



PROF. WILLEM NOLEN (Orcid ID : 0000-0002-2092-9891)  
DR. MAURICIO TOHEN (Orcid ID : 0000-0001-8049-4351)  
DR. EDUARD VIETA (Orcid ID : 0000-0002-0548-0053)  
DR. RENÉ ERNST NIELSEN (Orcid ID : 0000-0002-7982-6352)

Article type : Review

**Definite draft, April 1, 2019**

**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**

W.A. Nolen <sup>a</sup>, R.W. Licht <sup>b,c</sup>, A.H. Young <sup>d</sup>, G.S. Malhi <sup>e,f,g</sup>, M. Tohen <sup>h</sup>, E. Vieta <sup>i</sup>, R.W. Kupka <sup>j</sup>, C. Zarate <sup>k</sup>, R.E. Nielsen <sup>b,c</sup>, R. Baldessarini <sup>l</sup>, E. Severus <sup>m</sup>, on behalf of the ISBD/IGSLI Task Force on the treatment with lithium.

<sup>a</sup> Department of Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, Netherlands.

<sup>b</sup> Aalborg University Hospital, Psychiatry, Aalborg, Denmark.

<sup>c</sup> Department of Clinical Medicine, Aalborg University, Aalborg, Denmark.

<sup>d</sup> Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.

<sup>e</sup> The University of Sydney, Faculty of Medicine and Health, Northern Clinical School, Department of Psychiatry, Sydney, New South Wales, Australia.

<sup>f</sup> Academic Department of Psychiatry, Royal North Shore Hospital, Northern Sydney Local Health District, St Leonards, NSW 2065 Australia.

<sup>g</sup> CADE Clinic, Royal North Shore Hospital, Northern Sydney Local Health District, St Leonards, NSW 2065 Australia.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/bdi.12805

This article is protected by copyright. All rights reserved.

<sup>h</sup> Department of Psychiatry & Behavioral Sciences, University of New Mexico Health Sciences Center, Albuquerque, NM, USA.

<sup>i</sup> Bipolar Disorder Program, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, Catalonia, Spain.

<sup>j</sup> GGZ inGeest/VU University Medical Center, Department of Psychiatry, Amsterdam; Altrecht Institute for Mental Health Care, Utrecht, The Netherlands.

<sup>k</sup> National Institute of Mental Health (NIMH), Bethesda, MD, USA.

<sup>l</sup> Department of Psychiatry, Harvard Medical School, Boston; Mailman Research Center, McLean Hospital, Belmont, MA, USA.

<sup>m</sup> Department of Psychiatry and Psychotherapy, University Hospital Carl Gustav Carus, TU Dresden, Dresden, Germany.

ORCID ID:

Gin S Malhi <http://orcid.org/0000-0002-4524-9091>

#### **Corresponding author**

Willem A Nolen MD PhD  
Department of Psychiatry  
University Medical Center Groningen  
PO Box 30.001  
9700 RB Groningen  
Netherlands  
Tel +31624207437  
Email [w.a.nolen@umcg.nl](mailto:w.a.nolen@umcg.nl)

#### **Key Words**

Lithium, serum level, bipolar disorder, maintenance treatment

## Abstract

**Aims:** To systematically review existing trials on optimal serum levels for lithium for maintenance treatment of bipolar disorder and to develop clinical recommendations.

**Methods:** Systematic literature search. Discussion of major characteristics, limitations, methodological quality and results of selected trials. Delphi survey consisting of clinical questions and corresponding statements. For statements endorsed by at least 80% of the members, consensus was considered as having been achieved.

**Results:** With strict inclusion criteria no studies could be selected, making it difficult to formulate evidence-based recommendations. After loosening the inclusion criteria seven trials were selected addressing our aims at least to some extent. Four of these studies suggest better efficacy being associated with lithium serum levels in a range above a lower threshold around 0.45/0.60 and up to 0.80/1.00 mmol/L. These findings support the outcome of the Delphi survey.

**Conclusions:** For adults with bipolar disorder there was consensus that the standard lithium serum level should be 0.60-0.80 mmol/L with the option to reduce it to 0.40-0.60 mmol/L in case of good response but poor tolerance or to increase it to 0.80-1.00 mmol/L in case of insufficient response and good tolerance. For children and adolescents there was no consensus, but the majority of the members endorsed the same recommendation. For the elderly there was also no consensus, but the majority of the members endorsed a more conservative approach: usually 0.40-0.60 mmol/L, with the option to go to maximally 0.70 or 0.80 mmol/L at ages 65-79 years, and to maximally 0.70 mmol/L over age 80 years.

## Introduction

Lithium has been licensed in Europe and North America for the long-term maintenance treatment of bipolar disorder (BD) for more than 45 years and is justifiably considered one of the first line options in major global treatment guidelines, such as the WFSBP guideline <sup>1</sup>, the NICE guideline <sup>2</sup>, the RANZCP guideline <sup>3</sup>, the Dutch guideline <sup>4</sup>, the BAP guideline <sup>5</sup>, the CINP guideline <sup>6</sup> and the CANMAT/ISBD guideline <sup>7</sup>. However, lithium has a very low therapeutic index, i.e. a low ratio between the dose (or serum level) that is associated with toxicity (mainly CNS and renal toxicity) and the dose associated with therapeutic effect. The therapeutic index for lithium is approximately 2 <sup>8</sup>.

For maintenance treatment of BD several reviews recommend differing *minimum* effective lithium serum levels ranging from 0.40 mmol/L <sup>9-11</sup> and 0.50 mmol/L <sup>12</sup>, or 0.60 mmol/L <sup>13</sup> to as high as 0.80 mmol/L <sup>14-15</sup>. In addition, some major clinical practice guidelines recommend *optimal ranges*, but these also lack consistency. For instance, the NICE guideline <sup>2</sup>, the RANZCP guideline <sup>3,16</sup> and the BAP guideline <sup>5</sup> recommend 0.60-0.80 mmol, the CANMAT/ISBD guideline <sup>7</sup> 0.60-1.00 mmol/L, and the CINP guideline <sup>6</sup> 0.60-1.20 mmol/L, whereas other guidelines, e.g. the WFSBP guideline <sup>1</sup>, do not provide specific recommendations. Finally, some reviews more specifically recommend targeting serum levels for individual patients '*on the basis of efficacy and tolerability*' without further specification <sup>15</sup> or '*primarily at 0.60-0.75/0.80 guided by clinical response and tolerability*' <sup>16</sup> within a broad range of 0.40-1.20 mmol/L <sup>10</sup>, which is also recommended in the Dutch guideline <sup>4</sup>. The major problem underlying these different recommendations is that large studies, well designed to assess the optimal serum level in maintenance treatment, specifically the serum level that best balances efficacy and tolerability for most patients, are lacking.

In clinical pharmacology recommendations for the optimal dose (or serum level) of a drug should be based on studies in which patients are randomized to different groups with various *a priori* defined fixed doses/serum levels (or narrow dose/serum level ranges) of the drug in question. This is necessary as in studies with flexible doses/serum levels or with wide dose/serum level ranges, prescribers are inclined to increase the dose/serum level especially in those patients who are not responding to (and are tolerating) the original dose/serum level, resulting in a so called channeling effect: non-responding patients are

treated with higher doses/serum levels than patients who have already responded to lower doses/serum levels<sup>17</sup> or dosing following outcome rather than outcome following dose<sup>18</sup>.

To our knowledge, the most recent systematic review on optimal serum levels for lithium in the maintenance treatment of BD by Severus et al.<sup>10</sup> identified only five studies that addressed this issue, most with small samples sizes. A further limitation of this review is that several important methodological aspects and limitations of the studies were discussed only briefly. Moreover, many clinical questions have not been systematically addressed in the literature, e.g. what is the best timing of blood sampling? Should the recommendations be different for dosing once daily versus twice daily? Should the recommendations be different for immediate release versus extended release formulations? Should the optimal serum level be different for the prevention of manic (and mixed) episodes compared to depressive episodes? Should recommendations take ethnicity/race or gender into consideration? And should recommendations for children and adolescents or for elderly patients be different from the recommendations for non-elderly adults?

Based on the need for an updated and more critical systematic review of the literature on optimal serum levels for lithium in the maintenance treatment of BD, a joint International Society for Bipolar Disorders (ISBD) and International Study Group on Lithium (IGSLI) Task Force on the Treatment with Lithium was charged to address this topic (and other lithium related topics). In addition, it was decided to develop a consensus on recommendations for optimal serum levels that also attempted to answer key clinical questions such as those outlined above, for which there apparently was no direct evidence available, using the Delphi method.

## **Method**

### *Composition of the Task Force*

The ISBD/IGSLI Task Force on Lithium (further Task Force) was comprised of global experts active in ISBD or IGSLI and furthermore selected on their contributions to the literature on lithium. The Task Force started at the annual conference of the ISBD in Seoul (2014) while a subgroup of the Task Force (WN, ES, RL, GM, AY, MT) initiated this project on optimal serum levels for lithium at the annual ISBD conference in Washington (2017) and further discussed it at the annual ISBD conference in Mexico City (2018).

### *Selection of literature*

As a first step two members of the Task Force (WN and ES on August 28, 2017) initiated a systematic literature search in PubMed with the following search terms limited to 'clinical trials' and 'reviews': 'lithium' and 'bipolar disorder' and ['serum level' or 'serum concentration' or 'plasma level' or 'plasma concentration' or 'blood level' or 'blood concentration' or 'lithium level' or 'lithium concentration'].

From the retrieved trials we initially intended to examine and discuss only those which fulfilled a priori defined inclusion criteria (see table 1). The included trials were rated for their methodological quality regarding randomization, blinding and description of withdrawals and dropouts ensuring intent to treat analysis, according to the Jadad Scale<sup>19</sup> as poor (score 0-2), moderate (score 3-4) or good (score 5).

### *Delphi method*

Anticipating that our literature review would fall short of providing clear evidence-based conclusions and definite guidance regarding optimal lithium serum levels, we concurrently developed a survey consisting of clinical questions and corresponding statements using the Delphi method<sup>20</sup>. Informed by the findings of clinical trials that have been conducted and by existing reviews and guidelines we formulated 32 statements on 16 clinical questions (see supplemental table 1) presented to all 50 Task Force members, representing 15 countries. They were asked to score all statements as 'Agree' or 'Not agree', with the option in several statements to mark a preferred additional more specified response. For statements endorsed by at least 80% of the participating members consensus was considered as having been achieved.

Two survey rounds were initially planned. In the first survey round we tested the statements and besides asking members to choose among the additional responses, we also asked for comments on the statements and on the additional responses. After having received the responses and comments it was concluded that a second survey round was not needed as the comments received did not reveal a need to change statements or responses.

### *Further procedure*

The subgroup that initiated this project consisted of WN, ES, RL, GM, AY and MT. Selection of studies and data extraction was conducted by WN and ES. The first drafts of the paper and the Delphi statements were written by WN in collaboration with ES and RL and then circulated for comments and suggestions to GM, AY and MT resulting in a second draft that was sent to all Task Force members for comments and/or discussion at (or prior to) the annual ISBD conference in Mexico City (2018). This resulted in a third draft of the paper (including introduction, the method section and results from the systematic review but still without the results of the Delphi survey and the discussion section) that was developed in the same way and again send to all members together with the first round of the Delphi survey.

With the results of the Delphi survey a fourth complete draft (now also with the discussion section) was again written by WN, ES, RL, GM, AY and MT, and then sent to all Task Force members for comments and approval. As last step the final version was send for final approval to all authors (now also including EV, RK, CZ, RN and RB), while those Task Force members who at least had completed the Delphi survey and approved the manuscript were listed as members of the Task Force.

After having received the reviewers' comments, the definite text was written by WN, ES, RL, GM, AY and MT.

## **Results**

### *Selection of literature*

The results of our literature search and the further selection is presented in figure 1.

The literature search in PubMed resulted in 128 trials and 204 reviews, of which we based on the abstracts selected 7 trials and 5 reviews for full-text reading. In addition, we added another 12 papers from reference lists of the selected papers and already known to us: 9 trials and 3 reviews.

In the first instance we excluded all 16 selected trials, as none of the trials fulfilled all 6 inclusion criteria (see above). However, we then decided to modify the inclusion criteria by loosening them (see again Table 1).

After this adjustment seven trials could be included; in chronological order: Jerram & McDonald, 1978<sup>21</sup> (further Jerram); Waters et al., 1982<sup>22</sup> (further Waters); Coppen et al., 1983<sup>23</sup> (further Coppen); Maj et al., 1986<sup>24</sup> (further Maj); Gelenberg et al., 1989<sup>25</sup> (further Gelenberg); Vestergaard et al., 1998<sup>26</sup> (further Vestergaard); and Nolen & Weisler, 2012<sup>27</sup> (further Nolen). Nine other trials remained excluded on the following basis: because of possible channeling bias<sup>28-31</sup>, reporting duplicate information<sup>32-33</sup>, post-hoc analyses of original studies with potential channeling bias<sup>34-35</sup> or addressing a different question<sup>36</sup>.

In addition, eight literature reviews were selected: Sashidharan, 1982<sup>9</sup>; Hopkins & Gelenberg, 2000<sup>13</sup>; McIntyre et al., 2001<sup>14</sup>; Sproule, 2002<sup>15</sup>; Severus et al., 2005<sup>37</sup>; Severus et al., 2008<sup>10</sup>; Wijeratne & Draper, 2011<sup>12</sup>; and Dols et al., 2016<sup>38</sup>.

#### *Description of the included trials*

Jerram<sup>21</sup> included 'just over 80' patients with bipolar (BP) or unipolar depressive (UP) disorder, of whom 51 BP (and 22 UP) patients completed the one year study, indicating that about 10 BP or UP patients dropped out. At baseline, all patients were in remission and using lithium while potential use of other psychotropic medication is not reported. Patients were randomized to three different a priori defined 12-hour lithium level ranges: < 0.50; 0.50-0.69, and ≥0.70 mmol/L. Endpoint was '*need for additional psychotropic medication and/or admission*'.

Thus, this study fulfilled four of the six original inclusion criteria (2, 3, 5 and 6). Other limitations are that the minimum and maximum lithium levels were not defined for the lower and upper lithium level ranges respectively, that the randomization procedure is unclear, that the study was not blind, that blood was collected between 12-16 hours after the last intake and that only the number of BP patients who completed the study is presented, leaving unclear how many BP patients were included and how many dropped out.

Waters<sup>22</sup> included 36 BP patients of whom 29 completed the one year study. It is uncertain whether patients were in remission, as they '*had to have reported residual endogenous mood swings which had not required hospitalization*' in the past two years, indicating that this study was not a pure recurrence prevention study. At baseline randomization all were



using lithium while potential use of other psychotropic medication is not reported. Patients were randomized to two different a priori defined lithium level ranges: 0.30-0.80 and 0.80-1.40 mmol/L. Endpoint was '*change in mood of a sufficient degree to require clinically additional pharmacological intervention [...]*'.

Thus, this study fulfilled only three of the original inclusion criteria (1, 5 and 6). Further limitations are that the timing of blood collection was not reported and that the procedure used for randomization is unclear.

Coppen<sup>23</sup> included 88 BP and UP patients of whom 72 completed the one year study, including 23 BP patients. Prior to the trial all patients already were using lithium with a mean serum level of  $0.86 \pm 0.20$  mmol/L, and an unreported number were also using antipsychotics or antidepressants. The patients were probably also not in remission as the report presents pre-trial morbidity data which were compared with data during the trial, implying that it was not a pure recurrence prevention study. At baseline patients were randomized to continue lithium at the same dose or to receive a dose reduction of about 25% or 50% resulting in three different 12-hour lithium level ranges of 0.45-0.59, 0.60-0.79 and  $\geq 0.80$  mmol/L. Outcome measure was not a relapse or recurrence but the change in score on the *Affective Morbidity Index* during the trial compared to pre-trial ratings.

Thus, this study only fulfilled two of the original inclusion criteria (5 and 6). Further limitations are that only the number of BP patients who completed the study is presented, leaving unclear how many BP patients were included and how many dropped out and that the randomization procedure is unclear.

Maj<sup>24</sup> studied 80 BP patients (69 completers) with recurrent episodes (at least one episode during the two years '*preceding the index episode and the commencement of lithium prophylaxis*'), indicating that patients were not in remission (implying that it was not a pure recurrence prevention study), but also leaving unclear whether or not patients were already using lithium prior to the trial. At baseline, patients were divided (but not randomized, personal information to ES) to four different groups (n=20) with 12-hour lithium levels ranges of 0.30-0.45, 0.46-0.60, 0.61-0.75 and 0.76-0.90 mmol/L, while potential use of other psychotropic medication is not reported. Major endpoint was '*mean number of episodes*

and mean total morbidity (expressed in months)' during the two years of the trial compared with the two years prior to the trial.

Thus, this study only fulfilled two of the original inclusion criteria (1 and 6). Another limitation is that the study was not blind.

Gelenberg<sup>25</sup> studied 94 BP-1 patients of whom 59 completed the three-year trial. All patients had to be stable for at least 2 months since the last episode and also at least 2 months tolerating lithium at serum level 0.60-1.00 mmol/L, while potential use of other psychotropic medication at start or during the trial is not reported. At baseline patients were randomized to two different a priori defined lithium level ranges: 0.40-0.60 and 0.80-1.00 mmol/L. Endpoint was a recurrence of mania or major depression, defined by DSM-III or RDC criteria.

Thus, this study fulfilled five of the original inclusion criteria (1, 2, 3, 5 and 6). A source of bias could be the abrupt lowering of the lithium levels at randomization in patients previously responding to high levels resulting in a risk of drug-discontinuation related morbidity<sup>34</sup>. Another limitation is that only the patients but not the investigators/clinicians were blind to the dosage-assignment.

Vestergaard<sup>26</sup> included 57 BP and 34 UP patients in a two-year follow-up trial. At baseline all patients were still hospitalized (indicating that they were not in remission and implying that it was not a real recurrence prevention study) and about half of them were still using antipsychotics or antidepressants. Whether (some) patients were also already using lithium remains unclear in the report. At baseline patients were randomized double blind to two different a priori defined lithium level ranges of 0.50-0.80 and 0.80-1.00 mmol/L. Endpoint was recurrence (or relapse) defined as '*re-emergence of affective disorder of a severity that rendered re-hospitalization necessary*'.

Thus, this study fulfilled only two of the original inclusion criteria (5 and 6). Further limitations are that the study was not blind and that only the number of BP patients who were included is presented, leaving unclear how many BP patients completed the study and how many dropped out.

*Nolen*<sup>27</sup> published a post-hoc analysis of a long-term study intended to test the efficacy of quetiapine<sup>39</sup>. In this study patients with BP-1 disorder with a recent manic or depressive episode who had achieved remission during  $\geq 4$  weeks of open-label treatment with quetiapine monotherapy were randomized to two-years treatment of continued quetiapine or switch to placebo or lithium. According to the protocol, the 418 patients in the lithium arm should have received a lithium dose resulting in a 12-hours lithium level of 0.60-1.20 mmol/L, guided by a '*programmed automatic system*' to alert treating clinicians to adjust dose while maintain blind to the drug (lithium, quetiapine or placebo). However, in 54 patients no blood levels were measured while in 137 patients mean lithium levels remained below the minimum predefined level of 0.60 mmol/L, probably due to a failure of the programmed automatic system to adjust the dose, rather than due to adverse events or early response (being the most important indicators of channeling). This circumstance created the possibility to compare these 137 patients with the 201 patients receiving lithium within the predefined range of 0.60-1.20 mmol/L and with the patients receiving placebo (n=404). Endpoint was a recurrence defined as the '*initiation of [psychotropic medication] to treat a mood event, hospitalization for a mood event, or a YMRS or MADRS score of  $\geq 20$* '.

Thus, this study fulfilled four of the original inclusion criteria (1, 2, 3, and 4). Besides that the composition of both lithium arms was not the result of randomization and the levels were not a priori defined, another major limitation of this study is that it was discontinued prematurely by the sponsor after an interim analysis after 300 mood events had revealed superiority of quetiapine over placebo, which resulted in loss of probably 15-20% of patients before completing the two-years follow-up. Finally, although called a 'trough' level, it is unclear from the paper whether blood was sampled after  $12 \pm 1$  hour after the last evening intake.

In summary, none of the seven trial studies described above fulfilled all original inclusion criteria (Table 2). Four studies (Waters, Maj, Gelenberg, Nolen) addressed only BP patients (original criterion 1) and the remaining three studies (Jerram, Coppen, Vestergaard) included both BP and UP patients without a complete description of the results in de BP patients. Three studies (Jerram, Gelenberg, Nolen) included only patients who were in remission (original criterion 2) and so were pure recurrence prevention studies (original criterion 3), whilst the other four studies included patients with '*residual mood swings*'

(Waters), who were hospitalized (Vestergaard) or included patients with morbidity (Coppen, Maj) and thus also examined the direct/acute effects. Only one study (Nolen) investigated solely the effect of lithium as monotherapy (original criterion 4), two studies (Coppen, Vestergaard) reported that at least some patients used antipsychotics and/or antidepressants, and four studies (Jerram, Waters, Maj, Gelenberg) did not provide information on (possible) use of concurrent medication. Five studies (Jerram, Waters, Coppen, Gelenberg, Vestergaard) were randomized (original criterion 5), implying that lithium level ranges were not affected by channeling, while the remaining two studies (Maj, Nolen) were not randomized, but it is likely that the composition of the different lithium groups in these studies was not affected by channeling (see also below). Finally, six studies (Jerram, Waters, Coppen, Maj, Gelenberg, Vestergaard) compared a priori defined lithium level ranges (original criterion 6), while the seventh study (Nolen) involved two groups with lithium level ranges defined post-hoc, as well as a group receiving placebo.

Table 3 provides information on prior treatment with lithium; treatment at randomization; details regarding treatment with lithium; and treatment with other psychotropic medication at follow-up. No study provided adequate information on all potentially relevant points. It illustrates the heterogeneity of the included studies.

#### *Quality assessment of the included studies*

Regarding JADAD criteria to rate methodological quality, five of the seven studies involved randomized assignment to particular serum concentrations of lithium. However, the randomization was described adequately in only two studies (Gelenberg, Vestergaard) and not in the remaining three studies (Jerram, Waters, Coppen); two studies were not randomized (Maj, Nolen). Three studies had a double-blind design, which was well described in one study (Nolen) but not in the other two studies (Waters, Coppen); one study was blind for patients but not for investigators/clinicians (Gelenberg); three studies were open (Jerram, Maj, Vestergaard). Finally, four studies adequately reported the number and reasons for drop-outs (Waters, Maj, Gelenberg and Nolen), but the remaining three studies did not (Jerram, Coppen, Vestergaard).

Consequently, the JADAD quality scores for the included studies varied from low, i.e. score 1 (Jerram, Maj) or 2 (Waters, Coppen, Vestergaard) to moderate, i.e. score 3 (Gelenberg, Nolen). No study received score 4 or the highest score 5.

#### *Clinical outcomes of the included studies*

Table 4 summarizes the most important clinical outcomes of the included studies. Significant differences between lithium serum level ranges are indicated by green (better outcomes) versus red (poorer outcomes), while outcomes which were not significantly different are indicated by blue.

The findings of the seven studies are summarized in Table 4. Taken together, four studies (Waters, Maj, Gelenberg and Nolen) indicate better long-term efficacy in the prevention of any mood episode with mean lithium serum levels in a range above a lower threshold around 0.45/0.60 and up to 0.80/1.00 mmol/L. The other three studies (Jerram, Coppen, Vestergaard) did not find a significant difference between higher versus lower lithium levels.

In addition, four studies also looked at possible differences between the prevention of manic versus depressive recurrences. Only one study (Maj) suggested a lower threshold for the prevention of manic symptoms (around 0.45 mmol/L) than for depressive symptoms (around 0.60 mmol/L). Another study (Nolen) suggests the same threshold of 0.60 mmol/L for both the prevention of manic and depressive recurrences. The other two studies (Waters, Vestergaard) also found no difference in polarity of response and lithium levels.

Finally, three studies, addressed the association of lithium levels with adverse effects. Two studies, reported more side effects with levels 0.45-0.60 and higher versus 0.30-0.45 (Maj) and with levels 0.80-1.00 versus 0.40-0.60 (Gelenberg), while the third study (Nolen) did not report different drop-outs rates due to side effects between the three groups (lithium <0.60 mmol/L, lithium 0.60-1.20 mmol/L and placebo).

## Delphi survey

The Delphi survey was completed by 33 of the 50 (66.6%) Task Force members. Many statements received additional comments, most frequently that evidence regarding a statement or a specified response was lacking.

There was consensus on recommendations regarding timing of blood sampling to measure lithium serum levels during the day and on whether recommendations for optimal serum lithium levels should be the same with once daily dosing compared to twice daily dosing, as well as with different formulations of lithium medication (immediate release versus extended release).

There was a lack of consensus on the overall statement *“The higher the lithium serum level, the higher the likelihood of a preventative response”* (endorsed by only 54.5%) while several members commented that serum levels in the lower range (below 0.60 mmol/L) are associated with efficacy but in the higher range (above 1.00 mmol/L) with more adverse effects and toxicity. In addition, there was no consensus on the statement *“There is a minimum serum cut-off level below which essentially no patients are likely to experience a preventative effect”* (endorsed by 72.7%) while there was a wide range of specified responses regarding for a specific cut-off level (0.30, 0.40, 0.50 or 0.60 mmol/L, endorsed by 16.7, 41.7, 16.7 or 20.8%, respectively). In contrast, there was consensus on the statement *“There is a maximum lithium serum level that should ideally never be exceeded because of the risk of severe intolerance and/or intoxication”* (endorsed by 93.9%) but lack of agreement on a specific upper limit to recommend: most frequently supported limits being 1.00 and 1.20 mmol/L, endorsed by 35.5 and 38.7%, respectively.

The major recommendations including seven statements endorsed by at least 80% of the members (consensus) are summarized in Table 5.

Regarding recommended standard lithium serum levels in the maintenance treatment there was only consensus on the recommendation for adults, but not for children <12 years, adolescents 12-17 years, the elderly 65-79 and the elderly ≥80 years. However, the most frequently endorsed minimum level was 0.40 mmol/L for all age groups, whereas the maximum level varied with age; 1.00 mmol/L for adults, children and adolescents and between 0.70 and 1.00 mmol/L for the elderly.

Finally, for the prevention of specific episodes there was consensus for recommending similar levels for the prevention of mixed episodes versus manic episodes (endorsed by 87.9%) but not for recommending similar levels for the prevention of manic versus depressive episodes and for the prevention of hypomanic or subthreshold depressive episodes versus manic and depressive episodes, respectively.

## **Discussion**

The key conclusion from this review (and the major limitation) is that there is a lack of well-designed, double-blind randomized dose finding (or serum level finding) studies assessing optimal lithium serum level in the maintenance treatment of BP. This conclusion is truly remarkable given that lithium has been available for the treatment of BP since it was licensed in the early 1970s. This paucity of data makes it difficult if not impossible to formulate confident evidence-based recommendations for the dosing of lithium based on monitoring its serum level. Based on the initial inclusion criteria for our systematic review we were unable to identify any study that fulfilled all six original inclusion criteria for our review. Only after loosening these criteria, we identified seven studies that addressed, at least to some extent, the questions we posed at the outset, and outlined in the introduction.

### *Selection of studies*

A major issue in the selection of studies was that we wanted to compare similar groups of patients treated with lithium at different serum levels while the composition of the groups should not be influenced by channeling of non-responding patients with good tolerance to higher lithium level groups and early responding patients or patients with low tolerance to lower lithium level groups. Therefore, we initially intended to select only studies with a randomized design (criterion 5) in which patients were assigned to different treatment regimens with lithium resulting in different a priori defined fixed lithium serum level ranges (criterion 6). Five studies (Jerram, Waters, Coppen, Gelenberg, Vestergaard) were randomized. The remaining two studies (Maj, Nolen) were not randomized, but it is likely that the composition of the different lithium groups in these studies was not seriously

affected by channeling. In the Maj study the patients were allocated in an open way to four different groups at baseline, while in the Nolen study due to an organizational failure (see above) a large group of patients remained below the predefined minimum level of 0.60 mmol/L, resulting in the possibility to compare two groups below or above this level.

Nevertheless, prior to start of follow-up some channeling may have occurred in the pre-randomization open treatment phase in which all patients (Jerram, Waters, Coppén, Gelenberg) or possibly some patients (Vestergaard) were already using lithium. In particular, if lithium had been given prophylactically with at least some success regarding efficacy and tolerability (Jerram, Waters, Coppén) the probability that relevant channeling with regard to the topic of the trial, namely the prophylactic efficacy of lithium, may have occurred, cannot be dismissed. Similarly, an unknown number of patients in the Maj study were already using lithium prior to allocation to one of the groups/start of follow-up. This may have resulted in these studies in some selection bias: of poorer responding and better tolerating patients in the groups with higher lithium serum levels. In fact, the only study in which we can be confident that channeling did not happen regarding lithium levels prior to start of follow-up, is the (non-randomized) study by Nolen, as this is the only study where patients were not already using lithium at start of follow-up.

An example of the excluded studies is the study by Tohen et al.<sup>31</sup> which compared in a randomized controlled, double-blind design two groups of patients treated with olanzapine or lithium monotherapy after an open label phase in which they were treated for a manic episode with the combination of olanzapine and lithium. During the open label phase, lithium serum levels were targeted at 0.60-1.20 mmol/L, depending on overall response to and tolerance of the combination. Only responders were randomized to the double blind maintenance phase (enriched sampling). However, some patients remained below 0.60 mmol/L, for instance when they did not tolerate lithium (combined with olanzapine) at the intended level. In a post-hoc analysis of this study by Severus et al.<sup>35</sup> three subgroups were compared within the group randomized to continuing lithium at the same level as in the open treatment phase (<0.60, 0.60-0.79 and  $\geq$ 0.80 mmol/L). Although the original study protocol (31) prevented channeling in the double blind phase, channeling could have occurred during the open label phase as some patient had levels lower than 0.60 mmol/L and no new groups were constructed at the start of follow-up in the original study. Thus, the design differed from the Maj study, that did not include an acute phase while



most patients were allocated open label to different lithium levels from the start of treatment with lithium, and from the Nolen study (see above). However, as the post-hoc analysis (35) used proportional hazards Cox regression models and marginal structural models (MSMs) controlling for, among other variables, olanzapine and lithium dose at randomization as well as variables related to the disease severity and index episode features at baseline as well as variables related to the previous course of illness, the significance of the results of this post-hoc analysis may in fact be comparable to those of a randomized controlled trial .

*What is the optimal serum level for lithium in the maintenance treatment of bipolar disorder?*

The main finding of our review is that four of the seven included studies reported significantly fewer recurrences with lithium levels 0.80-1.40 mmol/L versus 0.30-0.80 (Waters); 0.61-0.90 versus 0.30-0.45 (Maj); 0.80-1.00 versus 0.40-0.60 (Gelenberg); and 0.60-1.00 versus <0.60 mmol/L (Nolen). In addition, the latter study found superiority of 0.60-1.20 versus placebo but not of <0.60 versus placebo. Taken together, these studies suggest better efficacy being associated with lithium levels in a therapeutic range with a lower threshold around 0.45/0.60 and up to 0.80/1.00 mmol/L. The other three studies (Jerram, Coppen, Vestergaard) did not find a significant difference between higher versus lower lithium levels.

There is also some additional information from some of the above studies. Hullin<sup>40</sup> performed a further second year follow-up of Jerram study of the combined group patients (n=68) with BP or UP. In this second year 5/28 patients (17.8%) with lithium levels of 0.60-1.00 mmol/L and 4/27 patients (14.8%) of those with levels between 0.40 and 0.59 had a recurrence. In contrast, in the group maintained on lithium levels between 0.25 and 0.39 mmol/L, 8/13 patients (61.5%) had a recurrence. Despite that no separate results were presented for the BP sample, this finding suggests that the cutoff for an effective lithium level may be around 0.40 mmol/L. In contrast however, a further subdivision of the two lithium groups in the Nolen study revealed that the Hazard ratio's (HR) versus placebo were not significant for lithium <0.40 mmol/L (HR 0.66; 95% CI: 0.42-1.03) and 0.40 to <0.60

mmol/L (HR 1.05; CI: 0.67-1.63) while they were significant for lithium 0.60 to <0.80 (HR 0.35; CI: 0.23-0.52) and 0.80 to <1.00 (HR 0.24; CI: 0.14-0.42), but not for lithium 1.00-1.20 (HR 0.50; CI: 0.21-1.22). These findings suggest that the cutoff is more likely to be around 0.60 mmol/L than around 0.40 mmol/L. In general, while lithium levels <0.60 mmol/L appear not as effective as levels >0.60 mmol/L, there may be some patients for whom levels between 0.40 and 0.60 mmol/L are effective. Therefore, this may be an option for those who do not tolerate levels >0.60 mmol/L as well as for patients with impaired renal function or with increased risk of intoxication, e.g. in the elderly (see below). In addition, levels <0.40 mmol/L do not seem to work at all. Therefore, there is no rationale for prescribing lithium at levels below 0.40 mmol/L.

Finally, the Nolen study also suggests no big additional benefit of levels 0.80 to <1.0 mmol/L over 0.60 to <0.85 mmol/L, which is in line with the Maj study which did not find a difference between levels 0.76-0.90 mmol/L and 0.61-0.75 mmol/L. Although we should recognize the possibility of a Type 2 error, i.e. that with larger numbers a difference may have been found, these studies suggest an upper effective cut-off around 0.80 mmol/L. Nevertheless, again there may be some patients for whom higher levels up to 1.00 mmol/L are more effective and well tolerated. In addition, however, there are no data supporting the notion that going beyond 1.00 mmol/L is associated with any further benefit, possibly with exception of its use in the treatment of acute manic episodes (not further being discussed in this paper). In addition, the upper limit of 0.80 or 1.00 mmol/L is of course also defined by the increased risk of side effects and toxicity at higher lithium serum levels <sup>41</sup>.

Taking all these above data into account, they do support the outcome of the Delphi survey: that *for adults (18-65 years) the standard therapeutic lithium serum level in the maintenance treatment of bipolar disorder should be 0.60-0.80 mmol/L with the option to reduce the level to the low therapeutic level of 0.40-0.60 mmol/L in case of sufficient response but poor tolerance or to increase the level to the high therapeutic level of 0.80-1.00 mmol/L in case of insufficient response and good tolerance* (figure 2). The recommendation for the standard therapeutic level of 0.60-0.80 mmol/L is in line with the NICE guideline <sup>2</sup>, the RANZCP guideline <sup>3</sup>, the BAP guideline <sup>5</sup> and the 'Lithiumeter 2.0' <sup>16</sup>.

*What is the best timing of blood sampling?*

No study was found that addressed this question but most of the included studies followed the generally accepted recommendation to sample blood  $12\pm 1$  hours after the last (i.e. evening) intake. There was also consensus on this practice among the Task Force members for patients taking lithium twice daily or taking lithium once daily in the evening. However, there was no consensus on whether to sample in the evening or in the next morning for patients taking lithium once daily in the morning, a dosing schedule probably hardly applied in clinical practice.

*Should the recommendations be different for dosing once daily versus twice daily and for immediate release versus extended release formulations?*

There is pharmacogenetic evidence that a total daily dose of lithium given once daily results in a 10-15% higher  $12\pm 1$  hours trough level than when given twice daily<sup>42</sup>. Nevertheless, there was consensus among the Task Force members that one could follow the same recommendations for optimal lithium serum levels for both dosing strategies. Similarly, the members also agreed that the recommendations should be the same for different formulations (immediate release versus extended release).

An important clinical implication of this approach is that the relative small changes in  $12\pm 1$  hours trough levels caused by changes in dosing schedule or lithium formulation, should not necessarily lead to adjustments of the aimed lithium levels. In the absence of solid evidence to support an alternative approach, this simple pragmatic approach is reasonable and also compatible with the likely intra-individual variations that inevitably occur with respect to subsequent lithium levels obtained with a stable dose. In this context, it should be born in mind that the average lithium level in a given individual with a constant renal function and a constant intake of fluid and sodium entirely depends on the total daily dose<sup>8</sup>.

*Should the optimal serum level be different for the prevention of manic (and mixed) episodes compared to depressive episodes?*

While lithium is effective in the prevention of both manic and depressive episodes<sup>43</sup> there is the question whether same lithium levels should be applied. The four studies (Waters, Maj, Vestergaard, Nolen) that addressed this issue did not find evidence that lithium serum levels for the prevention of specific episodes should be different than for other episodes. In the Delphi survey there was only consensus that the recommendations for the prevention of mixed episodes should be the same as for the prevention of manic episodes. The recommendations for other comparisons (i.e. manic versus depressive episodes; hypomanic versus manic episodes; subsyndromal depressive versus depressive episodes) did not result in consensus, although the majority of members supported the statements that the levels should be the same for the prevention of all types of episodes.

*Should recommendations take ethnicity or gender into consideration?*

We did not find any study addressing these topics in relation to optimal lithium serum levels. Therefore, we also did not formulate recommendations on these issues.

A specific group concerns pregnant women. Pregnant women can be treated with lithium, but the recommendations in this paper are not applicable for them. For more specific information, we refer to a recent paper by Wesseloo et al.<sup>44</sup>.

*Should recommendations for children and adolescents or for elderly patients be different from the recommendations for (younger) adults?*

Two of the selected studies (Coppen, Vestergaard) did not provide information on whether they had included patients below 18 years or above 65 years and two other studies did not include patients below 18 years or above 65 years (Waters, Maj). The remaining three studies did only include patients above 18 years including elderly patients, but did not provide separate information on the outcome of the elderly patients. Thus, we can conclude that evidence on optimal lithium serum levels in children, adolescents and the elderly (see also<sup>38</sup>) is lacking.

This article is protected by copyright. All rights reserved.

This is also reflected by outcome of the Delphi survey regarding the endorsement of the statements on optimal lithium serum levels in the various age groups of which only the statement on adults (18-65 years) was endorsed by at least 80% (in consensus) of the Task Force members. However, for all age groups there was consensus that the minimum effective lithium level for all age groups should be 0.40 mmol/L, while in adults the standard therapeutic level should be between 0.60-0.80 mmol/L within a broader range of 0.40-1.00 (figure 2). For children (< 12 years) and adolescents (12-17 years) there was no consensus, but the majority of the Task Force members endorsed the same recommendation as for adults (see also figure 2). For the elderly there was also no consensus. However, the majority of the members endorsed the recommendation that the standard therapeutic level should be 0.40-0.60 mmol/L, with the option to go to maximally 0.70 of 0.80 mmol/L in elderly 65-79 years and to maximally 0.7 mmol/L in elderly 80 years and older (figure 3), which is in line with the recommendations on the use of lithium in the elderly from the ISBD Task Force on Older Adults with Bipolar Disorder: that although *'the balance between lithium toxicity and efficacy has not been studied in older patients and the recommendations are mainly based on clinical judgement and fear of drug-related adverse events'*<sup>38</sup>, the upper limit should be lower than in younger adults<sup>45</sup>.

#### *Practical considerations regarding lithium level monitoring*

Due to lithium's low therapeutic index, its serum levels should always be monitored during treatment, at least when treatment is initiated, after each dose increase and every 3-6 months thereafter. The blood samples for measuring the lithium level must be drawn at steady state occurring after 5-6 times the half-life of lithium, which is approximately 24 hours in the non-elderly with normal renal function. Thus, the lithium level should be measured after 5-6 days at constant daily dosing. In the elderly or in patients with impaired renal function, the time to steady state increases, implying that the lithium level should also be measured after 10-12 days. In individual patients treated with a given schedule with a given lithium formulation, at steady state, the ratio of daily lithium dose and the 12 hours lithium level is constant. An additional requirement for safe use of lithium is to monitor renal function at baseline and also every 3-6 months.

Accepted Article

Finally, it should be taken into account that subjects with a relatively high renal function, in particular young men, may have an unexpectedly low 12-hours lithium level despite a relatively high daily lithium dose due to relative large variations in lithium levels over the day. If the dose is increased to obtain a lithium level within the recommended range, such patients may become intoxicated <sup>46</sup>.

## Conclusion

Based on our review and the responses obtained in our Delphi survey we recommend that the standard serum level in the maintenance treatment of adults with BD should be 0.60-0.80 mmol/L with the option to reduce the level to 0.40-0.60 mmol/L in case of sufficient response but poor tolerance, or to increase the level to 0.80-1.00 mmol/L in case of insufficient response and good tolerance. For children and adolescents the majority of the Task Force members endorsed the same recommendation, while for the elderly the majority endorsed a more conservative approach: usually 0.40-0.60 mmol/L, with the option to go to maximally 0.70 or 0.80 mmol/L at ages 65-79 years, and to maximally 0.70 mmol/L over age 80 years. Sampling should be in the morning 12±1 hours after intake of the evening dose. Tentative recommendations for optimal serum lithium levels should be the same with once and twice daily dosing, and with immediate release formulations compared to extended release formulations.

While the above recommendations reflect the available evidence, more high-quality research is still needed to provide more refined advices.

## Literature

- <sup>1</sup> Grunze H, Vieta E, Goodwin GM, Bowden C, Licht RW, Möller HJ, Kasper S; WFSBP Task Force on Treatment Guidelines for Bipolar Disorders. 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: 154-219.
- <sup>2</sup> NICE-2014. Bipolar Disorder: Assessment and Management (NICE2014). Available at: <https://www.nice.org.uk/guidance/cg185>
- <sup>3</sup> Malhi, G. S., Bassett, D., Boyce, P., Bryant, R., Fitzgerald, P. B., Fritz, K., et al. (2015). Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. *Aust N Z J Psychiatry*. 2015; 49: 1087-1206.
- <sup>4</sup> Kupka R, Goossens P, Van Bendegem M, Daemen P, Daggenvoorde T, Daniëls M, Dols A, Hillegers M, Hoogelander A, Ter Kulve E, Peetoom T, Schulte R, Stevens A, Van Duin D. Multidisciplinaire richtlijn bipolaire stoornissen, derde herziene versie (2015) Trimbos Instituut, Utrecht. Available at: <https://www.nvvp.net/stream/richtlijn-bipolaire-stoornissen-2015>.
- <sup>5</sup> Goodwin GM, Haddad PM, Ferrier IN, Aronson JK, Barnes T, Cipriani A, Coghill DR, Fazel S, Geddes JR, Grunze H, Holmes EA, Howes O, Hudson S, Hunt N, Jones I, Macmillan IC, McAllister-Williams H, Miklowitz DR, Morriss R, Munafò M, Paton C, Saharkian BJ, Saunders K, Sinclair J, Taylor D, Vieta E, Young AH. Evidence-based guidelines for treating bipolar disorder: Revised third edition recommendations from the British Association for Psychopharmacology. *J Psychopharmacol*. 2016; 30: 495-553.
- <sup>6</sup> Fountoulakis KN, Yatham L, Grunze H, Vieta E, Young A, Blier P, Kasper S, Moeller HJ. The International College of Neuro-Psychopharmacology (CINP) Treatment Guidelines for Bipolar Disorder in Adults (CINP-BD-2017), Part 2: Review, Grading of the Evidence, and a Precise Algorithm. *Int J Neuropsychopharmacol*. 2017; 20: 121-179.
- <sup>7</sup> Yatham LN, Kennedy SH, Parikh SV, Schaffer A, Bond DJ, Frey BN, Sharma V, Goldstein BI, Rej S, Beaulieu S, Alda M, MacQueen G, Milev RV, Ravindran A, O'Donovan C, McIntosh D, Lam RW, Vazquez G, Kapczinski F, McIntyre RS, Kozicky J, Kanba S, Lafer B, Suppes T, Calabrese JR, Vieta E, Malhi G, Post RM, Berk M. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD)

2018 guidelines for the management of patients with bipolar disorder. *Bipolar Disord.* 2018; 20: 97-170.

<sup>8</sup> Licht RW. Lithium: still a major option in the management of bipolar disorder. *CNS Neurosci Ther.* 2012; 18: 219-226.

<sup>9</sup> Sashidharan SP. The relationship between serum lithium levels and clinical response. *Ther Drug Monit.* 1982; 4: 249-264.

<sup>10</sup> Severus WE, Kleindienst N, Seemüller F, Frangou S, Möller HJ, Greil W. What is the optimal serum lithium level in the long-term treatment of bipolar disorder--a review? *Bipolar Disord.* 2008; 10: 231-237.

<sup>11</sup> Malhi GS, Gessler D, Outhred T. The use of lithium for the treatment of bipolar disorder: Recommendations from clinical practice guidelines. *J Affect Disord.* 2017; 217: 266-280.

<sup>12</sup> Wijeratne C, Draper B. Reformulation of current recommendations for target serum lithium concentration according to clinical indication, age and physical comorbidity. *Aust N Z J Psychiatry.* 2011; 45: 1026-1032.

<sup>13</sup> Hopkins HS, Gelenberg AJ. Serum lithium levels and the outcome of maintenance therapy of bipolar disorder. *Bipolar Disord.* 2000; 2(3 Pt 1): 174-179.

<sup>14</sup> McIntyre RS, Mancini DA, Parikh S, Kennedy SH. Lithium revisited. *Can J Psychiatry.* 2001; 46: 322-327.

<sup>15</sup> Sproule B. Lithium in bipolar disorder: can drug concentrations predict therapeutic effect? *Clin Pharmacokinet.* 2002; 41: 639-660.

<sup>16</sup> Malhi, G.S., Gershon, S., Outhred, T. Lithiummeter: Version 2.0. *Bipolar Disord.* 2016; 18: 631-641.

<sup>17</sup> Leufkens HG, Urquhart J. Variability in patterns of drug usage. *J Pharm Pharmacol.* 1994; 46 Suppl 1: 433-437.

<sup>18</sup> Tohen M, Wateraux CM, Tsuang MT. Outcome in Mania. A 4-year prospective follow-up of 75 patients utilizing survival analysis. *Arch Gen Psychiatry* 1990; 47: 1106-1111.

<sup>19</sup> Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials.* 1996; 17: 11-12.

<sup>20</sup> Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ.* 1995; 311: 376-380.



- 21 Jerram TC, McDonald R. Plasma lithium control with particular reference to minimum effective levels. In Johnson FN, Johnson S (Eds.) *Lithium in medical practice*. Lancaster: MTP Press Ltd. 1978: 407-413.
- 22 Waters B, Lapierre Y, Gagnon A, Cahudhry R, Tremblay A, Sarantidis D, Gray R. Determination of the optimal concentration of lithium for the prophylaxis of manic-depressive disorder. *Biol Psychiatry*. 1982; 17: 1323-1329.
- 23 Coppen A, Abou-Saleh M, Milln P, Bailey J, Wood K. Decreasing lithium dosage reduces morbidity and side-effects during prophylaxis. *J Affect Disord*. 1983; 5: 353-362.
- 24 Maj M, Starace F, Nolfi G, Kemali D. Minimum plasma lithium levels required for effective prophylaxis in DSM III bipolar disorder: a prospective study. *Pharmacopsychiatry*. 1986; 19: 420-423.
- 25 Gelenberg AJ, Kane JM, Keller MB, Lavori P, Rosenbaum JF, Cole K, Lavelle J. Comparison of standard and low serum levels of lithium for maintenance treatment of bipolar disorder. *N Engl J Med*. 1989; 321: 1489-1493.
- 26 Vestergaard P, Licht RW, Brodersen A, Rasmussen NA, Christensen H, Arnglim T, Grønvall B, Kristensen E, Poulstrup I. Outcome of lithium prophylaxis: a prospective follow-up of affective disorder patients assigned to high and low serum lithium levels. *Acta Psychiatr Scand*. 1998; 98: 310-315.
- 27 Nolen WA, Weisler RH. The association of the effect of lithium in the maintenance treatment of bipolar disorder with lithium plasma levels: a post hoc analysis of a double-blind study comparing switching to lithium or placebo in patients who responded to quetiapine (Trial 144). *Bipolar Disord*. 2013; 15: 100-109.
- 28 Sashidharan SP, McGuire RJ, Glen AI. Plasma lithium levels and therapeutic outcome in the prophylaxis of affective disorders: a retrospective study. *Br J Psychiatry*. 1982; 140: 619-622.
- 29 Goodnick PJ, Fieve RR. Plasma lithium level and interepisode functioning in bipolar disorder. *Am J Psychiatry*. 1985; 142: 761-762.
- 30 Kleindienst N, Severus WE, Möller HJ, Greil W. Is polarity of recurrence related to serum lithium level in patients with bipolar disorder? *Eur Arch Psychiatry Clin Neurosci*. 2005; 255: 72-74.
- 31 Tohen M, Greil W, Calabrese JR, Sachs GS, Yatham LN, Oerlinghausen BM, Koukopoulos A, Cassano GB, Grunze H, Licht RW, Dell'Osso L, Evans AR, Risser R, Baker RW, Crane H,

Dossenbach MR, Bowden CL. Olanzapine versus lithium in the maintenance treatment of bipolar disorder: a 12-month, randomized, double-blind, controlled clinical trial. *Am J Psychiatry*. 2005; 162: 1281-1290.

<sup>32</sup> Keller MB, Lavori PW, Kane JM, Gelenberg AJ, Rosenbaum JF, Walzer EA, Baker LA. Subsyndromal symptoms in bipolar disorder. A comparison of standard and low serum levels of lithium. *Arch Gen Psychiatry*. 1992; 49: 371-376.

<sup>33</sup> Solomon DA, Ristow WR, Keller MB, Kane JM, Gelenberg AJ, Rosenbaum JF, Warshaw MG. Serum lithium levels and psychosocial function in patients with bipolar I disorder. *Am J Psychiatry*. 1996; 153: 1301-1307.

<sup>34</sup> Perlis RH, Sachs GS, Lafer B, Otto MW, Faraone SV, Kane JM, Rosenbaum JF. Effect of abrupt change from standard to low serum levels of lithium: a reanalysis of double-blind lithium maintenance data. *Am J Psychiatry*. 2002; 159: 1155-1159.

<sup>35</sup> Severus WE, Lipkovich IA, Licht RW, Young AH, Greil W, Ketter T, Deberdt W, Tohen M. In search of optimal lithium levels and olanzapine doses in the long-term treatment of bipolar I disorder. A post-hoc analysis of the maintenance study by Tohen et al. 2005. *Eur Psychiatry*. 2010; 25: 443-449.

<sup>36</sup> Severus WE, Kleindienst N, Evoniuk G, Bowden C, Möller HJ, Bohus M, Frangou S, Greil W, Calabrese JR. Is the polarity of relapse/recurrence in bipolar-I disorder patients related to serum lithium levels? Results from an empirical study. *J Affect Disord*. 2009; 115: 466-470.

<sup>37</sup> Severus WE, Grunze H, Kleindienst N, Frangou S, Moeller HJ. Is the prophylactic antidepressant efficacy of lithium in bipolar I disorder dependent on study design and lithium level? *J Clin Psychopharmacol*. 2005; 25: 457-462.

<sup>38</sup> Dols A, Kessing LV, Strejilevich SA, Rej S, Tsai SY, Gildengers AG, Almeida OP, Shulman KI, Sajatovic M; International Society for Bipolar Disorders Task Force for Older Adults with Bipolar Disorder. Do current national and international guidelines have specific recommendations for older adults with bipolar disorder? A brief report. *Int J Geriatr Psychiatry*. 2016; 31: 1295-1300.

<sup>39</sup> Weisler RH, Nolen WA, Neijber A, Hellqvist A, Paulsson B; Trial 144 Study Investigators. Continuation of quetiapine versus switching to placebo or lithium for maintenance treatment of bipolar I disorder (Trial 144: a randomized controlled study). *J Clin Psychiatry*. 2011; 72: 1452-1464.

- 40 Hullin RP. Minimum serum lithium levels for effective prophylaxis. In: Johnson FN ed. Handbook of Lithium Therapy. Lancaster: MTP Press Limited. 1980: 243–247.
- 41 Wilting I, Heerdink ER, Mersch PP, Den Boer JA, Egberts AC, Nolen WA. Association between lithium serum level, mood state, and patient-reported adverse drug reactions during long-term lithium treatment: a naturalistic follow-up study. *Bipolar Disord*. 2009; 11: 434-440.
- 42 Amdisen A. Serum level monitoring and clinical pharmacokinetics of lithium. *Clin Pharmacokinet*. 1977; 2: 73-92.
- 43 Severus E, Taylor MJ, Sauer C, Pfennig A, Ritter P, Bauer M, Geddes JR. Lithium for prevention of mood episodes in bipolar disorders: systematic review and meta-analysis. *Int J Bipolar Disord*. 2014; 2: 15.
- 44 Wesseloo R, Wierdsma AI, van Kamp IL, Munk-Olsen T, Hoogendijk WJG, Kushner SA, Bergink V. Lithium dosing strategies during pregnancy and the postpartum period. *Br J Psychiatry*. 2017; 211: 31-36.
- 45 Shulman KI, Almeida OP, Herrmann N, Schaffer A, Strejilevich SA, Paternoster C, Amodeo S, Dols A, Sajatovic M. Delphi survey of maintenance lithium treatment in older adults with bipolar disorder: An ISBD task force report. *Bipolar Disord*. 2018, Oct 30.
- 46 Amdisen A. Lithium neurotoxicity – the reliability of serum lithium measurements. *Hum Psychopharmacol*. 1990; 5: 281-285.

## Legends to figures and tables

**Figure 1.** Prisma flow diagram of the selection of papers.

**Figure 2.** Recommendation for optimal lithium serum levels in the maintenance treatment of bipolar disorder: endorsed in consensus by the Task Force members for adults (18-65 years) and by the majority of the members for children (< 12 years) and adolescents (12-17 years).

**Figure 3.** Recommendation for optimal lithium serum levels in the maintenance treatment of bipolar disorder in elderly patients, endorsed by the majority of the Task Force members for elderly patients (65 years and older).

### Legends to tables 1-4 (see attachment)

**Table 1.** Overview of original inclusion criteria (see method) and modified inclusion criteria (see results)

**Table 2.** Original inclusion criteria met by the included studies

BP = Bipolar disorder; UP = Unipolar depressive disorder; APs = Antipsychotics; ADs = Antidepressants; Green = Criterion met; Orange = Unclear; Red = Criterion not met

**Table 3.** Treatment characteristics prior to the study, at randomization/start of the study and during follow-up

APs = Antipsychotics; ADs = Antidepressants; NA = Not applicable

**Table 4.** Summary of major clinical outcomes.

\* = Reconstructed with drop-out data; \*\* = One episode not stated;

HR = Hazard ratio; PP = Prior vs. Post randomization; CPRS = Comprehensive Psychiatric Rating Scale; SEs = Side effects; -> disc. = Led to discontinuation;

Significances: Green = Better outcome vs. Red = Poorer outcome; Blue = Not significant;

**Table 5:** Recommendations from the Delphi survey regarding lithium serum levels in the maintenance treatment of bipolar disorder endorsed by at least 80% (consensus) of the participating members of the task force, or by less than 80% (no consensus).

Domain	Recommendation
Timing of blood sampling to measure lithium serum levels	<p><b>Consensus</b></p> <ol style="list-style-type: none"> <li>1 With twice daily dosing: sampling should be in the morning 12±1 hours after intake of the (last) evening dose and before the morning dose (endorsed by 93.5%)</li> <li>2 With once daily dosing in the evening: sampling should be in the morning, 12±1 hours after intake of the (single) evening dose (endorsed by 93.5%)</li> </ol> <p><b>No consensus</b></p> <ul style="list-style-type: none"> <li>- With once daily dosing in the morning: sampling should be in the evening, 12±1 hours after intake of the (single) morning dose (endorsed by 74.2%)</li> </ul>
Once daily dosing versus twice daily dosing	<p><b>Consensus</b></p> <ol style="list-style-type: none"> <li>3 The recommendations for optimal serum lithium levels with once daily dosing of lithium should be the same as compared to twice daily dosing (endorsed by 90.9%)</li> </ol>
Immediate release formula versus extended release formulations	<p><b>Consensus</b></p> <ol style="list-style-type: none"> <li>4 The recommendations for optimal serum lithium levels with immediate release formulations should be the same as compared to extended release formulations (endorsed by 93.9%)</li> </ol>
Maximum lithium serum level	<p><b>Consensus</b></p> <ol style="list-style-type: none"> <li>5 There is a maximum serum lithium level that should ideally never be exceeded because of the risk of severe intolerance and/or intoxication (endorsed by 93.9%)</li> </ol> <p><b>No consensus</b></p> <ul style="list-style-type: none"> <li>- On specific maximum level</li> <li>- Most frequently mentioned: 1.00 mmol/L (endorsed by 35.5%) and 1.20 mmol/L (endorsed by 38.7%)</li> </ul>
Recommendations for optimal serum lithium levels in different age groups	<p><b>Consensus</b></p> <p><b><i>For adults (18-65 years):</i></b></p> <ol style="list-style-type: none"> <li>6 The standard serum level should be 0.60-0.80 mmol/L with the option to reduce the level to 0.40-0.60 mmol/L in case of poor tolerance or to increase the level to 0.80-1.00 mmol/L in case of insufficient response (endorsed by 93.9%)</li> </ol>

**No consensus**

- **For children (<12 years):**
  - The standard level should be 0.60-0.80 mmol/L, with the option to reduce the level to 0.40-0.60 mmol/L in case of poor tolerance or to increase the level to 0.80-1.00 mmol/L in case of insufficient response (endorsed by 57.6%)
- **For adolescents (12-17 years):**
  - The standard level should be 0.60-0.80 mmol/L, with the option to reduce the level to 0.40-0.60 mmol/L in case of poor tolerance or to increase the level to 0.80-1.00 mmol/L in case of insufficient response (endorsed by 75.8%)
- **For the elderly (65-79 years):**
  - The standard serum lithium level should be 0.40 - 0.60 mmol/L, with the option to increase to 0.60 - 0.80 mmol/L or to even 0.80 -1.00 mmol/L in case of insufficient response (unless there are somatic contra-indications and with close monitoring of emergent side effects) (endorsed by 63.6%)
- **For the elderly (80 years and older):**
  - The standard serum lithium level should be 0.40 - 0.60 mmol/L, with the option to increase to 0.60 - 0.70 mmol/L (and not higher) in case of insufficient response (unless there are somatic contra-indications and with close monitoring of emergent side effects) (endorsed by 48.5%)
  - The standard serum lithium level should be 0.40 - 0.60 mmol/L, with the option to increase to 0.60 - 0.80 mmol/L (and not higher) in case of insufficient response (unless there are somatic contra-indications and with close monitoring of emergent side effects) (endorsed by 48.5%)

---

Optimal lithium serum levels for the prevention of specific episodes

**Consensus**

- 7 The recommendations for the prevention of episodes of mania or depression with mixed features should be the same as the recommendations to prevent manic recurrences (endorsed by 87.9%)

**No consensus**

- The recommendations for the prevention of manic recurrences should be the same as the recommendations to prevent depressive recurrences (endorsed by 66.7%)
- The recommendations for the prevention of hypomanic recurrences should be the same as the recommendations to prevent manic recurrences (endorsed by 78.8%)
- The recommendations for the prevention of subsyndromal depressive recurrences should be the same as the recommendations to prevent depressive recurrences (endorsed by 78.8%)

**Members of the ISBD/IGSLI Task Force on the Treatment with Lithium who completed the Delphi survey and approved the paper**

M. Alda, Dalhousie University, Halifax, Canada

M. Bauer, University Hospital Carl Gustav Carus, TU Dresden, Dresden, Germany

A. Berghöfer, Charité – Universitätsmedizin, Berlin, Germany

M. Berk, Deakin University, Geelong, Australia

A. Bocchetta, University of Cagliari, Cagliari, Italy

C.V. Calkin, Dalhousie University, Halifax, NS, Canada

A. Duffy, Queens's University, Kingston, Canada

K.N. Fountoulakis, Aristotle University of Thessaloniki, Thessaloniki, Greece

A. Gonzalez-Pinto, CIBERSAM-BioAraba Research Institute, Vitoria, Spain

T. Hajek, Dalhousie University, Halifax, NS, Canada

L.V. Kessing, Psychiatric Center Copenhagen, Copenhagen, Denmark

C. López-Jaramillo, Universidad de Antioquia, Medellin, Colombia

R. Machado-Vieira University of Texas Health Science Center, Houston, TX, USA

R.M. Post, Bipolar Collaborative Network, Bethesda, MD, USA

J.K. Rybakowski, Poznan University of Medical Sciences, Poznan, Poland

C. Simhandl, Sigmund Freud Privat Universität, Wien, Austria

J.C. Soares, University of Texas Health Science Center, Houston, TX, USA

L. Tondo, Lucio Bini Mood Centers, Cagliari and Rome, Italy

## **Acknowledgements**

The following authors reported acknowledgements:

Dr. Baldessarini reports a grant from the Bruce J. Anderson Foundation and by the McLean Private Donors Psychopharmacology Research Fund.

Dr. Young mentions that this report represents independent research and that he is funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

## **Conflicts of Interest**

The following authors reported a potential conflict of interest:

Dr. Licht has received research grant from Glaxo Smith Kline, honoraria for lecturing from Pfizer, Glaxo Smith Kline, Eli Lilly, Astra-Zeneca, Bristol-Myers Squibb, Janssen Cilag, Lundbeck, Otsuka, Servier and honoraria from advisory board activity from Glaxo Smith Kline, Eli Lilly, Astra-Zeneca, Bristol-Myers Squibb, Janssen Cilag, and Sunovion.

Dr. Nielsen has received research grants from H. Lundbeck and Otsuka Pharmaceuticals for clinical trials, received speaking fees from Bristol-Myers Squibb, Astra Zeneca, Janssen & Cilag, Lundbeck, Servier, Otsuka Pharmaceuticals, and Eli Lilly and has acted as advisor to Astra Zeneca, Eli Lilly, Lundbeck, Otsuka Pharmaceuticals, Takeda, and Medivir.

Dr. Severus received speaker honoraria from Servier, Roche and Aristo in the past 3 years.

Dr. Vieta has received grants and served as consultant, advisor or CME speaker for the following entities: AB-Biotics, Abbott, Allergan, Angelini, AstraZeneca, Bristol-Myers Squibb, Dainippon Sumitomo Pharma, Farindustria, Ferrer, Forest Research Institute, Gedeon Richter, Glaxo-Smith-Kline, Janssen, Lundbeck, Otsuka, Pfizer, Roche, SAGE, Sanofi-Aventis, Servier, Shire, Sunovion, Takeda, the Brain and Behaviour Foundation, the Spanish Ministry of Science and Innovation (CIBERSAM), EU Horizon 2020, and the Stanley Medical Research Institute.

Dr. Zarate is a full-time U.S government employee and is an inventor on several patent and patent applications related to ketamine and ketamine metabolites that have all been assigned to the U.S. Government. He will share a percentage of any royalties that may be received by the Government in accordance to NIH policy.

This article is protected by copyright. All rights reserved.



**Table 1:** Overview of original inclusion criteria (see method) and modified inclusion criteria (see results)

#	Original inclusion criteria	Modified inclusion criteria
1	<p>Patients with BD, either type 1 or type 2 or both or unspecified.</p> <p>In instances where the trial also included other patient groups, the results of BD patients had to have been reported separately, or at least 80% of the total group should comprise BD patients</p>	Also allowing studies with patients with recurrent mood disorder, with at least some specified outcome data on patients with BD
2	Patients to be currently in remission	Also allowing studies with patients not in remission
3	Maintenance treatment with lithium aimed at preventing a new (treatment emergent) episode	Also allowing studies with lithium being used in the long-term treatment, i.e. in studies with a follow-up of at least 3 months
4	Evaluated the efficacy of lithium as monotherapy	Also allowing studies with lithium given in combination with other psychotropic drugs
5	Had a randomized controlled design	Also allowing non-randomized controlled studies under the condition of no clear indication of channeling bias
6	Patients were assigned to groups with different treatment regimens with lithium resulting in different a priori defined fixed lithium serum level ranges	Also allowing studies comparing different fixed lithium serum level ranges, but not a priori defined

**Table 2.** Original inclusion criteria met by the included studies

BP = Bipolar disorder; UP = Unipolar depressive disorder; APs = Antipsychotics; ADs = Antidepressants; Green = Criterion met; Orange = Unclear; Red = Criterion not met

Inclusion criterion	1	2	3	4	5	6			
Study Year	Included N	N BP randomized	N BP completed	Remission at randomization	Recurrence prevention	Lithium monotherapy	Randomization	A priori defined lithium levels	N inclusion criteria met
Jerram & McDonald 1978	BP and UP "over 80"	?	51	Yes	Yes	No info	Yes	Yes	4
Waters et al. 1982	BP 36	36	29	"residual mood swings"	No	No info	Yes	Yes	3
Coppen et al. 1983	BP and UP 88	?	23	No info	No ?	Total group: APs and ADs (%?)	Yes	Yes	2
Maj et al. 1986	BP 80	80	69	No info	No ?	No info	No	Yes	2
Gelenberg et al. 1989	BP-1 157	94	56	Yes	Yes	No info	Yes	Yes	5
Vestergaard et al. 1998	BP and UP 91	57	???	Still hospitalized	No	Total group: APs 35%, ADs 30%, Both 14%	Yes	Yes	2
Nolen & Weisler 2012	BP-1 742	742	295	Yes	Yes	Only zolpidem, benzo's, and chloralhydrate	No	Indirectly Yes ?	4

**Table 3.** Treatment characteristics prior to the study, at randomization/start of the study and during follow-up

APs = Antipsychotics; ADs = Antidepressants; NA = Not applicable

Study Year	Prior use of lithium	Medication at randomization		Lithium during follow-up				Other psychotropics during follow-up
		Lithium	Other medication	Abrupt change of dose	Immediate or slow Release	N doses/day	Sampling time	
Jerram & McDonald 1978	All	Yes	? (no info)	?	?	?	12-16 hrs	?
Waters et al. 1982	All	Yes	? (no info)	?	?	?	?	?
Coppen et al. 1983	All	Yes	Total group: APs and ADs (%?)	?	Slow	1dd	±12 hrs	Total group: APs and ADs (%?)
Maj et al. 1986	Probably not (no info)	? (no info)	? (no info)	?	Immediate ? (conventional formulation)	?	±12 hrs	?
Gelenberg et al. 1989	All	Yes	? (no info)	Abrupt	?	?	±12 hrs	?
Vestergaard et al. 1998	? (no info)	? (no info)	? (no info)	NA ? (no info)	Slow	2 dd	±12 hrs	Total group: APs 35%, ADs 30%, Both 14%
Nolen & Weisler 2012	Probably Yes (N unknown)	None	Quetiapine	NA	?	2dd	"Trough level"	Only zolpidem, benzo's, and chloralhydrate

**Table 4.** Summary of major clinical outcomes.

\* = Reconstructed with drop-out data; \*\* = One episode not stated; HR = Hazard ratio; PP = Prior vs. Post randomization;

CPRS = Comprehensive Psychiatric Rating Scale; SEs = Side effects; -> disc. = Led to discontinuation;

Significances: Green = Better outcome vs. Red = Poorer outcome; Blue = Not significant;

Study Year	Study arms (mmol/l)	Any recurrence	Manic recurrence	Depressive recurrence	Other outcome criteria			Tolerability/ side effects
Jerram & McDonald 1987	< 0.50 0.50-0.70 ≥ 0.70	4/16 (25%) 3/19 (15.8%) 2/16 (18.7%)	? ? ?	? ? ?				? ? ?
Waters et al. 1982	0.30-0.80 0.80-1.40	10/29 (34.5%) 02/19 (6.8%)	6/29 (20.7%) 1/29 (3.4%)	4/29 (13.8%) 1/29 (3.4%)				? ?
Coppen et al. 1983	< 0,80 ≥ 0,80	? ?	? ?	? ?	Affective morbidity index (PP) 0.12 > 0.09 (ns) 0.19 > 0.13 (ns)			? ?
Maj et al. 1986	0.30-0.45 0.46-0.60 0.61-0.75 0.76-0.90	11+2 DO/20 (65%) * 7+2 DO/20 (45%) * 6+1 DO/20 (35%) * 5+1 DO/20 (30%) *	? ? ? ?	? ? ? ?	CPRS mania score (PP) 1.47 0.69 0.64 0.37	CPRS depr. score (PP) 2.49 1.48 1.05 1.08		Side effects checklist score 0.79 1.43 2.03 2.22
Gelenberg et al. 1989	0.40-0.60 0.80-1.00	25/47 (53.2%) 15/47 (31.9%)	? ?	? ?	Time to recurrence HR low vs. high: 2.6			Less SEs and less withdrawals due to SEs in low level group vs. high level group
Vestergaard et al. 1998	0.50-0.80 0.80-1.00	In BP+UP group 8/41 (19.5%) 10/50 (20%)	In BP+UP group 2/41 (4.9%) ** 7/50 (14%)	In BP+UP group 5/41 (12.2%) ** 3/50 (6.0%)				? ?
Nolen & Weisler 2012	Low: < 0.60 High: 0.60-1.20 Placebo	? ? ?	? ? ?	? ? ?	Any mood event HR low vs. Pbo: 0.69 HR high vs. Pbo: 0.26 HR low vs. high: 2.23	Manic event HR low vs. Pbo: 0.88 HR high vs. Pbo: 0.42 HR low vs. high: 1.95	Depressive event HR low vs. Pbo: 0.78 HR high vs. Pbo: 0.32 HR low vs. high: 2.16	% with SEs -> % disc. Low: 62.0% -> disc.: 3.9% High: 59.2% -> disc.: 3.8% Placebo: 56.4% -> disc.: 3.0%



