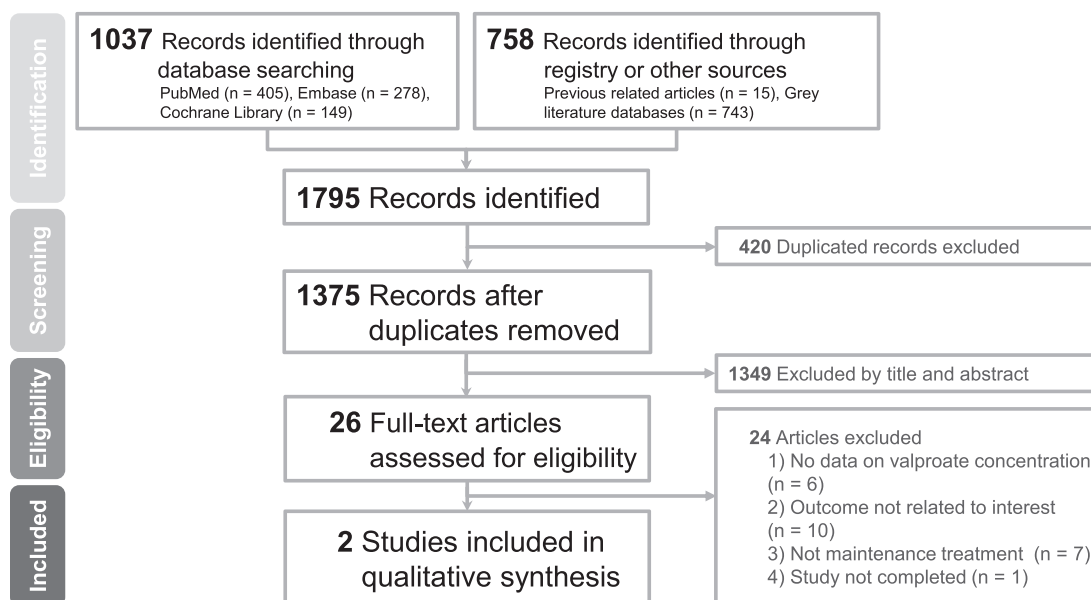
Characteristics	All participants (N = 896)	< 50 µg/ml (n = 404)	50–74 µg/ml (n = 344)	75–104 µg/ml (n = 148)
Age, year	43.6 (15.5)	44.4 (15.4)	43.6 (15.8)	41.7 (15.4)
Sex				
Female	444 (50%)	205 (50%)	155 (45%)	84 (57%)
Male	452 (50%)	199 (49%)	189 (55%)	64 (43%)
Psychiatry hospitalisations				
0	689 (77%)	320 (79%)	257 (75%)	112 (76%)
1–2	178 (20%)	79 (20%)	71 (21%)	28 (19%)
≥ 3	29 (3%)	5 (1%)	16 (5%)	8 (5%)
Comorbidities				
Hypertension	140 (16%)	54 (13%)	59 (17%)	27 (18%)
Diabetes mellitus	86 (10%)	34 (8%)	35 (10%)	17 (11%)
Hyperlipidaemia	91 (10%)	42 (10%)	34 (10%)	15 (10%)
Rheumatic diseases	6 (1%)	3 (1%)	3 (1%)	0 (0%)
Any cancer	17 (2%)	6 (1%)	6 (2%)	5 (3%)
Heart diseases	10 (1%)	4 (1%)	2 (1%)	4 (3%)
Peripheral vascular diseases	3 (< 1%)	1 (< 1%)	0 (0%)	2 (1%)
Cerebral vascular diseases	32 (4%)	9 (2%)	20 (6%)	3 (2%)
Pulmonary diseases	45 (5%)	17 (4%)	17 (5%)	11 (7%)
Liver diseases	57 (6%)	29 (7%)	14 (4%)	14 (9%)
Peptic ulcer diseases	62 (7%)	24 (6%)	26 (8%)	12 (8%)
Renal diseases	22 (2%)	8 (2%)	11 (3%)	3 (2%)
Any bleeding	91 (10%)	33 (8%)	40 (12%)	18 (12%)
Medications				
Mood stabilizers				
Lithium	215 (24%)	106 (26%)	80 (23%)	29 (20%)
Lamotrigine	42 (5%)	15 (4%)	19 (6%)	8 (5%)
Carbamazepine	10 (1%)	6 (1%)	4 (1%)	0 (0%)
Antipsychotics	662 (74%)	306 (76%)	245 (71%)	111 (75%)
Antidepressants	284 (32%)	147 (36%)	95 (28%)	42 (28%)
Benzodiazepines	654 (73%)	308 (76%)	245 (71%)	101 (68%)

**Table 1: Characteristics of the intention-to-treat population and valproic acid groups from 2001 to 2019.**  
Data are expressed as N (%) or mean (standard deviation).

followed the guidelines recommended by the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA, Appendix 10, Table S9).<sup>21</sup> Two investigators (YCC and CWH) independently searched the electronic databases PubMed/MEDLINE, Embase, and Cochrane CENTRAL from the inception of each database up to February 1, 2022, for studies that compared the effectiveness of different serum VPA concentrations on the maintenance treatment of BD patients. We also searched grey literature databases and reference lists of relevant reviews and articles.<sup>15</sup> No limits were set on the study design or language of publication. The detailed search strings and a flowchart of the study selection are provided in Appendix 11 and Figure 3. We included original papers that fulfilled the following criteria: (1) included patients with BD and (2) received VPA maintenance therapy for BD. The exclusion criteria were as follows: (1) patients who were not in BD remission or maintenance therapy; (2) VPA serum concentrations were not used for outcome comparison; and (3) the

study was not completed. Reviews, comments, letters-to-editor, expert opinions, guidelines, non-peer-reviewed articles, and conference papers were also excluded by title/abstract screening.

Data extraction from articles, if available, included characteristics of the study, study design, sample size, definition of VPA maintenance treatment for BD patients, age, range serum VPA concentration of each group, follow-up duration, concomitant medication, and definition of mood episode recurrences. The main outcome was the effectiveness of different serum VPA concentrations in preventing the recurrence of any mood episodes during the maintenance treatment of BD patients, such as HR. For studies that provided Kaplan-Meier survival curves rather than HR data, we estimated HR using a digitalised tool (Digitizelt software: version 2.5; Braunschweig, Germany) and the statistical procedure described by Tierney *et al.*<sup>22</sup> The risk of bias was assessed by two independent investigators (YCC and CWH) using the Newcastle-Ottawa Scale for



**Figure 3.** Flowchart of the selection process for the systematic review and meta-analysis.

cohort studies and the Cochrane risk-of-bias tool for randomised trials. Any discrepancies or disagreements between the two investigators during the processes of study inclusion/exclusion, data extraction, and study appraisal were resolved through discussion. We performed meta-analyses with random-effects models, calculated effect sizes using pooled HR and 95% CIs, and assessed heterogeneity using  $I^2$ . All analyses were performed using Comprehensive Meta-analysis software (version 3; Biostat, Englewood, NJ). A two-tailed  $p$ -value  $< 0.05$  was considered statistically significant.

#### Role of the funding source

The funders had no role in the study design, data collection, data analysis, data interpretation, or report preparation. All authors had full access to all data in the study and the decision to submit for publication.

## Results

#### Baseline characteristics of cohort patients

The intention-to-treat population included a total of 896 patients classified into the following three groups according to serum VPA concentration, represented as mean (standard deviation [SD]):  $< 50$   $\mu\text{g/ml}$ ,  $n = 404$ ,  $33.2$  ( $12.0$ );  $50$ – $74$   $\mu\text{g/ml}$ ,  $n = 344$ ,  $62.0$  ( $7.1$ );  $\geq 75$  ( $75$ – $104$ )  $\mu\text{g/ml}$ ,  $n = 148$ ,  $83.6$  ( $7.1$ ). Their basic characteristics are shown in Table 1. In general, the patients' mean age was  $43.6$  ( $SD = 15.5$ ) years old, 50% ( $n = 452$ ) of patients were male, 23% ( $n = 207$ ) of patients had

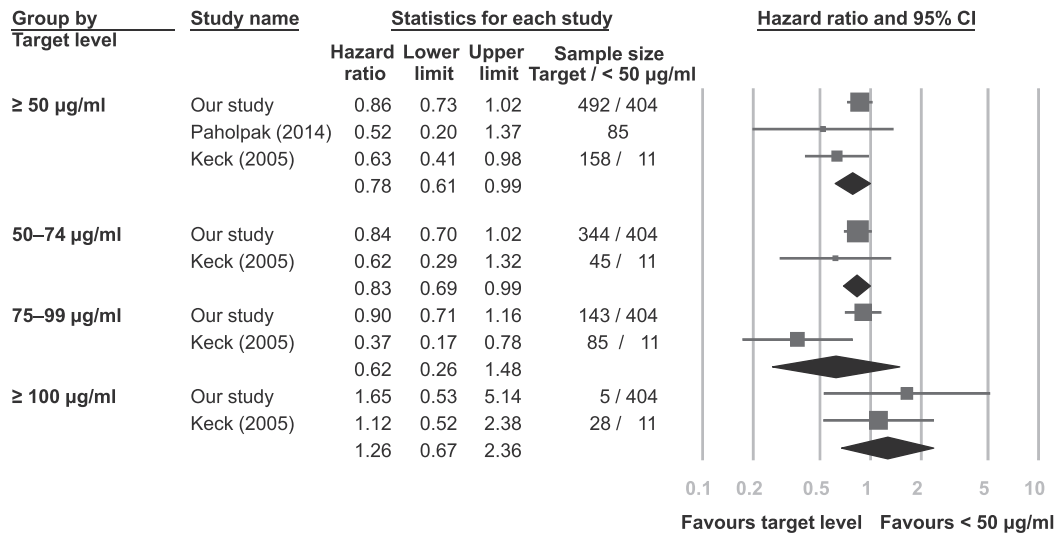
been hospitalised for psychiatric disorders within the two years before the index date, the most common comorbidity was hypertension (16%,  $n = 140$ ), and most patients received concomitant medication with antipsychotics (74%,  $n = 662$ ) and benzodiazepines (73%,  $n = 654$ ). Details of the baseline characteristics of the per-protocol population ( $n = 481$ ) are provided in Appendix 4 (Table S4). All characteristics after IPTW are shown in Appendix 5–6 (Table S5–S6).

#### Outcomes of this cohort study

In the intention-to-treat and per-protocol analyses, 526 and 335 patients experienced recurrence of any kind of mood episode at 820 and 314 person-years of follow-up, respectively. Their results are presented in Table 2 and Figure 2. Compared to patients with serum VPA concentration  $< 50$   $\mu\text{g/ml}$ , patients with serum VPA concentrations  $50$ – $74$   $\mu\text{g/ml}$  had a non-significantly lower rate of mood episode recurrence in the primary analysis (intention-to-treat aHR = 0.86, 95% CI = 0.71–1.05). The per-protocol analysis showed similar findings, with a significant benefit (aHR = 0.76, 95% CI = 0.60–0.97). Taking patients with serum VPA concentration  $< 50$   $\mu\text{g/ml}$  as a reference, those with serum VPA concentrations  $\geq 75$   $\mu\text{g/ml}$  showed inconclusive risks in mood episode recurrence (intention-to-treat aHR = 0.91, 95% CI = 0.71–1.18; per-protocol aHR = 1.03, 95% CI = 0.73–1.46). Neither analysis demonstrated statistical differences. Furthermore, the above findings were broadly similar across all sensitivity analyses ( $50$ – $74$   $\mu\text{g/ml}$ : intention-to-treat aHR between 0.55 and 0.

Group	Incident/total Cases	Person-years	Adjusted HR <sup>a</sup>
<b>Intent-to-treat</b>			
< 50 µg/ml	249/404	354	1.00 [Reference]
50–74 µg/ml	190/344	330	0.86 (0.71–1.05)
75–104 µg/ml	87/148	135	0.91 (0.71–1.18)
<b>Per-protocol</b>			
< 50 µg/ml	162/216	137	1.00 [Reference]
50–74 µg/ml	127/203	143	0.76 (0.60–0.97)*
75–102 µg/ml	46/62	35	1.03 (0.73–1.46)

**Table 2: Comparing the adjusted effectiveness of all patients in different valproic acid groups through the Cox proportional hazard model.**  
Abbreviations: HR, hazard ratio.  
<sup>a</sup> Hazard ratio was adjusted for age, sex, psychiatry hospitalisation, all comorbidities, and all medications; \* indicates  $p < 0.05$ .



**Figure 4.** Forest plots comparing groups with different serum valproic acid target levels and groups < 50 µg/ml.

86, per-protocol aHR between 0.40 and 0.78; ≥ 75 µg/ml: intention-to-treat aHR between 0.62 and 0.90, per-protocol aHR between 0.69 and 1.00) listed in Appendix 7–8 (Table S7 and Figure S1).

Additionally, the 50–74 µg/ml group had a lower risk compared with the ≥ 75 µg/ml group, but the difference did not reach significance (intention-to-treat aHR = 0.94, 95% CI = 0.72–1.22; per-protocol aHR = 0.74, 95% CI = 0.52–1.05; data not shown). Appendix 9 (Table S8) displays the results of the age-stratified subgroup analysis. The findings for the adult group (20–59 years old) were somewhat similar to the results for the entire cohort (50–74 µg/ml intention-to-treat/per-protocol aHR = 0.80/0.67; ≥ 75 µg/ml intention-to-treat/per-protocol aHR = 0.89/1.01). Limited sample sizes of the young (< 20 years old) and elderly (≥ 60 years old) groups prevented detailed commentary due to very wide CIs.

**Systematic review and meta-analysis**

Figure 3 shows a flowchart of the study selection. We found a total of 1795 references in the major databases and grey literature and excluded 420 duplicate articles. We further screened the remaining titles and abstracts and excluded 1349 articles, leaving 26 remaining articles for full-text review. An additional 24 articles were excluded for reasons detailed in Appendix 12 (Table S10). Finally, we included two articles in our systematic review, one retrospective cohort study (85 patients), and one *post-hoc* analysis of data from a randomised controlled trial (169 patients).<sup>16,17</sup> The characteristics of each study are provided in Appendix 13, and the risk-of-bias assessments are shown in Appendix 14.

Figure 4 shows the meta-analysis results using data from our intention-to-treat outcome and the two studies identified via the systematic review. Compared with patients with serum VPA concentrations < 50 µg/ml,



those with serum VPA concentrations  $\geq 50$   $\mu\text{g/ml}$  had a significantly lower rate of mood episode recurrence (HR = 0.78, 95% CI = 0.61–0.99,  $I^2 = 22\%$ ).<sup>16,17</sup> Moreover, compared to those with serum VPA concentrations  $< 50$   $\mu\text{g/ml}$ , a lower risk of recurrent mood episode was found for the 50–74  $\mu\text{g/ml}$  group (HR = 0.83, 95% CI = 0.69–0.99,  $I^2 = 0\%$ ) and the 75–99  $\mu\text{g/ml}$  group (HR = 0.62, 95% CI = 0.26–1.48,  $I^2 = 80\%$ ), but only the 50–74  $\mu\text{g/ml}$  group reached statistical significance.<sup>16</sup> Finally, those with serum VPA concentrations  $\geq 100$   $\mu\text{g/ml}$  had non-significantly higher risk of recurrent mood episode (HR = 1.26, 95% CI = 0.67–2.36,  $I^2 = 0\%$ ).

## Discussion

To the best of our knowledge, this real-world retrospective cohort study is the largest to date relevant to the topic of evaluating the therapeutic effectiveness of different serum VPA concentrations during maintenance treatment of patients with BD for the prophylaxis of mood episodes of any polarity. The primary (intention-to-treat) outcome of our cohort and current meta-analysis indicated that the 50–74  $\mu\text{g/ml}$  and 75–99  $\mu\text{g/ml}$  groups had non-significantly lower risks of mood episode recurrence when compared to the  $< 50$   $\mu\text{g/ml}$  group. However, only the 50–74  $\mu\text{g/ml}$  group had lower risk which reached significant differences in the subsequent per-protocol analysis and current meta-analysis. Moreover, age subgroup analysis showed that 50–74  $\mu\text{g/ml}$  was an effective serum VPA concentration range for the adult group (20–59 years old) compared with  $< 50$   $\mu\text{g/ml}$ .

Our cohort showed that the risk of mood recurrence was consistently lower in the 50–74  $\mu\text{g/ml}$  group than in the  $< 50$   $\mu\text{g/ml}$  group by both the intention-to-treat and per-protocol analyses and their sensitivity tests, but only the per-protocol analysis reached significance. The differences between the results obtained from the analysis of the two populations could be explained by the selection of our subjects. Our intention-to-treat population included patients with VPA dose reduction during the follow-up period, while these patients were excluded from the per-protocol population by the additional exclusion criterion 1. Some of these patients may have been initially assigned to the 50–74  $\mu\text{g/ml}$  group but ended the study with serum VPA concentration  $< 50$   $\mu\text{g/ml}$  due to VPA dose reduction. Meanwhile, the 50–74  $\mu\text{g/ml}$  group of the per-protocol group would not suffer from this problem, and the patients would have more constant levels of VPA dosing and serum concentration. Based on this notion, it is reasonable to find a greater risk reduction for the 50–74  $\mu\text{g/ml}$  group compared to the  $< 50$   $\mu\text{g/ml}$  group, with the per-protocol analysis. Moreover, the  $\geq 75$   $\mu\text{g/ml}$  group had inconclusive and non-significant results for risk of mood recurrence in the intention-to-treat and per-protocol analyses.

The pharmacokinetics of VPA may explain this phenomenon. At the current standard therapeutic range for VPA, most of the drug (approximately 90%) binds to serum proteins.<sup>23</sup> *In vitro* and animal<sup>8,24</sup> studies have shown that the serum VPA concentration reaches the threshold for binding sites at 70  $\mu\text{g/ml}$  and 80  $\mu\text{g/ml}$ , respectively, coinciding with our upper limit cut-off point for the maintenance serum VPA concentration of 75  $\mu\text{g/ml}$ . Past this threshold, free serum VPA concentration as a function of serum VPA concentration transforms from linear to quadratic growth.<sup>25</sup> This non-linearly increasing VPA concentration might have unpredictable effects (beneficial or adverse) in patients treated above 75  $\mu\text{g/ml}$ , thus leading to reduced case numbers in the  $\geq 75$   $\mu\text{g/ml}$  group and inconsistent risk results (as revealed by the analyses in our cohort and the high heterogeneity in our meta-analysis). Finally, although we found that the 50–74  $\mu\text{g/ml}$  group had a lower point-estimated risk than the  $\geq 75$   $\mu\text{g/ml}$  group, the difference has not been proven. Therefore, our cohort could not conclude which of these two VPA concentrations was more effective in preventing recurrent mood episodes.

While examining for missing data in our cohort, we noticed that many patients were not tested for trough VPA concentration and therefore could not be included for analysis. This may reflect the lack of VPA concentration recommendations for maintenance therapy of BD in the current treatment guidelines. The optimal VPA concentration is uncertain; therefore, clinicians may check a rough serum VPA concentration for toxicity ( $> 100$   $\mu\text{g/ml}$ ) or ineffectiveness ( $< 50$   $\mu\text{g/ml}$ ). In fact, measuring trough levels is an important concept in therapeutic drug monitoring as the trough levels are less likely to be influenced by absorption and distribution problems.<sup>26</sup> Past studies have proven that a trough VPA concentration of at least 50  $\mu\text{g/ml}$  is required to achieve therapeutic efficacy in the treatment of BD.<sup>27</sup> Therefore, given these missing values and the reasons mentioned above, our study may also alert clinicians to ensure that therapeutic drug monitoring in VPA is at a trough level. Furthermore, our subgroup analysis with stratification by age showed that only adults (20–59 years old) treated with VPA concentrations of 50–74  $\mu\text{g/ml}$  had a significantly lower risk of mood episode recurrence when compared to the  $< 50$   $\mu\text{g/ml}$  group. The current findings do not show any consistent trends or benefits for any specific VPA serum concentration in the other age groups. The relatively small sample sizes of the young ( $< 20$  years old) and elderly ( $\geq 60$  years old) groups resulted in lower power to detect statistical significance. In this study, we report these preliminary data and hope to increase research interest in these areas to benefit future large trials investigating the effective serum VPA concentrations in young and elderly patients with BD.

We found only two studies in our systematic review, Paholpak *et al.* and Keck *et al.*<sup>16,17</sup> This reflects the current weak consensus on the optimal serum VPA concentration required for the maintenance treatment of BD patients. For example, the 2018 CANMAT and ISBD guidelines<sup>10</sup> simply recommended a VPA concentration within the accepted laboratory range values, most commonly 50–100 µg/ml,<sup>5</sup> during the maintenance treatment of patients with BD, and the 2013 World Federation of Societies of Biological Psychiatry (WFSBP) Guideline<sup>7</sup> recommended a VPA concentration of 45–100 µg/ml. Our findings support the results of Paholpak *et al.*<sup>17</sup> which showed that patients with serum VPA concentrations  $\geq 50$  µg/ml had a lower risk of recurrence of mood episodes than those with serum VPA concentrations  $< 50$  µg/ml. However, our finding is inconsistent with the result from Keck *et al.*<sup>16</sup> which found 75–99 µg/ml to be the better serum VPA concentration for the prophylactic treatment of patients with BD to prevent mood episode recurrence. There are two possible reasons for this finding. First, the follow-up period for Keck *et al.*<sup>16</sup> was one year, while our follow-up period was two years, and their patients were started on maintenance treatment with an average of 74.2 days from the last mood episode, while our patients had been in remission for at least 180 days. The stricter patient selection in our study could mean that our patients had been in a more stable condition with VPA treatment before the initiation of the study period and outcome observation. Second, except for limited lorazepam and haloperidol, all other medications were discontinued before maintenance treatment in the study by Keck *et al.*<sup>16</sup> while many of our patients were concomitantly administered antipsychotics, antidepressants, and/or benzodiazepines.

The strength of the study is that we included a larger sample of patients compared with previous studies<sup>16,17</sup> to improve the representativeness and statistical power and reduce random errors for a more definitive conclusion. This study has some limitations. First, the study may be subject to indication bias (i.e. patients with higher VPA dosage may have had more severe illness, which may have increased risk of relapse) and residual confounding because of non-random allocation through retrospective data from registry-based electronic health records.<sup>28</sup> We considered a large number of covariates, such as concomitant psychotropic agents, in the regression model to minimise the potential confounding effects. However, some patients lacked data on serum concentrations of concomitant mood stabilisers for model adjustment (no data were available for 38 of 215 lithium users, 42 of 42 lamotrigine users, and 5 of 10 carbamazepine users). Second, studies have shown that serum albumin level may affect a patient's serum VPA-free fraction,<sup>23</sup> which is a better indicator of drug activity. However, very few clinical laboratories routinely analyse free serum VPA concentration. Data based on

total serum VPA concentrations may be more practical in a real-world setting. Future studies are warranted to examine the correlation between free serum VPA concentration and total serum VPA concentration or whether serum albumin level affects the therapeutic effectiveness of VPA in patients with BD. Third, the registry-based data led to some limitations. The CGRD did not provide information regarding validated rating scales for depressive or mania symptoms (e.g., Young Mania Rating Scale or Hamilton Depression Rating Scale), so we were unable to differentiate recurrent mood episodes as mania or depressive. Moreover, BD patients identified in CGRD were included over a long period of time (from 2001 to 2019), came from different medical institutions, were evaluated by different clinicians, and used two different diagnostic systems (ICD-9 and ICD-10). Under these circumstances, CGRD has not yet demonstrated the validity or reliability of these BD diagnoses. However, a post-hoc analysis indicated that all patients in our study were diagnosed by psychiatrists, and psychiatrists in Taiwan are well trained by the Taiwanese Society of Psychiatry to make standard and consistent coding behaviours. Accordingly, the diagnosis of BD we defined should be relatively sound. Fourth, information outside the CGRD system was unavailable, and we may have underestimated the incidence of mood episode recurrences. However, we considered this situation to be non-differential among the groups and, as a result, the potential bias was not informative. Fifth, it is noteworthy that our findings are only valid for VPA use in BD patients and do not provide any inference for off-label use of VPA in patients with other diseases. Besides, our study did not address the benefit-risk of VPA prescription; hence, a reminder of the teratogenic risks associated with VPA in women of child-bearing age is warranted,<sup>29</sup> especially given that its prescription in this population remains an issue in many countries.<sup>30</sup> Finally, our study design used exclusion criteria that referenced previous studies on lithium and VPA,<sup>15-18</sup> but findings in patients remaining after these exclusion criteria may not be generalisable to the general patient population. Furthermore, we did not adjust for statistical significance when performing multiple tests in this study (50–74 µg/ml vs.  $< 50$  µg/ml and  $\geq 75$  µg/ml vs.  $< 50$  µg/ml), which might have resulted in the significance no longer being the nominal 5%. To this end, we performed several sensitivity analyses and a meta-analysis, and the 50–74 µg/ml group showed the most consistent trends.

This large retrospective cohort study based on real-world situations showed that compared to those with  $< 50$  µg/ml serum VPA concentration, BD patients with 50–74 µg/ml serum VPA concentrations for maintenance treatment had a consistently lower risk of mood episode across all analyses. Results remained robust throughout the current meta-analysis. However, our study did not prove any difference between serum VPA

concentrations 75–99 µg/ml and < 50 µg/ml for the maintenance treatment. These results may serve as the basis for future large prospective studies to confirm these findings.

### Contributors

C.W.H. conceived the research idea for the study. C.W.H. led the study design, with Y.C.C. and C.C.L. C.W.H. contributed to data acquisition and extraction, and Y.C.C. verified the underlying data. C.W.H. performed statistical analysis. Y.C.C. drafted the manuscript first, then C.S.L., L.J.W., K.C.H., A.F.C., M.S., E.V., P.T.T., P.Y.L., Y.K.T., C.W.H., and C.C.L. revised the manuscript. All authors contributed important intellectual content during manuscript revision, had full access to all the data in the study, and accept responsibility to submit for publication.

### Data sharing statement

The data that support the findings of this study are available from the corresponding author, C.W.H., upon reasonable request.

### Declaration of interests

Prof. Vieta has received grants and served as consultant, advisor or CME speaker for the following entities: AB-Biotics, AbbVie, Angelini, Biogen, Boehringer-Ingelheim, Celon Pharma, Dainippon Sumitomo Pharma, Ferrer, Gedeon Richter, GH Research, Glaxo-Smith Kline, Janssen, Lundbeck, Novartis, Orion Corporation, Organon, Otsuka, Sage, Sanofi-Aventis, Sunovion, Takeda, and Viatrix, outside the submitted work. Dr. Solmi received honoraria and has been a consultant for Angelini, Lundbeck, Otsuka, outside the submitted work. All the other authors declared no conflict of interest.

### Acknowledgements

The authors would like to thank Ms. Pei-Ying Yang and Mr. Chien-An Hu for the technical support. This study is supported by grants from the Ministry of Science and Technology, Taiwan (MOST 109-2314-B-182A-009-MY2 and 111-2314-B-182A-027-) and the Chang Gung Medical Research Project (CMRPG8L0871 and CMRPG8M0531). The funding sources had no role in the design of the study.

### Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.eclinm.2022.101678.

### References

- Carvalho AF, Firth J, Vieta E. Bipolar disorder. *N Engl J Med*. 2020;383(1):58–66.
- Thase ME. Maintenance therapy for bipolar disorder. *J Clin Psychiatry*. 2008;69(11):e32.
- Revicki DA, Hirschfeld RM, Ahearn EP, Weisler RH, Palmer C, Keck Jr. PE. Effectiveness and medical costs of divalproex versus

- lithium in the treatment of bipolar disorder: results of a naturalistic clinical trial. *J Affect Disord*. 2005;86(2-3):183–193.
- Cipriani A, Reid K, Young AH, Macritchie K, Geddes J. Valproic acid, valproate and divalproex in the maintenance treatment of bipolar disorder. *Cochrane Database Syst Rev*. 2013(10):CD003196. <https://doi.org/10.1002/14651858.CD003196.pub2>.
- Patsalos PN, Spencer EP, Berry DJ. Therapeutic drug monitoring of antiepileptic drugs in epilepsy: a 2018 update. *Ther Drug Monit*. 2018;40(5):526–548.
- Tseng YJ, Huang SY, Kuo CH, Wang CY, Wang KC, Wu CC. Safety range of free valproic acid serum concentration in adult patients. *PLoS One*. 2020;15(9):e0238201.
- Grunze H, Vieta E, Goodwin GM, et al. The World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for the biological treatment of bipolar disorders: update 2012 on the long-term treatment of bipolar disorder. *World J Biol Psychiatry*. 2013;14(3):154–219.
- Kodama Y, Kodama H, Kuranari M, et al. Gender- or age-related binding characteristics of valproic acid to serum proteins in adult patients with epilepsy. *Eur J Pharm Biopharm*. 2001;52(1):57–63.
- Goodwin GM, Haddad PM, Ferrier IN, et al. Evidence-based guidelines for treating bipolar disorder: revised third edition recommendations from the British Association for Psychopharmacology. *J Psychopharmacol*. 2016;30(6):495–553.
- Yatham LN, Kennedy SH, Parikh SV, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) 2018 guidelines for the management of patients with bipolar disorder. *Bipolar Disord*. 2018;20(2):97–170.
- Bipolar disorder: assessment and management (CG185)*. London: NICE; 2018.
- Nolen WA, Licht RW, Young AH, et al. What is the optimal serum level for lithium in the maintenance treatment of bipolar disorder? A systematic review and recommendations from the ISBD/IGSLI Task Force on treatment with lithium. *Bipolar Disord*. 2019;21(5):394–409.
- Benchimol EI, Smeeth L, Guttman A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med*. 2015;12(10):e1001885.
- Hernan MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758–764.
- Hsu C-W, Tsai S-Y, Tseng P-T, et al. Differences in the prophylactic effect of serum lithium levels on depression and mania in bipolar disorder: a dose-response meta-analysis. *Eur Neuropsychopharmacol*. 2022;58:20–29.
- Keck Jr. PE, Bowden CL, Meinhold JM, et al. Relationship between serum valproate and lithium levels and efficacy and tolerability in bipolar maintenance therapy. *Int J Psychiatry Clin Pract*. 2005;9(4):271–277.
- Paholpak S, Jaikasemwong S, Roumcharoenkiat A, Paholpak P, Rangseekajee P. Clinical outcome of valproate maintenance treatment in bipolar I disorder at Srinagarind Hospital. *J Med Assoc Thai*. 2014;97(4):431–438.
- Hsu CW, Carvalho AF, Tsai SY, et al. Lithium concentration and recurrence risk during maintenance treatment of bipolar disorder: multicenter cohort and meta-analysis. *Acta Psychiatr Scand*. 2021;144(4):368–378.
- Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000;11(5):561–570.
- Hsu CW, Tseng WT, Wang LJ, Yang YH, Kao HY, Lin PY. Comparative effectiveness of antidepressants on geriatric depression: real-world evidence from a population-based study. *J Affect Disord*. 2022;296:609–615.
- Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
- Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials*. 2007;8:16.
- Gugler R, von Unruh GE. Clinical pharmacokinetics of valproic acid. *Clin Pharmacokinet*. 1980;5(1):67–83.
- Loscher W. Serum protein binding and pharmacokinetics of valproate in man, dog, rat and mouse. *J Pharmacol Exp Ther*. 1978;204(2):255–261.

- 25 Nasreddine W, Dirani M, Atweh S, Makki A, Beydoun A. Determinants of free serum valproate concentration: a prospective study in patients on divalproex sodium monotherapy. *Seizure*. 2018;59:24–27.
- 26 Kang JS, Lee MH. Overview of therapeutic drug monitoring. *Korean J Intern Med*. 2009;24(1):1–10.
- 27 Fleming J, Chetty M. Therapeutic monitoring of valproate in psychiatry: how far have we progressed? *Clin Neuropharmacol*. 2006;29(6):350–360.
- 28 Taipale H, Tiihonen J. Registry-based studies: what they can tell us, and what they cannot. *Eur Neuropsychopharmacol*. 2021;45:35–37.
- 29 Anmella G, Pacchiarotti I, Cubala WJ, et al. Expert advice on the management of valproate in women with bipolar disorder at childbearing age. *Eur Neuropsychopharmacol*. 2019;29(11):1199–1212.
- 30 Kan ACO, Chan JKN, Wong CSM, Chen EYH, Chang WC. Psychotropic drug utilization patterns in pregnant women with bipolar disorder: a 16-year population-based cohort study. *Eur Neuropsychopharmacol*. 2022;57:75–85.