ABSTRACT

Introduction  Bipolar disorder is an often disabling mental illness with a lifetime prevalence of 1%–2%, a high risk of recurrence of manic and depressive episodes, a lifelong elevated risk of suicide and a substantial heritability. The course of illness is frequently characterised by progressive shortening of interepisode intervals with each recurrence and increasing cognitive dysfunction in a subset of individuals with this condition. Clinically, diagnostic boundaries between bipolar disorder and other psychiatric disorders such as unipolar depression are unclear although pharmacological and psychological treatment strategies differ substantially. Patients with bipolar disorder are often misdiagnosed and the mean delay between onset and diagnosis is 5–10 years. Although the risk of relapse of depression and mania is high it is for most patients impossible to predict and consequently prevent upcoming episodes in an individual tailored way. The identification of objective biomarkers can both inform bipolar disorder diagnosis and provide biological targets for the development of new and personalised treatments. Accurate diagnosis of bipolar disorder in its early stages could help prevent the long-term detrimental effects of the illness. The present Bipolar Illness Onset study aims to identify (1) a composite blood-based biomarker, (2) a composite electronic smartphone-based biomarker and (3) a neurocognitive signature for bipolar disorder.

Methods and analysis  The study will include 300 patients with newly diagnosed/first-episode bipolar disorder, 200 of their healthy siblings or offspring and 100 healthy individuals without a family history of affective disorder. All participants will be followed longitudinally with repeated blood samples and other biological tissues, self-monitored and automatically generated smartphone data, neuropsychological tests and a subset of the cohort with neuromaging during a 5 to 10-year study period.

Ethics and dissemination  The study has been approved by the Local Ethical Committee (H-7-2014-007) and the data agency, Capital Region of Copenhagen (RHP-2015-023), and the findings will be widely disseminated at international conferences and meetings including conferences for the International Society for Bipolar Disorders and the World Federation of Societies for Biological Psychiatry and in scientific peer-reviewed papers.

Trial registration number  NCT02888262.

Strengths and limitations of this study

► The Bipolar Illness Onset (BIO) study is the first study aiming to identify (1) a composite blood-based biomarker, (2) a composite electronic smartphone-based biomarker and (3) a neurocognitive signature for bipolar disorder.
► The same biomarkers will be measured longitudinally in newly diagnosed/first-episode patients with bipolar disorder and their healthy first-generation relatives.
► The BIO study will be performed by an experienced international research group, combining expertise from all areas of the study.
► Extensive initial assessment of study procedures may result in selection of participants who are intrinsically positive towards clinical research.
► Confounding effects of psychotropic medication may influence results.
► Lack of randomised interventions may limit causal interpretations of the results.
disorders such as unipolar disorder are unclear although some pharmacological and psychological treatment strategies differ substantially. Patients with bipolar disorder are often misdiagnosed as having unipolar disorder, transient psychosis, reaction to stress/adjustment disorder or psychoactive substance abuse, and the mean delay between onset and diagnosis is 5–10 years. Although the risk of relapse of depression and mania is high, it is for most patients impossible to predict and consequently prevent upcoming episodes in an individual tailored way. The identification of objective biomarkers as measures of pathophysiological processes can both inform bipolar disorder diagnosis and provide biological targets for the development of new and personalised treatments.

Accurate diagnosis of bipolar disorder in its early stages could help prevent the long-term detrimental effects of the illness. Recently, promising results have been presented regarding a diagnostic test for unipolar depression comprising levels of nine biomarkers in peripheral blood. Although the nature of bipolar disorder seems more biologically driven than the nature of major depression with a higher heritability, there has been no or few attempts to identify a similar composite biomarker for bipolar disorder.

Onset of illness and staging in bipolar disorder

Although the course of illness is heterogeneous, there is a body of evidence for clinical progression on average of unipolar and bipolar disorders as increasing number of affective episodes seem associated with (1) increasing risk of recurrence, (2) increasing duration of episodes, (3) increasing symptomatic severity of episodes, (4) decreasing threshold for developing episodes and (5) increasing risk of developing dementia. It is likely that this clinical progression with deteriorating effects of affective episodes and duration of illness is associated with neurobiological changes over the course of illness. Unfortunately, results of all longitudinal studies on the biology of bipolar disorder are hitherto hampered by three major limitations: (1) Only few studies have recruited patients with bipolar illness from onset of the illness and most of these have used first-onset mania as inclusion criteria thereby excluding patients with a hypomanic episode (bipolar disorder, type II). As the biology of bipolar disorder—based on cross-sectional studies—seems to change over the course of illness from first episode to first relapse and recurrent relapses to an unremitting or rapid cycling course, this is a major limitation in our knowledge internationally, (2) the number of patients included in prior studies is less than 200, which is a limitation taking the heterogeneity of bipolar disorder into account (eg, bipolar disorder types I and II may have different biology, and so on), and (3) the prospective follow-up period is less than a few years in all studies.

In the Bipolar Illness Onset (BIO) study, we will establish a large cohort comprising three subcohorts that will be followed long term with systematic diagnostic, blood-based biomarkers, smartphone data and cognitive and brain imaging assessment. The three subcohorts will consist of: (1) patients with newly diagnosed/first-episode bipolar disorder and (2) their healthy first-generation siblings and offspring and (3) healthy individuals without a family history of affective disorder.

Overall aims

1. To identify a composite blood-based biomarker measure as well as a composite electronic smartphone-based biomarker from onset of bipolar disorder during progression and in later stages.
2. To investigate if the composite blood-based biomarker measure and electronic smartphone-based biomarker identified among patients with bipolar disorder predict onset of illness (depression or mania) among these patients’ healthy first-generation siblings and offspring.
3. By applying an integrated systems approach, to identify patterns of ‘cerebral signatures’ across neurocircuitry and cognitive levels and to validate the composite blood-based biomarker and electronic smartphone-based biomarkers against these biosignatures.
4. To investigate long-term developmental trajectories in neurocognitive function and brain imaging from the high-risk state to onset of bipolar disorder following first relapse and recurrent relapses and in the late stage with an unremitting, multiphase or rapid cycling course.
5. To investigate whether the course of illness is progressive on average in bipolar disorder and to identify corresponding changes in biomarkers during the course of illness within BIO-1 to BIO-4.

METHODS AND ANALYSIS

Overall methods

The BIO study is a long-term cohort study that started on April 2015 and planned to include 300 newly diagnosed/first-episode patients with bipolar disorder from the Copenhagen Affective Disorder Clinic, 200 of these patients’ healthy first-generation relatives and 100 healthy individuals without a first-generation family history of affective disorders.

The Copenhagen Affective Disorder Clinic

The Copenhagen Affective Disorder Clinic is a mood disorder clinic that provides treatment service for patients with newly diagnosed/first-episode bipolar disorder. The Copenhagen Affective Disorder Clinic receives patients from the entire Capital Region of Denmark covering a catchment area of 1.6 million people and all psychiatric centres in the region. All patients referred to the Clinic as newly diagnosed/first-episode patients, that is, onset of first manic or hypomanic episode or when the ICD-10 (International Statistical Classification of Diseases and Related Health Problems 10th Revision) diagnosis of bipolar disorder is made for the first time, will routinely...
be asked for inclusion in the BIO study. Nearly all patients treated in the Clinic have a diagnosis of bipolar disorder type I or type II whereas patients with bipolar disorder not otherwise specified or patients with cyclothymia are not treated in the Clinic and consequently not included in the BIO study.

Recruitment of the three cohorts
1. Three hundred patients (aged 15–70 years) referred to the Copenhagen Affective Disorder Clinic as newly diagnosed/first-episode patients with bipolar disorder, that is, onset of first manic or hypomanic episode or when the diagnosis of bipolar disorder is made for the first time. The Clinic receives more than 100 newly diagnosed/first-episode patients with bipolar disorder each year and we expect that nearly all will accept participation in the BIO study as this implies an extensive clinical evaluation.
2. Two hundred first-generation relatives (siblings and children aged 15–40 years) to the recruited newly diagnosed/first-episode patients with bipolar disorder.
3. One hundred age and gender-matched healthy individuals without a first-generation family history of affective disorders recruited among blood donors from the Blood Bank at Rigshospitalet, Copenhagen, as in prior studies.

Diagnostic assessments at inclusion
The initial diagnostic assessment will be done using the Structured Clinical Interview for DSM-IV-TR Axis I Disorders, categorising patients into bipolar disorder type I or type II as part of daily praxis by the experienced specialists in psychiatry during the patients’ 2-year stay in the Copenhagen Affective Disorder Clinic. This clinical diagnosis of bipolar disorder will be confirmed in a semistructured research-based interview using the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) providing an ICD-10 diagnosis. There will be no attempt to balance the prevalence of bipolar subtypes in the BIO study.

Follow-up
Besides the assessments at inclusion, patients will be assessed during remitted, depressive and manic/mixed phases. Patients and healthy control individuals will be initially assessed face-to-face and at least every year during the first 4 years and after this, every second year for 5 years.

As part of daily clinical praxis in Copenhagen Affective Disorder Clinic and as part of the BIO-2 substudy, all patients will get access to a smartphone app for electronic continuous monitoring of illness activity during a 5-year follow-up period (see substudy BIO-2). Additionally, research assistants will contact all participants every third month to identify upcoming episodes/onset of illness and to ensure continued participation in the BIO study. At each assessment, the present clinical state (remission, manic, hypomanic, depressive, mixed episode) of all participants will be established according to ICD-10. The severity of depressive and manic symptoms (if present) will be assessed using the 17-item Hamilton Depression Rating Scale (HAMD-17) and the Young Mania Rating Scale (YMRS) with a time period of 3 days applied. Remission is defined as score of $\leq 7$ on the HAMD-17 and the YMRS. In this way, upcoming episodes and onset of illness (for healthy individuals) will be assessed with great certainty.

Based on prior findings, it is estimated that patients will develop four to five affective episodes on average during the follow-up period, that is, relapse or recurrence, defined as above. It estimated that 20%–30% of the healthy first-generation relatives will develop onset of affective illness compared with 2%–5% among the healthy individuals without a first-generation family history of affective disorders. Finally, linkage to Danish nationwide register-based data will be included for all individuals on psychiatric hospitalisations, prescribed medication and socioeconomic variables during the 5-year follow-up.

Investigations
We will include state-of-the-art methods within clinical assessment as well as biomarkers including a range of new methods:
1. clinical assessments using the Hamilton Depression Scale-17 items (HAMD-17), the YMRS, the Functional Assessment Short Test (FAST) (a measure of psychosocial function);
2. questionnaires including the Hypomania Check list, Standardized Assessment of Personality-Abbreviated Scale, the WHO (Five) well-being index and the Verona Satisfaction Scale-Affective Disorder;
3. standardised fasting blood tests for a large number of potential biomarkers (see the BIO-1, Biological tests section);
4. spot urine samples for oxidative generated damage to DNA and RNA;
5. hair cortisol as a valid and reliable index of long-term systemic cortisol levels;
6. combined heart rate and movement sensor mounted at the thorax (Actiheart);
7. daily electronic smartphone-based self-monitoring of depressive and manic symptoms;
8. daily electronic smartphone-based automatically generated data (eg, data on phone usage, social activity and physical activity);
9. neuropsychological assessment (only during full or partial remission);
10. structural MRI focusing on prefrontal cortex and hippocampus as well as functional MRI (fMRI) (only during full or partial remission).

Clinical assessments will be performed by six PhD students (master’s degree as a Medical Doctor or psychologist). They will be certified in a PhD training course in the SCAN interview and will be trained in using rating scales (HAMD-YMRS, FAST). Inter-rater sessions supervised by senior clinicians from the Copenhagen Affective Disorder Clinic and senior researchers in the BIO study.
will be conducted among the researchers at regular time points during the entire study period.

The BIO study includes four separate but interacting substudies (BIO-1, BIO-2, BIO-3, BIO-4), as illustrated in table 1 with specified aims, background and theoretical basis, and methods, as described in the following. Background and reasons for selection of putative biomarkers is an iterative process presented in each substudy. Nevertheless, as the area of potential biomarkers is constantly evolving and as the substudies cover four different areas, the research protocol does not include systematic reviews of the literature.

**BIO-1: peripheral blood-based biomarker in bipolar disorder**

Aims
To identify a composite blood-based biomarker that (1) discriminates individuals with bipolar disorder from healthy control individuals, (2) discriminates between manic, depressive and remitted states, (3) predicts emerging affective episodes, and (4) to validate the composite blood-based biomarker against the smartphone-based biomarker, and the neurocognitive and brain imaging signature, and (5) to investigate the change in individual biomarkers as well as the composite blood-based biomarker following onset of first manic episode, during successive relapses and in the end late stages of the illness.15

Background and theoretical basis
In a series of meta-analyses, we concluded that although a number of candidate peripheral biomarkers related to neuroplasticity, inflammation, oxidative stress and gene expression seem promising, findings are limited to neuroplasticity, inflammation, oxidative stress and a number of candidate peripheral biomarkers related to illness stage, it is necessary to include assessment before onset of illness and during first and recurrent relapses and in the late stages of the illness according to the staging model of bipolar disorder.49

We have in two longitudinal studies with repeated assessments of patients with bipolar disorder during affective states (manic/mixed, depressive and euthymic) and healthy control individuals found that brain-derived neurotrophic factor (BDNF),39 increased oxidative DNA and RNA damage40 41 and decreased mRNA expression of the PTGDS gene encoding the prostaglandin D synthase enzyme42 are markers related to the illness trait in bipolar disorder. The level of the cytokines IL-6 and IL-18 (interleukin) was related to manic episodes only and the activity of GSK-3beta varied with affective states,43 suggesting that these may be state markers.44 These results have contributed to the research area of biomarkers in bipolar disorder, moving the area closer towards identifying clinically applicable biomarkers.

Nevertheless, it is unlikely that one single biomarker will provide a useful diagnostic tool; instead, a composite of several relevant biomarkers appears as more viable approach.45 Recently, promising results have been presented regarding a diagnostic test in unipolar depression comprising serum levels of nine individual biomarkers in peripheral blood.15 Similarly, preliminary studies have suggested composite biomarkers for bipolar disorder.46-48

We identified a composite biomarker consisting of gene expression from 19 candidate genes for bipolar disorder that accurately discriminated patients with bipolar disorder from healthy control individuals.46 Thus, such approaches highly increase the chances of obtaining a high specificity and sensitivity of the composite blood-based biomarker.46-48

In order to establish relevant markers of risk and markers related to illness stage, it is necessary to include assessment before onset of illness and during first and recurrent relapses and in the late stages of the illness according to the staging model of bipolar disorder.49 Further, the study of early-onset individuals is necessary to evaluate biomarker levels without influence from medication, which may otherwise limit the validity of identified biomarkers.

### Methods

BIO-1 will include repeated clinical assessments and corresponding samples of blood and other tissues among...
the 300 newly diagnosed/first-episode patients with bipolar disorder, the 200 first-generation relatives and the 100 healthy individuals without a first-generation family history of affective disorder.

We will estimate a composite blood-based biomarker based on a number of individual markers including BDNF, neutrophin-3, five different cytokines, gene expression, additional 30 candidate genes and other potential biomarkers (see the Biological tests section), and identify the composite biomarker that correlates best with affective states and Hamilton Depression Rating Scale and YMRS scores of depression and mania, respectively. A final blood-based biomarker will be chosen based on its ability to (1) discriminate patients with bipolar disorder from healthy control individuals and to (2) discriminate between manic, depressive and remitted states in bipolar disorder.

**Laboratory procedures**

We will obtain careful standardisation of blood sampling and laboratory analysis by obtaining blood samples in a fasting state and in a 1-hour interval in the morning. At the same day as blood sampling, smoking status, alcohol use, body mass index, menstrual cycle, and so on will be assessed. Blood sampling and all phases of laboratory processing for plasma and DNA/RNA analyses will be done at the Department of Clinical Biochemistry, Rigshospitalet, using standard operational methods conducted by a team of technicians blinded with respect to participant status. All plasma samples will be stored at −80°C. The BIO study will include a total of 2400 blood samples: 300 patients × 5 + 200 healthy first-generation relatives × 3 + 100 healthy participants without a first-generation family history × 3.

**Biological tests**

We will use a multianalyte panel including a large number of potential biomarkers such as plasma levels of Neutrophins3, GSK-3, B-amyloid Aβ40 and Aβ42, BDNF, inflammatory markers, high-sensitive C-reactive protein, lipoproteins (very-low-density lipoprotein, low-density lipoprotein, high-density lipoprotein) and specific apolipoproteins (eg, apoE, apoA-I, apoA-II and apoM) as potential markers of low-grade inflammation particularly salient in the early stages of bipolar disorder. Total RNA, microRNA, genomic DNA and histones are isolated from peripheral blood mononuclear cells. Gene expression and alternative slicing of RNA transcripts are analysed using array real-time reverse transcriptase quantitative PCR and next-generation sequencing. Epigenetic modifications of the DNA (eg, methylation) are measured using antibody-based methods or bisulphite treatments in combination with next-generation sequencing. The genomic positions of histones with specific modifications are detected using chromatin immunoprecipitation sequencing. The degree of histone modifications is measured by semiquantitative antibody-based detection.

Measurements of DNA and RNA damage by oxidation are obtained from spot urine samples and analysed using ultra performance liquid chromatography and mass spectrometry.

Hair cortisol will be included as a valid and reliable index of long-term systemic cortisol levels.

We will report on these individual biological tests including comparing patients, first-degree relatives and healthy control individuals, when appropriate. The BIO study sample of 600 participants is rather small for genetic analyses discriminating patients with bipolar disorder from healthy controls but some genetic analyses, including the CACNA1C gene, can be conducted in cooperation with national and international genetic network groups.

**Statistical analyses**

Data represent repeated measures within and between individuals and will be analysed using a combination of generalised linear mixed models, integrated data analysis and penalised regression approaches to facilitate the combined feature selection and prediction of the available high-dimensional data. Integrative data analysis ensures that we are able to identify an improved composite blood-based biomarker since data from different molecular levels are combined in a simultaneous analysis that closely resembles the biological system. Further, we will use cross validation or alternatively split sample designs in the development and validation of the composite biomarker. Finally, if possible we will evaluate the biomarker(s) in external non-Danish data sets in collaboration with other international researchers.

A general principle that pertains to statistical analyses of all four substudies is the ‘intention to treat’ principle. Accordingly, participants who during follow-up get a diagnosis with a higher diagnostic validity than bipolar disorder (ie, a lower ICD-10 diagnostic number, DF00, DF10 and DF20) that may substantially influence the biomarker measures are included in the analyses until onset of symptoms from the disorder but excluded from subsequent analyses. These disorders include significant neurologic disorders such as dementia, stroke, brain tumour, multiple sclerosis, Parkinson’s disease as well as disorders due to significant psychoactive substance use and schizophrenia.

Furthermore, in all four substudies, the problem of missing data will be alleviated by the use of mixed effect model (for longitudinal measurements) and multiple imputations using chained equations when applicable. If possible, joint modelling will be considered, depending on the missing mechanism observed.

**Statistical power**

The study has a power of 80% to detect a minimum increment of 6.5 percentage points in sensitivity if we assume that the existing diagnostic tools have a sensitivity of 70% to diagnose bipolar disorder for a patient who has the disorder (see ref 52). Thus, if the composite
biomarker score increases the sensitivity by a minimum of 6.5 percentage points then we have a power of 80% to detect that increase based on 300 patients with bipolar disorder using a one-sided exact binomial test (for fixed specificity).

**BIO-2: smartphone-based electronic biomarker in bipolar disorder**

**Aims**
To identify a composite smartphone-based electronic biomarker that (1) discriminates patients with bipolar disorder from healthy control individuals, (2) discriminates between manic, depressive and remitted states, (3) predicts emerging affective episodes, and (4) to investigate the change in the composite smartphone-based biomarker following onset of first manic episode, during successive relapses and in the end stages of the illness as according to the staging system by Berk et al.\(^\text{15}\).

**Background and theoretical basis**
Recently, electronic self-monitoring of the severity of depressive and manic symptoms using text messages has been suggested as an easy and inexpensive way to identify early signs of affective episodes, providing opportunities for mental healthcare providers to intervene shortly after prodromal symptoms first appear.\(^\text{53}\) We have in the MONARCA project developed and tested a smartphone-based electronic monitoring system including daily subjective self-assessments of illness activity in bipolar disorder as well as a bidirectional feedback loop between the patient and clinicians (the MONARCA system\(^\text{54–58}\)).

Using the MONARCA system, fine-grained electronic data were collected during everyday life in naturalistic settings in patients with bipolar disorder. The MONARCA system was reported highly useable and useful by patients with bipolar disorder with a high self-assessment adherence (87%–95%), and the patients reported that the MONARCA system helped them to better manage their disease.\(^\text{54, 55}\) Further, the severity of depressive and manic symptoms was found to correlate with automatically generated smartphone data including (1) physical activity as reflected by the number of changes in cell tower ID per day,\(^\text{59}\) (2) social activity as reflected by the number of incoming and outgoing calls per day, the duration of incoming and outgoing calls per day and the number of outgoing text messages per day,\(^\text{60}\) and (3) voice features collected during phone calls.\(^\text{61}\)

Although these findings are encouraging there is a need to integrate self-monitored smartphone data with automatically generated smartphone data on social and physical activities, speech and sleep into one composite smartphone-generated electronic biomarker measure. This composite measure should be modelled to (1) discriminate patients with bipolar disorder from healthy control individuals, (2) have a high correlation with depressive and manic symptoms, (3) discriminate between euthymic, manic and depressive states and (4) early predict emerging affective episodes for the individual patient to increase the possibility for early intervention.

To validate smartphone-based measures of physical activity and sleep, a subset of patients will wear a combined heart rate and movement sensor mounted at the thorax that has been shown to correlate with mood symptoms and affective states and differentiate between patients and controls (Actiheart\(^\text{32, 33}\)) like other wearable actigraphs.\(^\text{62–64}\)

**Methods**
All newly diagnosed/first-episode patients will have access to a smartphone-based system (Monsenso that is developed from the MONARCA system) for continuous self-monitoring as well as fine-grained automatically monitoring of behavioural activity and early identification of emerging affective episodes during the first 2 years and following relapse of episodes during a 5-year follow-up period.

**Data analyses**
In contrast to data in BIO-1, data in BIO-2 represent big data collected on a daily basis within individuals. We will use hierarchical Bayesian predictive models that can handle big data through sampling and visualisation techniques that summarise data.

**BIO-3: neurocognitive and brain imaging signatures in bipolar disorder**

**Aims**
(1) To identify an integrated brain-based biomarker of bipolar disorder including neurocognitive and neuroimaging measures tapping into ‘hot’ (ie, emotion laden) and ‘cold’ (non-emotional) cognition, (2) to examine whether the degree of abnormality in these measures predicts illness onset in the high-risk group and/or relapse in the patient group, (3) to identify developmental trajectories in ‘hot’ and ‘cold’ cognitive dysfunction and to identify structural MRI and fMRI correlates in bipolar disorder via longitudinal assessments of high-risk individuals to remission after onset of first manic episode and following successive relapses, and (4) to identify associations between aberrant ‘hot’ and ‘cold’ neurocognitive functions, structural and functional brain changes and the composite blood-based and smartphone-based biomarkers. Such ‘integrated systems approach’ involving identification of patterns of biomarkers (biosignatures) across these multiple levels of investigation is considered imperative for deeper understanding of the dimensions of underlying pathophysiological processes in bipolar disorder.\(^\text{12}\)

**Background and theoretical basis**
Results from a number of meta-analyses of a large number of cross-sectional studies of patients with bipolar disorder in remission suggest trait-related ‘cold’ cognitive deficits in attention/processing speed, memory and executive function compared with healthy controls\(^\text{65–68}\), that correlate with everyday functioning.\(^\text{69}\) Cross-sectional comparison of patients at different illness stages revealed more pronounced cognitive deficits during late
stage compared with early stage in line with the staging hypothesis of bipolar disorder.70 However, there are only a few longitudinal studies of neuropsychological functioning with small sample sizes (12 studies including a total of 152 patients with bipolar disorder71). A meta-analysis of these studies found no support for a progressive nature of cognitive deficits.71 However, results from these studies are hampered by a number of limitations including small sample sizes, short follow-up (mean follow-up period of 4.6 years) and high attrition rates (up to 45%).71 Consequently, it is unclear whether cognitive function assessed with neuropsychological tests deteriorates with the number and duration of illness episodes in bipolar disorder although epidemiological studies consistently revealed increased risk of dementia with the number of episodes7 (see also ref 9). Risk of developing dementia long-term6 8 72–74 and there is some evidence for

Deficits in ‘hot’ cognition are closely linked to emotional disturbances75 and difficulties in socioemotional behaviour and interpersonal relations in bipolar disorder.76 ‘Hot’ cognition abnormalities in bipolar disorder have been observed within three domains: (1) emotional processing, (2) reward processing and (3) emotion regulation (reviews in refs 77 78).

Results from a large number of cross-sectional structural imaging studies suggest that patients also show increased lateral ventricular volumes and greater prevalence of white matter hyperintensities.70 While these findings are rather unpecific, studies also suggest that treatment with lithium increases the grey matter volume of prefrontal cortex, amygdala and hippocampus.79 In addition, a number of functional imaging studies suggest that bipolar disorder is associated with abnormalities within fronto-limbic-subcortical structures.38

As long-term, integrative studies are lacking, it is unclear how neurocognitive and brain imaging abnormalities correlate with the staging of bipolar disorder, illness progression and treatment38 80 or with changes in biological markers such as neurotropic, inflammatory and oxidative stress markers.81 82

Methods

Using a comprehensive neurocognitive test battery, we will assess all participants from the three groups (patients with bipolar disorder, first-generation relatives and healthy individuals without a family history of affective disorders). Patients will be followed from first onset of affective disorder and during successive periods of remission or at an annual basis (patients with bipolar disorder with no relapse, first-generation relatives with no onset and healthy controls). Among these, a subgroup of 60 patients, 60 healthy relatives ( ultra risk) and 30 healthy individuals without a family history of affective disorders (healthy controls) will undergo functional and structural MRI at these time points.

Neurocognitive testing

Within ‘cold’ cognition verbal learning/memory and executive function have been highlighted as the most suitable candidates for biomarkers of bipolar disorder.83 84 ‘Cold’ cognition will therefore be assessed with neurocognitive tests probing verbal memory, attention and executive function including the Rey Auditory Verbal Learning Test85 86 and the WAIS-III letter-number sequencing, RBANS digit span, n-back working memory, verbal fluency and Trail Making Test B. Verbal intelligence will be estimated with the Danish Adult Reading Test.

‘Hot’ cognition will be assessed with a comprehensive battery of computerised neurocognitive tests outside the scanner probing (1) emotional processing, (2) reward processing and (3) emotion regulation. These include the facial expression recognition and faces using dot-probe tasks from the Emotional Test Battery (P1Vital, Oxford), and an ecologically valid social scenarios test developed by our group.87 During fMRI we will also administer the following experimental paradigms: (1) an emotional face processing task using face stimuli from the Nimstim (http://www.macbrain.org/resources.htm), (2) a monetary reward processing task88 and (3) a negative affective picture task using validated stimuli sets from the International Affective Picture System developed in collaboration with researchers at the University of Chicago. In addition, we will explore the neuronal basis for ‘cold’ cognition (executive function) using n-back working memory and picture encoding tasks programmed in-house.89 Finally, self-report measures (BIS/BAS, and the CERQ46) are used to assess reward responsiveness and habitual emotional regulation strategies.

Structural MRI and fMRI

Structural MRI

Using T1-weighted images acquired at a 3T Siemens scanner at the Copenhagen University Hospital, Rigshospitalet, we will focus on lateral ventricular volumes, grey matter volume of prefrontal cortex, amygdala and hippocampus, relative to whole brain volume. Specifically, segmentation and analysis of subcortical and regional cortical volume, shape and grey matter density will be conducted with FMRIB Software Library (FSL) tools, including the FMRIB’s Integrated Registration and Segmentation Tool, the FSL-VBM tool and FSL vertex (shape) analysis (http://fsl.fmrib.ox.ac.uk/).

Functional MRI

T2-weighted images will be acquired to investigate white matter hyperintensities. We will also use fMRI to investigate neuronal underpinnings of ‘hot’ and ‘cold’ cognition with the previously described experimental paradigms. fMRI data processing will be carried out with the FMRI Expert Analysis Tool, part of FMRIB’s Software Library (www.fmrib.ox.ac.uk/fsl). We will examine mean per cent BOLD signal change within predefined hippocampal and amygdala regions of interest obtained in standard space
with mri3dX (http://www.idoimaging.com/program/160). In addition, whole-brain exploratory analysis will be conducted to explore neural activity differences in other cortical regions. For this group analysis, individual contrasts of interest will be included in separate general linear models with non-parametric permutation inference (n=5000) using the ‘randomize’ algorithm implemented in FSL.91

Statistical power
The above sample size for participants undergoing fMRI assessments is determined based on our previous fMRI studies. In particular, inclusion of about 17–22 participants per treatment/diagnostic group (matched for age and gender) had a power of >0.8 to detect differences between groups in neural and cognitive response to emotional faces (eg, refs 92 93 and Miskowiak et al, under review) at an alpha level of <0.05 for cross-sectional designs. For longitudinal designs, we were able to demonstrate differences between groups in the change in task-related neural activity and in hippocampal structure with a sample of about 40 participants per group.94–96 Given this, our inclusion of 60 participants per group is expected to ensure adequate statistical power for both the cross-sectional and longitudinal parts of the fMRI study.

BIO-4: at risk or prodromal phase for bipolar disorder
Aims
To test whether (1) the composite blood-based biomarker, (2) the composite electronic smartphone-based biomarker and (3) the neurocognitive signature for bipolar disorder predict onset of illness (depression or mania) among a healthy (ie, non-syndromal level) high-risk population of first-degree siblings and offspring to patients with newly diagnosed/first-episode mania/bipolar disorder, included in BIO-1, BIO-2 and BIO-3.

Background and theoretical basis
First-generation relatives to patients with bipolar disorder have a ninefold increased risk of developing bipolar disorder and a twofold to threefold increased risk of developing unipolar disorder.1 Although there is a great need for early detection and primary prevention of onset of illness among relatives to patients with bipolar disorder, recent prior attempts have not been successful due to retrospective designs, poor characterisation of high-risk individuals and small sample sizes.97 This is the first time a composite blood-based, a composite smartphone-based and a composite neurocognitive biomarker identified among patients with bipolar disorder will be investigated among their healthy relatives as a predictor measure of onset of illness.18 This approach increases the chances of obtaining a high specificity and sensitivity of the composite biomarkers.

Methods
This BIO-4 includes first-generation relatives (siblings and offspring aged 15–40 years) to the recruited patients with first manic or bipolar diagnosis. All recruited patients will be asked about the lifetime psychiatric history of first-degree relatives (their biological parents, siblings and offspring) based on the Brief Screening for Family Psychiatric History questionnaire described by Weissman and colleagues.98 We expect that at least 200 first-generation relatives will be asymptomatic or present mild symptoms or prodromal patterns to affective disorders and will be included in the study. Biological tissues will be drawn (as part of BIO-1) and neurocognitive function will be assessed on all individuals and brain imaging will be done on 60 individuals with longitudinal assessments.

Statistical analyses
BIO-4 will use a combination of penalised regression techniques, and random forests to infer the importance of the original and combined markers and to compare the similarity of prediction error from the models with and without combining the markers.

Feasibility
The BIO study is fully feasible as patients are recruited as part of daily healthcare for patients referred to the Copenhagen Affective Disorder Clinic.

Ethical considerations
The BIO study has been approved by the Local Ethical Committee (H-7-2014-007) and the data agency, Capital Region of Copenhagen (RHP-2015–023). According to these specifications, adolescents aged 15–18 years will be invited only if parents have given consent. Data will be saved, encrypted and assessed according to the regulations from the Capital Region of Denmark. The study complies with the Declaration of Helsinki principles (Seoul, October 2008).

Dissemination
Study results will be presented in peer-reviewed journals and at international conferences in accordance with relevant reporting guidelines.99 100

DISCUSSION
Summary
It is expected that the BIO cohort will provide valid biological, electronic, neurocognitive and neuroimaging data and for the first time longitudinally identify changes in biomarkers during different stages of bipolar illness, that is, at risk stage, following onset, during first relapse and recurrent relapses and in the late stages of the illness as according to Berk et al.15

Limitations
Some limitations of the BIO study should be noted beforehand. First, the rather extensive initial assessment study procedure may potentially result in selection of participants who are intrinsically positive towards clinical research and readily willing to cooperate. Nevertheless, we expect that such selection will be decreased as it is likely that the vast majority of more than 100 newly diagnosed/
first-episode patients with bipolar disorder referred to the Copenhagen Affective Disorder Clinic each year will accept participation in the BIO study as this implies an extensive clinical evaluation.

Second, attrition may increase during long-term follow-up and patients who stay in the study may adhere more to treatment in general. Such selection is inherent in clinical longitudinal research, and the large number of participants that will be included will increase external validity. Third, potential confounding effects of psychotropic medication may influence comparisons between patients with bipolar disorder and healthy control individuals as well as comparisons within patients as the vast majority of patients will get medication that may change during the course of illness. Lithium, mood stabilisers and antipsychotics may have effects on the collected biological, smartphone-based, neuropsychological and brain imaging data. Effects of medication on biological measures are not clear although analyses from systematic reviews and meta-analyses involving only patients with bipolar disorder have not found clear effects of medication on cytokines, BDNF and gene expression nor have subsequent individual studies on cytokines, BDNF, gene expression or DNA and RNA damage. Effects of medication on electronic smartphone-generated data as well as on neuropsychological and brain imaging data are poorly investigated and warrant further studies.

Fourth, due to the large number of biological and statistical tests included in the BIO study, chance findings may occur in relation to the individual biological test. However, the aim of the BIO study is to identify a composite biomarker measure related to bipolar illness, depression and mania using cross validation or alternatively split sample designs in the development and validation of the composite biomarker.

Finally, the BIO study does not include (randomised) interventions limiting causal interpretations of the results. Nevertheless, with the BIO prospective, repeated measures design it is possible to identify valid associations between the composite measures (of biological, electronic, neuropsychological and brain imaging data) and depressive and manic symptoms and states.

Strengths
The BIO study is the first study aiming to identify (1) a composite blood-based biomarker, (2) a composite electronic smartphone-based biomarker and (3) a neurocognitive signature for bipolar disorder as well as to measure the same biomarkers in newly diagnosed/first-episode patients with bipolar disorder and their healthy first-generation relatives. It is possible to recruit newly diagnosed/first-episode patients with mania/bipolar disorder as all such patients from the entire Capital Region of Denmark are referred to the Copenhagen Affective Disorder Clinic and routinely asked for inclusion in the BIO study. Long-term attrition is supposed to be low as all patients will be followed by the Copenhagen Affective Disorder Clinic for the first 2 years and subsequently in other treatment settings in the Capital Region of Denmark. Including longitudinal assessment of healthy individuals is of paramount importance to control for the effect of timing and ageing, but among all studies on biomarkers in bipolar disorder, this has been done only in two studies from our group. The study will be performed by an experienced international research group, combining expertise from all areas of the study.

CONCLUSION
The BIO study is a large long-term cohort study on biomarkers in bipolar disorder and we expect that the findings for the first time will be representative of biomarkers in bipolar disorder in general as no prior study on newly diagnosed/first-episode bipolar disorder has been conducted. It is expected that the BIO cohort will provide valid biological, electronic, neuropsychological and brain imaging longitudinal data.

Author affiliations
1Department of Psychiatry, Psychiatric Center Copenhagen, Copenhagen, Denmark
2Institute of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark
3Department of Clinical Biochemistry, Rigshospitalet, Copenhagen, Denmark
4Department of Biostatistics, University of Copenhagen, Copenhagen, Denmark
5Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Denmark
6Gene Regulation Bioinformatics at the Bioinformatics Centre, Department of Biology/BRIC, University of Copenhagen, Copenhagen, Denmark
7The Centre of Inflammation and Metabolism at Department of Infectious Diseases, Rigshospitalet, Copenhagen, Denmark
8Laboratory of Clinical Pharmacology, Rigshospitalet, Copenhagen, Denmark
9Department of Psychiatry and Pharmacology, University of Toronto, Brain and Cognition Discovery Foundation, Toronto, Ontario, Canada
10McMaster University, Hamilton, Ontario, Canada
11Department and Institute of Psychiatry, and Laboratory of Neuroscience (LIM27), University of São Paulo Medical School, São Paulo, Brazil
12IT University Copenhagen, Copenhagen, Denmark
13Create-Net: Center for Research and Telecommunications Experimentation for Networked Communities, Trento, Italy
14Neurobiology Research Unit and Center for Integrated Molecular Brain Imaging, Rigshospitalet, Copenhagen, Denmark
15Department of Psychiatry, University of Pittsburgh, Western Psychiatric Institute and Clinic, Pittsburgh, Pennsylvania, USA

Correction notice This paper has been amended since it was published Online First. Owing to a scripting error, some of the publisher names in the references were replaced with 'BMJ Publishing Group'. This only affected the full text version, not the PDF. We have since corrected these errors and the correct publishers have been inserted into the references.

Acknowledgements The validity of the research cannot be influenced by any of these potential secondary interests (such as financial gain or personal relationship).

Contributors LVK designed the study together with MV, KM, KW and MFJ. LVK drafted the study protocol and the manuscript. All authors contributed to development of the study protocol and to editing the manuscript and read and approved the final version. KM, LBN, RF5, CE, BKP, HEP, RSM, FK, WFG and MV contributed specifically to BIO-1 on peripheral blood-based biomarkers. MFJ, OW, JB, MF and OM contributed specifically to BIO-2 on smartphone-based electronic biomarkers. KWM, GMK and MP contributed specifically to BIO-3 on neurocognitive and brain imaging signatures. MV contributed specifically to BIO-4 on risk or prodromal phase of bipolar disorder (in addition to BIO).

Funding The study is funded by grants from the Mental Health Services, Capital Region of Denmark, TheDanish Council for Independent Research, Medical
Sciences (DFF-4183-00570), Weimans Fund, Markedsmodningsfonden (the Market Development Fund 2015-310), Gangstefonden (A29594), Helsefonden (16-8063), Innovation Fund Denmark (the Innovation Fund, Denmark, 5164-00001B), Copenhagen Center for Health Technology (CACHET), EU H2020 ITN (EU project 722561), Augustinusfonden (16-0083), Lundbeck Foundation (R215-2015-4121).

Competing interests UK has within the preceding three years been a consultant for Lundbeck, AstraZeneca and Sunovion. KWM has received consultancy fees in the past three years from Lundbeck and Allergan. MFJ has been a consultant for Eli-Lilly and Lundbeck. MV has within the preceding three years been a consultant for AstraZeneca and Servier. FK has been a speaker for Ache, Daiichi Sankyoand Janssen. MF is a consultant for Roche Pharmaceuticals. RSM: Advisory boards: Lundbeck, Pfizer, AstraZeneca, Eli-Lilly, Janssen, Ortho Purdue, Johnson & Johnson, Moskha,s Sunovion, Mitsubishi, Takeda, Forest, Otsuka, Bristol-Myers Squibb, Shire. Speaker fees: Lundbeck, Pfizer, AstraZeneca, Eli-Lilly, Janssen Ortho, Purdue, Johnson & Johnson, Moskha,s Sunovion, Mitsubishi, Takeda, Forest, Otsuka, Bristol-Myers Squibb, Shire. Research grants: Lundbeck, Janssen Ortho, Shire, Purdue, AstraZeneca, Pfizer, Otsuka, Allergan. HEP has received a research grant from Boehringer Ingelheim. GMK has received honoraria as Field Editor of the International Journal of Neuropsychopharmacology and as scientific advisor for H Lundbeck A/S. JB and MF are co-founders and shareholders in Monsenso. LBN, RFS and WFG declare no competing interests. The validity of the research cannot be influenced by any of these potential secondary interests (such as financial gain or personal relationship).

Patient consent The informed consent form from the Local Ethical Committee in Copenhagen is signed by the patients and the researcher.

Ethics approval The study has been approved by the Local Ethical Committee (H-7-2014-007) and the data agency, Capital Region of Copenhagen (H-2015-023).

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, provided the original work is permitted under any of these potential secondary interests (such as financial gain or personal relationship).

REFERENCES


