Lithium toxicity profile: a systematic review and meta-analysis

Rebecca F McKnight, Marc Adida, Katie Budge, Sarah Stockton, Guy M Goodwin, John R Geddes

Summary

Background Lithium is a widely used and effective treatment for mood disorders. There has been concern about its safety but no adequate synthesis of the evidence for adverse effects. We aimed to undertake a clinically informative, systematic toxicity profile of lithium.

Methods We undertook a systematic review and meta-analysis of randomised controlled trials and observational studies. We searched electronic databases, specialist journals, reference lists, textbooks, and conference abstracts. We used a hierarchy of evidence which considered randomised controlled trials, cohort studies, case-control studies, and case reports that included patients with mood disorders given lithium. Outcome measures were renal, thyroid, and parathyroid function; weight change; skin disorders; hair disorders; and teratogenicity.

Findings We screened 5988 abstracts for eligibility and included 385 studies in the analysis. On average, glomerular filtration rate was reduced by −6·22 mL/min (95% CI −14·65 to 2·20, p=0·148) and urinary concentrating ability by 15% of normal maximum (weighted mean difference −158·43 mOsm/kg, 95% CI −229·78 to −87·07, p<0·0001). Lithium might increase risk of renal failure, but the absolute risk was small (18 of 3369 [0·5%] patients received renal replacement therapy). The prevalence of clinical hypothyroidism was increased in patients taking lithium compared with those given placebo (odds ratio [OR] 5·78, 95% CI 2·00–16·67; p=0·001), and thyroid stimulating hormone was increased on average by 4·00 IU/mL (95% CI 3·90–4·10, p<0·0001). Lithium treatment was associated with increased blood calcium (+0·09 mmol/L, 95% CI 0·02–0·17, p=0·009), and parathyroid hormone (+7·32 pg/mL, 3·42–11·23, p<0·0001). Patients receiving lithium gained more weight than did those receiving placebo (OR 1·89, 1·27–2·82, p=0·002), but not those receiving olanzapine (0·32, 0·21–0·49, p<0·0001). We recorded no significant increased risk of congenital malformations, alopecia, or skin disorders.

Interpretation Lithium is associated with increased risk of reduced urinary concentrating ability, hypothyroidism, hyperparathyroidism, and weight gain. There is little evidence for a clinically significant reduction in renal function in most patients, and the risk of end-stage renal failure is low. The risk of congenital malformations is uncertain; the balance of risks should be considered before lithium is withdrawn during pregnancy. Because of the consistent finding of a high prevalence of hyperparathyroidism, calcium concentrations should be checked before and during treatment.

Funding National Institute for Health Research Programme Grant for Applied Research.

Introduction Lithium is the most effective long-term therapy for bipolar disorder, protecting against both depression and mania and reducing the risk of suicide and short-term mortality.1–3 Although efficacious, lithium has some clinical disadvantages: it has a narrow therapeutic index requiring routine monitoring of serum concentrations and endocrine and renal function; a slow onset of action in acute mania; and acute effects of thirst, unpleasant taste, and tremor. Because lithium has always been an unpatented, cheap drug, it is not commercially promoted and the potential for adverse effects has been a substantial deterrent to use. Alternative drugs for bipolar disorder have increasingly been proposed, licensed, and adopted into clinical practice even when evidence for efficacy is modest and often limited to one pole of bipolar illness.

Particular concerns have been the effect of lithium on renal function and the risk from teratogenicity. Lithium commonly induces a clinically evident nephrogenic diabetes insipidus,4 which would be explained by actions on tubular renal function, but of more concern is the speculative description of a specific lithium nephropathy5 or disease of the renal glomerulus. However, the extent of reduction of glomerular renal function in a typical patient and the true long-term increase in risk of renal failure in well monitored patients remain poorly quantified. The risk of congenital malformations is generally thought to be high. One study reported a 400-fold increased risk of Ebstein’s anomaly,6 and present clinical practice recommendations have been to avoid lithium in pregnancy when possible. An up-to-date estimate of the true risks of lithium together with a systematic assessment of the associated renal problems has not been available.

Evidence has confirmed the important therapeutic benefits of lithium relative to some of the alternative drugs that have replaced it, which might lead to wider use of lithium.7 Clinicians and patients therefore need accurate evidence of harms and benefits. We report a systematic review and meta-analysis of studies investigating the association between lithium and all
reported major adverse effects, to provide a clinically informative systematic toxicity profile for lithium.

Methods
Search strategy and selection criteria
We searched Medline (1966–2010), Medline In-Process and other non-indexed citations (from 1966 to October, 2010), Embase (1980–2010), the Cumulative Index to Nursing and Allied Health Literature (1982–2010), PsycINFO (1806–2010), the Cochrane Library database (inception–2010), Biosis Previews (1926–2010), TOXNET database (inception–2010, weappendix), and archives of the journals Lithium, Lithium Therapy Monographs, and Teratology (search terms listed in weappendix). All relevant references were checked for additional and unpublished citations. Major textbooks of mood disorders and conference abstracts were hand-searched. We contacted pharmaceutical companies that market lithium, relevant clinicians, and authors of trials with incompletely reported data. All studies were assessed for meeting inclusion criteria, and those used for analysis were reviewed by a second researcher.

Studies were included in the review if they investigated one or more of the adverse events of interest. Randomised controlled trials (RCTs) comparing lithium with placebo, no treatment, or other drug therapies in patients with depression or bipolar disorder were considered most reliable if they included safety data for adverse effects, followed by prospective cohort studies comparing patients given lithium with those not given lithium, and then case-control studies. In the absence of controlled studies, we included uncontrolled prospective studies following up patients with depression or bipolar disorder given lithium and, finally, individual case reports. For each outcome, all studies meeting inclusion criteria were assessed and tabulated, but only the highest available form of evidence was included in the formal analysis. When only poor quality data from a higher level of evidence were available, we routinely included the next level down. For adverse events that often occur after months or even years of treatment, observational studies are often more informative than are RCTs.

Outcomes
The main outcomes investigated were: renal function (glomerular filtration rate [GFR, normal >90 mL/min], renal concentrating ability [maximum urinary concentrating ability, normal 800–1200 mOsm/kg]); thyroid function (thyroid stimulating hormone [TSH, normal 0.5–5.7 IU/mL], subclinical hypothyroidism [raised TSH with normal thyroxine] or clinical hypothyroidism [raised TSH and low thyroxine], or hyperthyroidism [depressed TSH and high thyroxine]); parathyroid function (total calcium [normal 2.1–2.8 mmol/L] and parathyroid hormone [PTH, normal 10–70 pg/mL]); bodyweight (clinically significant change in bodyweight ≥7% total weight in kg]); hair disorders; skin disorders; and teratogenicity (risk of major congenital and cardiac malformations in infants exposed to lithium in utero).

We judged study quality by assessing design aspects likely to induce bias—ie, method of randomisation and concealment of treatment allocation, blinding, length of follow-up, reporting withdrawals and dropouts, and method of analysis for RCTs; and likelihood of measurement bias, handling of confounding, and loss to follow-up for observational studies. Authors were contacted when published reports did not contain adequate details.

Statistical analysis
When appropriate, data from individual trials were pooled by meta-analysis with STATA (version 11.1). Both Mantel-Haenszel fixed and DerSimonian and Laird random effects models were used to assess the degree to which results were robust to the choice of statistical model. Non-standard units were converted to standard international units. Continuous data were combined to produce weighted mean differences (WMDs, for common measures) and standardised WMDs (for heterogeneous measures). Dichotomous and categorical data were combined to produce odds ratios (ORs) and absolute risk differences. Heterogeneity between study-specific estimates was investigated and, when important heterogeneity was expected or identified, sources for such variation were sought with meta-regression. We undertook sensitivity analyses to investigate the effect of exclusion of studies of inferior quality or with highly discrepant results.

Role of the funding source
The sponsor of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. RFMcK and JRG had full access to all the data, and JRG had final responsibility for the decision to submit for publication.

Results
The search process identified 5988 records, 385 of which fitted the inclusion criteria and were included for analysis (figure 1). The quality of evidence available varied between outcomes. High quality evidence was sparse: we identified only one systematic review (which was excluded from analysis because the data was used from the original studies included within it) and 22 RCTs. Most studies were case-control, uncontrolled cohort, or cross-sectional studies (n=197) or case reports (166). When cohort studies were reported in several publications, we analysed only the most complete set of data to avoid double counting cases. Studies published in English, French, and German were included; no studies in other languages met inclusion criteria.

30 studies investigated the effect of lithium GFR or maximum urinary concentrating ability, or both: nine case-control studies and 21 uncontrolled-cohort studies (weappendix). The uncontrolled cohorts could not be used for quantitative analysis because of the absence of...
data for within-patient change. The data were not adequate to analyse the effect of age or concomitant drugs (including diuretics). Overall, however, the results showed a small (0–5 mL/min) reduction in GFR over a mean observation time of 1 year (webappendix). Meta-analysis of case-control studies (cases=372, controls=307) showed that the GFR of patients taking lithium was lower than that of matched controls (figure 2). Maximum urinary concentrating ability was reduced by about 15% in patients taking lithium compared with controls (figure 3).

Data for the most clinically important outcome, renal failure, were scarce. The only substantial cohort study of patients on a lithium register reported 18 of 3369 (0·5%) as being treated with renal replacement therapy, compared with 0·2% of the Swedish general population.10,11

We identified 77 studies that reported the effects of lithium on thyroid function (webappendix): four RCTs, 16 case-control studies, 15 cohort studies, 20 cross-sectional reports, and 22 case reports. Because the RCTs collected heterogeneous data and the cohorts were uncontrolled, we used the case-control studies for analysis. Many studies before 1980 reported measures that were incompatible with more recent studies; these studies are shown in the webappendix but could not be used for analysis.

Eight studies compared the prevalence of subclinical or clinical hypothyroidism in patients given lithium (n=1402) for a mean of 70·1 months (SD 2·6) with the prevalence in controls (n=1032). Meta-analysis showed more hypothyroidism in patients given lithium than in controls (figure 4). The relative risk increased when only cases of clinical hypothyroidism were included (OR 6·05, 95% CI 2·72–13·37, p<0·0001; heterogeneity χ²=7·09 [df=5], p=0·21).
15 uncontrolled cohort studies (n=1085) measured a change in TSH over a mean of 18.5 months (SD 1.4); meta-analysis was not possible because of insufficient data (webappendix). Meta-analysis of the case-control studies (cases=645, controls=377) showed an increase in TSH concentrations in patients given lithium compared with controls (WMD 4.00 iU/mL, 95% CI 3.90–4.10, p<0.0001; heterogeneity χ²=1868.59 [df 10], p<0.0001). Four case-control studies reported possible increased thyroid function (webappendix). Meta-analysis showed no evidence of a difference between those taking lithium (n=178) and controls (n=181; OR 1.46, 95% CI 0.23–9.35, p=0.69; heterogeneity χ²=1.34 [df 2], p=0.51). Data from the RCTs accorded with that from the observational studies: a meta-analysis of lithium versus placebo trials reported that 4% of patients given lithium developed hypothyroidism compared with none given placebo (webappendix).1

60 studies (no RCTs) reported the effect of lithium on parathyroid function, and results were consistent. We identified four cohort studies, 14 case-control studies, 36 case reports, and six cross-sectional studies (webappendix). Calcium and PTH were increased by 10% compared with normal values in patients given lithium (n=730) compared with controls (n=699; figures 5, 6).

Weight change was included in 14 RCTs comparing lithium with placebo or other drug treatment (webappendix). Clinically significant weight gain (>7%) was more frequent in patients receiving lithium than in those receiving placebo (OR 1.89, 95% CI 1.27–2.82, p=0.002; heterogeneity χ²=2.28 [df 4], p=0.69; webappendix). Weight gain was lower with lithium than with olanzapine (n=285; OR 0.32, 95% CI 0.21–0.49, p<0.0001; heterogeneity χ²=0.72 [df 1], p=0.39; webappendix).

24 publications reported an adverse effect of lithium on hair, 14 of which were case reports (webappendix). One RCT of lithium (n=91) versus placebo (n=94) for 12 months reported hair loss in seven of 91 (8%) patients in the lithium group compared with six of 94 (6%) in the placebo group,12 whereas another reported hair loss in one of 32 (3%) patients given lithium versus none of 28 given placebo.13

We identified little high quality evidence supporting the association between lithium and skin disorders. 77 publications met inclusion criteria, 68 of which were case reports (webappendix). Two RCTs reported skin disorders within one combined analysis (webappendix). Meta-analysis showed no significant difference in the prevalence of skin disorders between patients given lithium and those given placebo (OR 1.28, 95% CI 0.49–3.36, p=0.62; heterogeneity χ²=0.29 [df 1], p=0.69).11,13

We identified 62 studies of the teratogenic potential of lithium: seven cohort studies, seven case-control studies, and 48 case reports (webappendix). Six case-control studies (n=264) measured the association between
Ebstein’s anomaly and lithium. The odds of exposure to lithium in cases of Ebstein’s anomaly did not differ significantly from controls; however, estimates are unstable because of the low number of events (Peto OR 0·27, 95% CI 0·004–18·17, p=0·54; heterogeneity $\chi^2=0·00$ [df 1], p=0·96; Mantel-Haenszel OR 2·0, 95% CI 0·20–20·6, p=0·54; heterogeneity $\chi^2=1·98$ [df 1], p=0·16).

A case-control study of 10 698 infants born with any major congenital abnormality and 21 546 healthy controls showed no significant association between lithium and congenital abnormalities (Peto OR 2·62, 95% CI 0·74–9·20, p=0·132; webappendix). The number of infants exposed to lithium was low in cases (six of 10 698) and controls (five of 21 546).

**Discussion**

The objective of this review was to synthesise what is known about the harmful effects of lithium. Findings from our study have shown that lithium is associated with increased risk of reduced urinary concentrating ability, hypothyroidism, hyperparathyroidism, and weight gain. We recorded no significant increased risk of congenital malformations, alopecia, or skin disorders, and little evidence for a clinically significant reduction in renal function in most patients.

The main limitations of this study are the quality and quantity of the primary evidence. High-quality data from long-term randomised or controlled cohort studies were sparse, and the sample size of most included observational studies was quite small. Although included studies reported doses and concentrations of lithium that are consistent with modern use, and data mainly represent the effects of lithium within the generally accepted therapeutic range rather than at concentrations of toxicity, dose information was incompletely reported and any potential effect of dose could not be specifically addressed in the meta-analysis. This review cannot, therefore, establish the relative safety of low doses or
dosing on alternative days. Furthermore, most studies excluded patients with a history of lithium toxicity or did not provide appropriate information to separate out these individuals or link their clinical presentation to number of episodes of toxicity or dosing regimens.

The studies were published over 60 years from 1950, and were highly variable in design (webappendix) and execution (data not shown). Diagnostic criteria, standard treatments, methods, and accuracy of measurement of physiological parameters have changed during that period. Moreover, because most cohort studies and RCTs did not use a patient group that was new to lithium or did not provide this information, length of follow-up was usually poorly defined so the average interval between first starting lithium and the onset of adverse events is unknown or approximate.

Many of the important cohort studies had a high dropout rate with little explanation of the cause of withdrawal. Although we made every effort to include studies reporting the same parameter investigated with a similar methodology, differences could be attributable to unidentified confounders.

We could not identify or obtain any unpublished data; therefore, there is a risk of publication bias. Nonetheless, we were able to locate a reasonable amount of evidence that allows cautious conclusions to be drawn about the safety of lithium. The panel shows our recommendations for clinical practice.

Although GFR is impaired by lithium treatment, impairment is not clinically significant in most patients. A maximum reduction in GFR of 5 mL/min represents only 5% of the minimum normal GFR. The pathological mechanism underlying the effects of lithium on glomerular function is not understood.

Progressive reductions in glomerular function can lead to end-stage renal failure, and lithium is thought to play a direct part in this process. In the 1970s, chronic tubulointerstitial nephropathy was described in patients with lithium-related end-stage renal failure, but this pathology is non-specific and not reliably linked to lithium.\(^6,7\) The risk of end-stage renal failure might be increased compared with healthy controls but the absolute risk seems to be low (0.5%). The incidence of chronic kidney disease is rising, especially in ageing populations, with an excess in women and an association with hypertension and diabetes. Chronic kidney disease can lead to end-stage renal failure in 2% of cases. Identification of the potential causal effect of lithium is difficult because of the confounding effects of diabetes and cardiovascular disease, which might lead to end-stage renal failure; but these disorders are also increased in patients with bipolar disorder compared with the general population.\(^8\) Large-scale epidemiological studies are needed that control for confounders (including age and sex) and model the effects of lithium dose, concomitant drugs (eg, angiotensin-converting-enzyme inhibitors, diuretics), treatment length, and repeated episodes of toxicity. Present clinical recommendations include recording of renal function before start of lithium therapy, and henceforth monitoring at intervals as short as 6 weeks. Because the absolute risk of end-stage renal failure is so low, yearly testing is probably sufficient in the absence of clinical reasons to monitor more frequently.

Renal, parathyroid, and thyroid function (at least GFR, TSH, calcium) should be repeated, at a minimum interval of every 12 months, more frequently if an abnormal result is found or the patient has a family history of endocrine disease.

Blood tests should all be repeated immediately if there is a change in mood state (eg, mania).

Occurrence of adverse effects (including skin and hair disorders) should be routinely recorded.

Women who would like to conceive or have become pregnant while receiving lithium should be advised that the increased risk of congenital malformations to women of childbearing age should be explained.

GFR=glomerular filtration rate. TSH=thyroid-stimulating hormone. *Changes to present therapy that we recommend; previous standard practice refers to UK guidelines.

Panel: Summary of recommended monitoring of lithium therapy in clinical practice

**Before start of lithium therapy**

- The risk of major adverse events (as summarised in this Article) should be discussed with the patient.
- A serum calcium should be added to baseline blood tests.
- Uncertainty about risk of congenital malformations to women of childbearing age should be explained.

**During lithium therapy**

- Renal, parathyroid, and thyroid function (at least GFR, TSH, calcium) should be repeated, at a minimum interval of every 12 months, more frequently if an abnormal result is found or the patient has a family history of endocrine disease.
- Blood tests should all be repeated immediately if there is a change in mood state (eg, mania).
- Occurrence of adverse effects (including skin and hair disorders) should be routinely recorded.
- Women who would like to conceive or have become pregnant while receiving lithium should be advised that the increased risk of congenital malformations is uncertain; patient and clinician should discuss the balance of risks between harm to the baby and maternal mood instability before making a decision to stop lithium therapy.

---

**References**


8. The Lancet 2012; 379: 1383-91


17. The Lancet 2012; 379: 1448-54


20. The Lancet 2012; 379: 1469-75

---

**Figure:**

A maximum reduction in GFR of 5 mL/min represents only 5% of the minimum normal GFR. The pathological mechanism underlying the effects of lithium on glomerular function is not understood.

**Table:**

| GFR=glomerular filtration rate | TSH=thyroid-stimulating hormone | *Changes to present therapy that we recommend; previous standard practice refers to UK guidelines. |
led to normalisation of increases in T4 and decreased TSH. However, mood symptoms can be harder to treat when patients are in the low normal range of thyroid function, so treatment might be warranted on psychiatric grounds. Lithium is concentrated by the thyroid gland and has four potentially negative effects on thyroid function: inhibition of iodine uptake, inhibition of iodotyrosine coupling, alteration in thyroglobulin structure, and inhibition of thyroxine secretion. TSH concentrations tend to be increased in response to the inhibitory effect on thyroxine availability.

Primary hyperparathyroidism was quite frequent in patients receiving lithium: an absolute risk of 10% (vs 0·1% of the general population) is probably attributable to lithium’s inactivation of the calcium-sensing receptor and interference with intracellular second messenger signalling. This effect leads to an increased release of parathyroid hormone, which raises calcium concentrations in blood.

Thyroid and parathyroid abnormalities occur in about 25% of patients receiving lithium therapy, and clinical monitoring should reflect this finding. Guidelines for bipolar disorder make no mention of monitoring of calcium, which seems to be an important omission in view of the high absolute risk of hyperparathyroidism. Baseline blood tests before lithium is given should include TSH and calcium, and should be monitored every year or more frequently if clinical symptoms are reported. We recorded no evidence for a toxic effect of hypercalcaemia on renal function in patients given lithium, but it could contribute in view of the known risk of a decrease in renal function in long-term hypercalcaemia. More research is needed to clarify the relation between lithium, calcium, and the kidney.

Several factors might explain the association between lithium and weight gain: its insulin-like properties in increasing cellular glucose uptake, increased thirst, direct stimulation of the hypothalamic appetite centre, and the induction of hypothyroidism. Lithium might also affect relevant neurotransmitter receptor function, although the effect is less than, for example, olanzapine, which inhibits histaminergic and serotonergic receptors in the brain.

The evidence that exposure to lithium is teratogenic is quite weak, and our findings accord with the notion that the risk has been overestimated. Thus, the risk estimates were not significant, although the upper confidence limit is consistent with a clinically significant result. Estimates were not significant, although the upper confidence limit is consistent with a clinically significant result.

In conclusion, clinical practice guidelines have long recommended lithium as a first-line long-term treatment for bipolar disorder but its use has decreased, partly because of safety concerns. Evidence confirming its efficacy has led to suggestions that lithium should again be more widely used. This review provides a comprehensive synthesis of the evidence of harm that should inform clinical decisions and draw attention to key questions in urgent need of further clarification.

Contributors
RMcK located references, extracted data, assisted with analyses and results interpretation, and drafted the report. KB assisted with locating references and data recording. MA initiated the review and helped to design the methodology, located references, and assessed quality. SS ran the electronic database searches. GMG initiated the review and helped to design the methodology. JRG initiated the review, helped to design the methodology, assessed the quality of studies, and did the analyses. All authors helped to interpret findings and write the final report. JRG is guarantor.

Conflicts of interest
We declare that we have no conflicts of interest.

Acknowledgments
This Article presents independent research commissioned by the National Institute for Health Research (NIHR) under Research Programme Grant for Applied Research: RP-PG-0108-10087. The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health. The review was designed, conducted, analysed, and interpreted by the authors entirely independently of the funding sources.

References