Figure S1: Differences in cortical thickness (top), area (middle) and volume (bottom) for patients using lithium (Li) versus patients who did not use lithium medication (left panel), and for patients using antiepileptic drugs (AE) versus No-AE-users (right panel). Significance after
Multiple comparison correction is represented on a log(p-value) scale, where positive values (warm colors) are assigned to users>100-users, and negative values (cold colors) to users<100-users.
Figure S2: Differences in cortical thickness (top), area (middle) and volume (bottom) for BDI patients versus controls corrected for lithium (Li) use (left panel) and use of antiepileptic drugs (AE, right panel). Significance after multiple comparison correction is represented on a log(p-value) scale, where positive values (warm colors) are assigned to BDI<controls, and negative values (cold colors) to BDI>controls clusters.
Figure S3: a): Differences in cortical thickness (top), area (middle) and volume (bottom) for BDII patients versus controls corrected for use of antiepileptic drugs (AE, left panel). In BDII patients versus controls the results are the same with and without correction for lithium use (see

DOI: 10.1503/jpn.150093

Copyright © 2015, 8872147 Canada Inc.

*Online appendices are unedited and posted as supplied by the authors.*

main article). b): Differences in cortical thickness (top) and volume (bottom) for BDI versus BDII patients correcting for lithium use (Li) use. No differences were observed for cortical area. c): Differences in cortical thickness (top) for BDI versus BDII patients correcting for AE use. No differences were observed for cortical area or volume. Significance after multiple comparison correction is represented on a log(p-value) scale, where positive values (warm colors) are assigned to BDII<controls in a) (BDI<BDII in b) and c)), and negative values (cold colors) to BDII>controls clusters in a) (BDI>BDII in b) and c)).
Figure S4: The only difference in pial surface area was observed in the left hemisphere, where patients with BDI revealed lower surface area than controls. No other differences in pial surface area were observed in any of the group comparisons. The color code is equivalent to Figures S1, S2 and S3.

DOI: 10.1503/jpn.150093

Copyright © 2015, 8872147 Canada Inc.

*Online appendices are unedited and posted as supplied by the authors.*

Figure S5: We compared cortical thickness between BDII patients and controls, BDII versus BDI patients, and BDI patients versus controls in females only. The color code is equivalent to
Figures S1, S2 and S3. This reanalysis perfectly matched the results obtained for both sexes. Importantly - and in agreement with results obtained in the main analysis - we observed a case-control difference in the right medial prefrontal regions only when comparing BDI patients with controls, but not when comparing BDII patients with controls. Also in line with the results reported for both sexes, cluster sizes were larger in the BDI patients versus controls comparison than in the BDII patients versus controls comparisons. Moreover, the original suggestive finding that patients with BDI showed lower cortical thickness than patients with BDII in right temporal and medial prefrontal areas is also present when analyzing females only.

With these extended analyses, we feel confident that the results were not confounded by sex differences, and it is unlikely that the larger proportion of females in the BDII group could account for the suggested lesser extent of cortical thickness abnormalities in patients with BDII.
Inclusion and exclusion criteria:

Patients were recruited from the St. Göran bipolar project, enrolling patients from the bipolar unit at the Northern Stockholm Psychiatric Clinic, Stockholm, Sweden. All patients were assessed by a psychiatrist or resident psychiatrist using a standardized interview protocol: the Affective disorders evaluation that has previously used in the Systematic Treatment Enhancement Program of Bipolar Disorder program (Sachs et al, 2003). The Affective disorders evaluation guides the interviewer through a systematic assessment of the patient’s current mental state, past history, and diagnosis according to DSM-IV criteria as contained in the Structured Clinical Interview for DSM-IV (SCID). Co-morbid psychiatric disorders were screened for using Mini International Neuropsychiatric Interview (M.I.N.I.) (Sheehan et al, 1998). The full diagnostic assessment was based on all available sources of information, including patient interview, case records and, if possible, interviews with the next of kin. The diagnoses were set at a diagnostic case conference, where all information at the time of admission was presented. A consensus panel of experienced board certified psychiatrists specialized in bipolar disorder made a best-estimate diagnostic decision. Using this procedure, the risk of inter-rater bias in the inclusion process was reduced. The general criteria for inclusion were patients at least 18 years
old and who met the DSM-IV criteria for any bipolar disorder, i.e., type I, II, NOS, cyclothymia, or schizoaffective syndrome manic type. Here, only bipolar 1 and 2 patients were included. Information was collected about number of depressive, manic, and mixed episodes, history of suicide attempts, family history (first or second degree relatives with bipolar disorder), history of abuse (alcohol or substances), comorbid anxiety disorders (i.e., panic disorder, social phobia, post-traumatic stress disorder, generalized anxiety disorder, obsessive-compulsive disorder, and agoraphobia), body mass index, and history of psychosis. The lifetime severity of bipolar disorder was rated using the Clinical Global Impression (CGI) rating scales. This seven point scale reflects the clinician’s rate of the severity: 1 normal or not at all ill, 2 borderline mentally ill, 3 mildly ill, 4 moderately ill, 5 markedly ill, 6 severely ill, and 7 extremely ill. In order to determine euthymia, the Montgomery-Asberg Depression Rating Scale (MADRS) and the young mania rating scale (YMRS) were used (euthymia defined as MADRS<14 and YMRS<14). For ethical reasons, patients continued to take their prescribed medications at the time of scanning.

Healthy control participants were recruited from the general population in the same catchment area as the bipolar subjects. Statistics Sweden (www.scb.se) randomly selected individuals matched for age and gender to the bipolar subjects, seven persons per subject, and sent them letters of invitation to participate in the study. Responders were screened for mental illness and other medical conditions by a research nurse through a telephone interview. The telephone interview screened for and excluded subjects with drug abuse, somatic illness, mental illness, pregnancy, first-degree relative with bipolar disorder or schizophrenia, metal objects in the body excluding MRI-scan. Eligible persons were scheduled for a visit and further investigated for
mental illness by a psychiatrist using the M.I.N.I. and selected parts of the Affective Disorder
Evaluation. Subjects also completed the SCID-2 self-rating form (Structured Clinical Interview
for DSM-IV Axis II Personality Disorders) and the ASRS (The World Health Organization
Adult ADHD Self-Report Scale). Controls presenting potentially pathological findings were
discussed between examining clinician, primary investigator, and study coordinator at case
conferences where a decision was made whether to exclude the subject or not. Further exclusion
criteria were: neurological conditions except mild migraines, untreated endocrine disorders,
dementia, recurrent depressive disorder, severe personality disorder. Controls with past history of
an isolated depressive episode, isolated episode of panic disorder, eating disorder, or obsessive
compulsive disorder that remitted spontaneously or with brief psychotherapeutic counselling
were included in the study.