

INTEGRATION OF COMPLEMENTARY BIOMARKERS IN PATIENTS WITH FIRST EPISODE PSYCHOSIS: RESEARCH PROTOCOL OF A PROSPECTIVE FOLLOW UP STUDY

Martina Rojnic Kuzman^{1*}, Porin Makaric^{2*}, Dina Bosnjak Kuharic², Ivana Kekin¹,
Linda Rossini Gajsak³, Marina Boban¹, Nada Bozina¹, Tamara Bozina⁴, Mirela Celic Ruzic³,
Sanja Darmopil⁴, Igor Filipcic⁵, Lana Ganoci¹, Ana Hladnik⁴, Zoran Madzarac¹, Branko Malojcic¹,
Alma Mihaljevic Peles¹, Daniel J. Mueller⁶, Drazenka Ostojic², Zdravko Petanjek⁴, Ratimir Petrovic¹,
Zeljka Vogrinc¹, Aleksandar Savic², Ante Silic⁷, Marina Sagud¹, Maja Zivkovic¹ & Zarko Bajic⁸

¹Zagreb University Hospital Centre, Zagreb, Croatia

²University Psychiatric Hospital Vrapce, Zagreb, Croatia

³Neuropsychiatric Hospital "Dr. Ivan Barbot", Popovaca, Croatia

⁴Zagreb School of medicine, Zagreb, Croatia

⁵Neuropsychiatric Hospital "Sveti Ivan", Zagreb, Croatia

⁶Pharmacogenetics Research Clinic, Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON, Canada and Department of Psychiatry, University of Toronto, Toronto, ON, Canada

⁷Department of Psychiatry, Clinical Hospital Centre "Sestre Milosrdnice", Zagreb, Croatia

⁸Biometrika Healthcare Research, Zagreb, Croatia

* These authors have contributed equally to this manuscript

received: 20.11.2018;

revised: 14.2.2019;

accepted: 11.4.2019

SUMMARY

In this project, we recruited a sample of 150 patients with first episode of psychosis with schizophrenia features (FEP) and 100 healthy controls. We assessed the differences between these two groups, as well as the changes between the acute phase of illness and subsequent remission among patients over 18-month longitudinal follow-up. The assessments were divided into four work packages (WP): WP1- psychopathological status, neurocognitive functioning and emotional recognition; WP2- stress response measured by saliva cortisol during a stress paradigm; cerebral blood perfusion in the resting state (with single photon emission computed tomography (SPECT) and during activation paradigm (with Transcranial Ultrasonography Doppler (TCD)); WP3-post mortem analysis in histologically prepared human cortical tissue of post mortem samples of subjects with schizophrenia in the region that synaptic alteration was suggested by WP1 and WP2; WP4- pharmacogenetic analysis (single gene polymorphisms and genome wide association study (GWAS)). We expect that the analysis of these data will identify a set of markers that differentiate healthy controls from patients with FEP, and serve as an additional diagnostic tool in the first episode of psychosis, and prediction tool which can be then used to help tailoring individualized treatment options. In this paper, we describe the project protocol including aims and methods and provide a brief description of planned post mortem studies and pharmacogenetic analysis.

Key words: schizophrenia - first episode psychosis - biological markers - neurocognition - TCD - SPECT - stress - cortisol

* * * * *

INTRODUCTION

Schizophrenia is one of the most complex psychiatric illnesses, affecting about 1% of population worldwide. Its complexity arises both from the difficulty in diagnosing the disorder and achieving functional recovery of affected persons, which in turn all contributes to high emotional, social and financial burden to patients and their families, and increased stigma surrounding the term "schizophrenia".

In both clinical and research settings, the diagnosis is made based on a clinical interview, following the criteria of the Diagnostic and Statistical Manual, 5th revision (DSM 5) (American Psychiatric Association 2013) or International Classification of Diseases, 10th revision (World Health Organization 1992). In the majority of cases, schizophrenia is a chronic illness with

a recurrent course, characterized by alternating periods of acute psychotic illness and their remission (an der Heiden & Häfner 2000). About a third of patients will respond to treatment with fair treatment response, about a third will have poor treatment outcome, while a third will have a variable outcome, somewhere in between (Levine et al. 2012). However, the clinical presentation as well as treatment outcomes are influenced by a number of factors such as the comorbid use of drugs, by adherence to treatment, by side effect of medication, availability of treatment services, etc. (Weiden et al. 2004, Miller et al. 2009, Tessier et al. 2017, Correll et al. 2018).

Thus, despite the fact that huge efforts have been made to identify a biological correlate of schizophrenia that can objectify the diagnosis and course of treatment, there are still no straightforward results. This implies

that there are still no reliable “tests” for the diagnosis of schizophrenia, nor “tests” that helps clinicians to objectify the course of neither treatment nor prognostic outcomes.

The most replicated findings on biomarkers that differentiate patients with schizophrenia from controls or make distinction among specific subsets of patients include psychopathological (e.g. negative symptoms) (Fusar-Poli et al. 2014) and neuropsychological assessments (e.g. verbal memory, verbal fluency) (Lepage et al. 2014, Mesholam-Gately et al. 2009), peripheral blood (e.g. prolactine levels, homocystein levels) (Riecher-Rössler et al. 2013), neuroimaging methods (e.g. brain volume, gray matter) (Fusar-Poli et al. 2013), post mortem studies (Glausier & Lewis 2013, Lewis 2014) and more recently genetic markers (Juraeva et al. 2014, Ripke et al. 2014). In research that looked after the factors influencing treatment response/progression of illness, several factors seem to be relevant, such as the value of neurocognitive tests (e.g. stabile deficits in verbal fluency) (Mesholam-Gately et al. 2009), neuroimaging data (Radua et al. 2012), as well as pharamacogenetic markers (Åberg et al. 2010, McClay et al. 2011, Nicodemus et al. 2014). However, there are significant differences among results of those studies, some arising from the confounders in those studies. New research, designed to decrease the number of confounders is finally emerging in the literature, such as studies reporting on longitudinal studies of patients with first episode psychosis (FEP) (Cotton et al. 2017, Kahn et al. 2008). Patients with FEP are more homogenous group of patients with schizophrenia compared to patients with multiple episodes as they are usually younger, drug naive, or treated with less medication, with a more homogenous clinical presentation (predominantly acute positive symptoms, followed by subacute phase), and more homogenous response to medication then patients with multiple-episodes (Emsley et al. 2013), although not exclusively (Rosen et al. 2012, Slotema et al. 2018). Thus, the inclusion of different features which may help the differentiation of type of FEP (for example affective vs. non affective FEP) and longitudinal follow up of these patients further increase the reliability of results and produce relevant new knowledge.

Secondly, all research conducted so far has led to the conclusion of a complex, heterogeneous disorder with possible various causes which is why none of the simplistic approaches have yielded much and why we therefore need integration of data and different approaches.

In this study, we examined a set of promising biomarkers including the psychopathological status, neurocognitive functioning, and laboratory biomarkers stress response by using different methods during resting state and under activation in 150 FEP patients vs. controls and their changes during the FEP and subsequent remission at 18-month longitudinal follow-up. Specific

objectives and expected results relate to four complementary areas of research, performed through different work packages (WP1-WP4). In this paper, we describe the project protocol, including aims and methods, focusing exclusively on clinical data (WP1, WP2). We provide only a brief description of planned post mortem studies in histologically prepared human cortical tissue of subjects with schizophrenia (WP3) and analysis of association pharmacogenetic markers and the treatment response (WP4).

With this study design two main hypotheses relating to WP1 and WP2 were tested:

- Patients with FEP have significant neurocognitive deficits and deficits in facial emotional recognition, as well as deficits in brain cerebral perfusion and altered stress response compared to healthy controls;
- Psychopathology, neurocognition and facial emotional recognition, as well as deficits in brain cerebral perfusion and stress response in patients with FEP significantly improve with 18-month treatment compared to their initial results.

SUBJECTS AND METHODS

Participants and protocol

The sample consisted of 150 patients with first episode psychosis and 100 healthy individuals. Patients were recruited from four hospitals in Croatia, Zagreb University Hospital Centre (ZUHC), University Psychiatric Hospital Vrapce (UPHV), Psychiatric Hospital “Sveti Ivan” (PHSI) and the Psychiatric hospital “Dr. Ivan Barbot” (PHIB), in the period from October 2015 until February 2018. Inclusion criteria for FEP were: no history of antipsychotic treatment prior to admission to the hospital, first episode of psychosis and meeting of the criteria for psychotic episode (codes F23, F29) according to the criteria of ICD-10 (World Health Organization 1992). Exclusion criteria were: age <18 years, mental retardation, mental illness in childhood that can present with psychosis, neurological disorders, pregnancy and lactation, organic psychosis, the use of medications that can produce psychotic reactions, comorbid alcoholism or other addictions, use of drugs (more than up to 3 times a year). A convenient sample of healthy volunteers, with no personal or family history of psychiatric disorders was chosen that matched study subjects in respect to age and sex. The study protocol was approved by the Ethics Committees of all hospitals participating in the study. All participants have signed an informed consent form before the enrollment. The study was performed in accordance with the World Medical Association Declaration of Helsinki 2013 (World Medical Association 2013).

Specific tasks were divided according to areas of research, performed through different work packages (WP1-WP4).

WP1

WP1 was performed among all participants and incorporated data obtained with clinical and neurocognitive testing and laboratory features in FEP patients. All participants filled sociodemographic data, and a set of self-assessments scales. In addition, trained researchers performed the clinical rating, emotional and neurocognitive assessments, as described below. All assessments were performed at two time points, first during the first three weeks of treatment at the first admission to psychiatric services for their acute phase of the illness, and the second after the period of 18 months of treatment. Also, all patients had their blood taken for laboratory analyses and genotyping. Primary outcomes were results from psychopathology scales, number of correctly identified emotions, and neurocognitive tests in two time points. Secondary outcomes are laboratory findings (hormones and lipids) in several time points and results from other psychiatric and self-assessment scales.

Self-assessments scales

a) Inventory of Depressive Symptomatology-Self Report, IDS-SR (Rush et al. 1996) is a 30-item questionnaire for self-assessment of depressive symptoms with total score ranging from 0 to 84 (higher score indicating higher levels of depressive symptoms).

b) Barratt Impulsiveness Scale-11, BIS-11 (Patton & Stanford 1995) was used for assessment of impulsiveness. It is composed of 30 items designed to assess three dimensions of impulsiveness: attention (ability to concentrate/ to focus attention), motor (tendency to act without previous thinking) and non-planning (lack of planning the future). It includes BIS-11 Total score and 3 subscores represented with adequate items: Attention (8 items), Motor (11 items) and Non-planning (11 items). Higher scores represent more impulsive behavior.

c) Aggression questionnaire, AQ (Buss & Perry 1992) is a 34-item questionnaire measuring aggression in adult population. Besides overall aggression presented with the AQ Total score, items are organized in four factor subscores: Physical aggression (9 items), Verbal aggression (5 items), Anger (8 items) and Hostility (8 items). Higher scores represent higher aggression rates.

d) World Health Organization Quality of Life Assessment, WHOQOL-BREF (Whoqol Group 1998) was used for assessment of overall perception of quality of life and health, as well as for four specific domains: physical health, psychological, social relationships and environment. Raw scores were transformed by the instructions of the World Health Organization with higher score meaning higher perceived quality of life.

e) Parental Bonding Instrument, PBI (Parker et al. 1979) is a 25-item scale measuring bonding and attach-

ment between child and parents during the first 16 years of child's life. The results of two domains, Care (12 items) and Overprotection/ Control (13 items), are combined into four quadrants: Affectionate constraint (high care and high overprotection), Optimal parenting (high care and low overprotection), Affectionless control (low care and high overprotection) and Neglectful parenting (low care and low overprotection), presenting parental styles as perceived by the child.

f) The Holmes-Rahe Stress Inventory (The Social Readjustment Rating Scale), STRESS (Holmes & Rahe 1967) is a questionnaire used for assessment of levels of stress. It consists of 43 items, with scores above 150 implying a higher chance of major stress-related health problem within two years.

g) Questionnaire on suicide ideation and behaviour, SUICIDE (Marušič et al. 2007) was used for the assessment of suicidality, It is a 14-item questionnaire divided in three sections. First section is composed of nine questions in yes/no format assessing passive and active suicide thoughts, suicide behaviour and suicide attempt, second is a question regarding desire for attempting suicide and third is comprised of questions regarding personal and family history of attempted and completed suicides.

Clinical rating of symptoms

a) Positive and Negative Syndrome Scale, PANSS (Kay et al. 1987) is a 30-item scale used for assessment of overall psychotic symptoms and three subdomains: Positive (7 items), Negative (7 items) and General (16 items). Higher scores represent higher levels of psychopathology.

b) Calgary Depression Scale for Schizophrenia, CDSS (Addington et al. 1990) is a nine-item instrument measuring depressive symptoms in population of patients with schizophrenia. Score higher than 6 shows 82% specificity and 85% sensitivity for predicting the major depressive episode (Addington et al. 1993).

c) Young Mania Rating Scale, YMRS (Young et al. 1978) is an 11-item scale used to assess symptoms of mania, with higher scores indicating more severe manic symptoms.

d) The Global Assessment of Functioning, GAF (American Psychiatric Association 1994) was used for measuring social, occupational and psychological functioning (impairment). Higher scores represent better functioning.

Neurocognitive testing

Neurocognitive assessment composed of Mini Mental Status Examination (MMSE), a 30-item screening test for overall cognitive impairment (Folstein et al. 1975) and a number of specific neurocognitive tests previously used in Croatian population:

a) Rey Auditory Verbal Learning Test (RAVLT) (Schmidt 1996) included two subtests for measuring immediate and one for measuring delayed verbal memory. Scores were composed of number of correctly recalled words from a 15-word list, higher score meaning better results.

b) Wechsler verbal paired associates (Wechsler 1945) included two subtests for assessing immediate and delayed recall of series of verbal paired associates. Higher score represented better results.

c) Digit span test (Lichtenberger & Kaufman 2009) was also used for assessment of verbal memory of numbers. We used two subtests, one for forward and one for backward repeating series of numbers. Better performance was presented with higher score (more correctly repeated series).

d) Block design test (Block design) (Hutt 1932) was used for assessment of executive functions. In this test, participants use blocks to recreate different pictures shown to them while the rater measures time in seconds necessary to finish the task. Results are transformed to numbers, with higher number meaning better results.

e) Frontal assessment battery (FAB) (Dubois et al. 2000) measures conceptualization, mental flexibility, programming, sensitivity to interference, inhibitory control and environmental autonomy. Higher scores represent better performance.

f) Clock drawing test (CDT) (Freedman et al. 1994) was used for assessment of executive functioning. Scores range from 0 to 10 with higher results showing better performance.

g) Stroop test (Golden 1976) included three subtests, one for assessing attention and processing speed (Stroop words) and two for assessing executive functioning (Stroop colours and Stroop word-colours). Result is presented in seconds need for finishing the test with higher results representing worse performance.

h) Trail Making Test (Tombaugh 2004) included two subtests. In Trail making test A used for assessment of attention and processing speed, participant links series of numbers, while in Trail making test B participants links series of numbers and letters. Results are presented as seconds needed for finishing the task (higher result= worse performance).

i) Digit symbol test (Digit symbol) (Lichtenberger & Kaufman 2009) was used for assessment of attention and processing speed. Performance is scored as a number of correctly coded list of numbers.

j) Rey-Osterrieth Complex Figure Test (Fastenau et al. 1999) was used for assessment of immediate and delayed visuospatial abilities. Better performance was presented with higher results.

k) Semantic (category) and Phonetic fluency test, (Semantic, Phonetic) (Lichtenberger & Kaufman 2009) were used for the assessment of language functions. Scores was calculated from the number of different words said for each category (categories “animals”, “vegetables” and “supermarket” for semantic, and letters “a” and “f” for phonetic fluency).

Emotion recognition assessment

a) Penn Emotion Recognition Task, ER40 (Gur et al. 2002) used for assessment of facial emotion recognition, is composed of 40 color photographs shown to participant on a monitor, one photograph at a time in random order. Photographs present male and female faces expressing five emotions (happiness, sadness, anger, fear and neutral). Results are presented as correct identification of each of the five emotions.

b) The Infant Expressions of Emotions from Looking at Pictures Task, IFEEL (Emde et al. 1993) is a projective test for evaluation of individual differences in attributions of emotions consisting of 30 pictures showing ambiguous facial expressions in infants. Participants are asked to describe with one word the emotion that the infant in the photograph expresses. The word used for description of the emotional expression is allocated to one of twelve categories of affect (Surprise, Interest, Joy, Content, Passive, Sad, Cautious-Shy, Shame-Guilt, Disgust-Dislike, Anger, Distress, Fear) and category for non-decodable or unspecified descriptions, and the total sum of words used per category was counted.

WP 2 subset

WP2 was performed in a subset of participants who were treated at the Zagreb University Hospital Centre only, as described below. Specific aims included: 1) examination of alterations of the cerebral blood flow velocities (CBFV) in the main intracranial arteries using Transcranial Doppler (TCD) ultrasonography in relation to neurocognitive and stress paradigm in FEP subjects compared to control subjects, and then comparison of the results of the FEP patients at baseline and after 1.5 years of treatment; 2) examination of the baseline alterations of the regional cerebral flow (rCBF) using SPECT and its changes in relation to 1.5-year treatment with medication in FEP patients; 3) examination of the difference in stress response patterns in patients and controls measured from saliva during the stress paradigm and then reviewing the results of the FEP patients at baseline and after 1.5 years of treatment. Primary outcome measures were: 1) velocity of the cerebral blood flow (CBFV) with TCD; 2) brain perfusion deficits across brain regions with SPECT, 3) levels of saliva cortisol in five time points during stress paradigm. Secondary outcomes measures are results of psychiatric and self-assessment scales as described below.

Subjects

The sample consisted of 46 patients and 45 controls. We have chosen a consecutive sample of patients by the order of their admission to the hospital at the Department of Psychiatry, University Hospital Centre Zagreb in the period from January 2016 to December 2017. A convenient sample of healthy volunteers, with no personal or family history of psychiatric disorders was chosen that matched study subjects in respect to age and sex. Inclusion as well as exclusion criteria were the same as for WP1.

Protocol

Participants followed the protocol described for all patients in WP1, with several exceptions: all participants who were involved in the stress paradigm, also filled Life Events Questionnaire (Norbeck 1984), International Personality Item Pool (IPIP) (Ashton et al. 2007), Rosenberg (Rosenberg & Press. 1965); All participants involved in the TCD testing were tested for handedness by using Edinburgh Handedness Inventory (EHI) (Oldfield 1971).

Experimental paradigm for TCD

Before and after computer paradigm was performed, blood pressure (BP), heart rate (HR) and visual analogue anxiety scale (VAS) were assessed. All cognitive tasks were verbally explained to subjects prior to testing. During performance of cognitive tasks subjects were sitting in a quiet room looking at a computer screen showing ongoing paradigm lasting approximately 25 minutes. TCD monitoring of CBFVs in both MCAs and ACAs as well as HR were performed during testing.

The paradigm consisted of three cognitive tasks: Phonemic verbal fluency test (pVFT), Stroop test with incongruent stimuli (Golden 1976) and Trial Making Test B (Tombaugh 2004). Prior to testing breath holding test (BHT) was performed on all of the subjects for 30 seconds. Resting periods were two minutes between each major task and one minute between each subtask. Resting intervals between tasks and subtasks were proved (on preliminary testing) to be sufficient for BFV and heart rate to return to baseline values. The paradigm was chosen base on the previous work of the research group who showed that these tests elicited CBV changes in ACAs and MCAs in a healthy population (Boban et al. 2014a, Boban et al. 2014b).

The monitoring was done with the 2 MHz pulse-wave TCD device (Doppler-BoxX, DWL) probes placed bilaterally over trans temporal window at a 55-60 mm insonation depth which is related to M1 segments of MCAs (Aaslid et al. 1982, von Reutern et al. 2000). The probes were held in a place by the head frame. Time resolution of the device was 0.01 s.

Mean values were calculated automatically by the system using equation: $MBFV (V_{sis} + 2V_{dis})/3$; (MBFV = mean blood flow velocity, V_{sis} = peak systolic

velocity, V_{dis} = peak diastolic velocity). Averaging of MBFV values was done for sequential 5-s time intervals for each subtask separately reaching a final mean value for each task and each subject. Averaging of 5 s intervals was done with 0.5 s of overlapping between intervals (Szirmai et al. 2005). Distance between V_{sis} peaks was used to calculate the heart rate which was averaged using the same 5-s intervals (Szirmai et al. 2005). Considering significant difference in actual values of MBFV among participants we used relative MBFVs ($MBFV_{rel}$) that were calculated by the following equation: $MBFV_{rel} = (MBFV_{act}/MBFV_{ref}) \times 100$; MBFV act being the actual MBFV during the observed interval and MBFV ref being the MBFV velocity during the -15 to -3 prestimulus period of the resting interval (Knecht et al. 1998, Knecht et al. 2000).

Protocol for the SPECT

SPECT of the brain was performed in accordance with the guidelines of the European Association of Nuclear Medicine (Kapucu et al. 2009). Patients were scanned with Symbia[®] T cameras, Version 4.1., which introduce TruePoint SPECT[™]/CT (CT-computer tomography), a new technology combining SPECT with the precision of multi-layer CT. This technology allows us to precisely determine the location, size, type, and spread of the disease. Symbia cameras are the result of two different modalities working as one, allowing three types of scans: SPECT, CT, and SPECT/CT. Symbia cameras use HD digital detectors for SPECT and ultra-fast ceramic (UFC[™]) CT detectors for higher quality CT scans. The system includes a network of workstations for data acquisition and analysis. The technical specifications of Symbia T - TruePoint SPECT/CT are as follows (2 slices, 1-10 mm slice thickness, 0.8 sec rotation time, 3.5 MHU tube, 40 kW generator, 30-240 mA tube voltage, CARE Dose, SureView). The cameras are housed at the Department of Nuclear Medicine and Radiation Protection, Zagreb University Hospital Centre, in spacious well-lit rooms that are adequately equipped and in line with radiation protection regulations.

Protocol for the stress paradigm

We used the modification of the Montreal Imaging Stress Task (MIST) (Dedovic et al. 2005) and Trier Social Stress Test (TSST) (Dressendorfer et al. 1992, Kirschbaum et al. 1993). The combination of a public speaking and a cognitive task produces a robust cortisol increase (Kirschbaum et al. 1993) probably determined by their association with social evaluative threat and uncontrollability, two important characteristics of psychological stressors to induce strong cortisol responses (Dickerson & Kemeny 2004).

In short, the test was performed as follows: first after a saliva sample was collected for basal levels of free salivatory cortisol, the participants were given the instruction to prepare (3 minutes). Afterward, the partici-

participants returned to the TSST room on two separate occasions, where they took part in a simulated job interview (5 minutes) followed by a mental arithmetic task (5 minutes) in front of an audience of three non-biased mental healthcare professionals of both sexes who were not permitted to express any kind of facial expression neither emotional reaction and were not dressed in formal white coats. To assess salivatory cortisol levels, a saliva sample was taken at five time points during the test, immediately before and after the TSST, with further samples taken at 30 minutes after the TSST.

In order to minimize the impact of diurnal rhythm on the cortisol responses, all participants were seen in the afternoon between 13:00 and 16:00 h. Participants were informed about the nature of the task after baseline saliva samples had been collected.

Saliva cortisol was analysed at Department of Laboratory Diagnostics Zagreb University Hospital Centre, on fully automated analyzer Roche cobas c6000, by electrochemiluminescence (ECL) technology (Chiu et al. 2003);. The test that was used in study is intended for use for the in vitro quantitative determination of cortisol in human serum, plasma, urine, and saliva. Collecting of saliva samples was performed according to previously described protocol (Hanrahan et al. 2006). Based on the recommendation of the manufacturer, using the Elecsys Cortisol methods normal salivatory cortisol values were <19.1 nmol/l in the morning 8.00-10.00 a.m., and <11.9 nmol/L in the afternoon 14.30-15.30 p.m. (Chiu et al. 2003). WP 1 and WP 2 protocol are summarized in Figure 1.

Statistical analysis

Sample size calculation

Data required to calculate the needed sample size were obtained by pilot study on 14 FEP patients treated at the Clinical Hospital Center in Zagreb in 2015 and 17 patients with no psychotic symptoms. Those pilot study participants were not enrolled in the main study described by this protocol. The required sample size was calculated for the most demanding objective and the outcome with the planned statistical power of 0.80, the two-tailed statistical significance $p < 0.05$. Under these conditions we needed $n = 98$ participants in each group. Anticipating up to 10% of incorrectly collected data and participants who will be lost for follow-up, the initially required sample size was estimated at $n = 109$ participants in each group. We performed the power analysis using the PASS 14 Power Analysis and Sample Size Software (2015). NCSS, LLC. Kaysville, Utah, USA, ncss.com/software/pass.

For the WP2 subset, data required to calculate the needed sample size were obtained by pilot study on 14 FEP patients treated at the Clinical Hospital Center in Zagreb in 2015 and 17 patients with no psychotic symptoms. Those pilot study participants were not enrolled in the main study described by this protocol. The required sample size was calculated for the interaction of the measurement time during the Trier's psychosocial stress test and the study group using a mixed between-within subjects analysis of covariance, with the planned statistical power of 0.80, the two-tailed statistical significance $p < 0.05$, two repeated measurements, the minimum

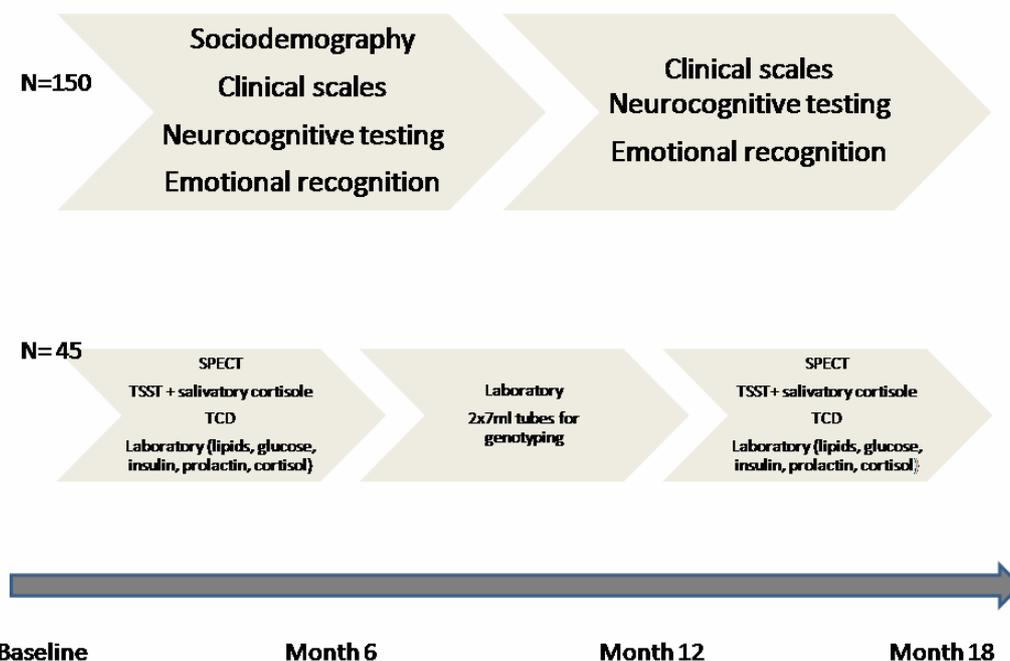


Figure 1. Protocol for WP1 and WP2

expected correlation between the repeated measurements of $r \geq 0.60$ and the minimum magnitude of the effect we consider to be clinically relevant and which we want to be able to determine statistically significant of: partial $\eta^2 = 0.04$, corresponding to the Cohen's "small" to "moderate" standardized effect size of $f = 0.20$. Under these conditions we need $n = 40$ participants in each group. Anticipating up to 10% of incorrectly collected data and participants who will be lost for follow-up, the initially required sample size was estimated at $n = 45$ participants in each group. We performed the power analysis using the PASS 14 Power Analysis and Sample Size Software (2015). NCSS, LLC. Kaysville, Utah, USA, ncss.com/software/pass.

Data analysis

We will perform the primary analysis in per-protocol populations. In the intention-to-treat population and as the sensitivity analysis that we will perform if the proportion of lost-to-follow up will exceed 5% of participants, we will impute the missing data at the 18th month of follow-up with the baseline values that promotes the null hypothesis of no change/no difference. We will check whether the data/participants are missing completely at random using the Little's test. We will check the normality of residuals distributions using the Shapiro Wilk test. Depending on distribution, we will present the data with descriptive statistics: means and standard deviations, or medians and interquartile ranges. In the within groups analysis we will present the results by medians of absolute differences with their 95% confidence intervals, medians of differences relative to the baseline values and their 95% confidence intervals. In the between groups analysis we will present the results by differences between groups' medians, and the differences in the targeted (e.g. FEP) group relative to the median in the control group, both with their 95% confidence intervals as the nonparametric standardized effect size we will calculate Cliff's Delta with its 95% confidence interval, and Hedges g as the parametric effect-size. In the univariable, unadjusted analysis we will calculate the statistical significances of the differences using a two-sided Sign test or the Wilcoxon signed-rank test and in the adjusted, multivariable analysis we will use the quantile (median) regression controlling for variables with the possible confounding effects. In the analysis of primary outcomes in WP2, we will analyze the differences between groups in cortisol response to TSST using the mixed, between-within subject analysis of covariance. BFV was originally measured by TCD in milliseconds. We will transform this to seconds before the analysis. We will assess the differences in CBF in different brain regions measured by SPECT using a χ^2 test. As the standardized effect size we will calculate the Cramer's V . In the analysis of longitudinal data, comparison will be performed by comparing the medians of differences in assessment tests from the baseline to the

18 months of standard treatment (neurocognitive and emotional recognition) we will adjust all analysis of primary outcomes for age, gender and education. We will correct the statistical significances for multiple testing using the sequential Holm-Bonferroni method. The level of statistical significance will be set at a two-tailed $p < 0.05$, and all confidence intervals (CI) at the 95% level. Data analysis will be performed using the statistical programming environment R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

WP 3

After the analysis of in-vivo data, obtained with clinical assessments of psychopathology, neurocognitive tests and functional and structural imaging (SPECT) (data obtained from WP1 and WP2), we will study the alteration in number and distribution, as well as morphology and chemical properties of specific neuron subclasses in histologically prepared human cortical tissue of subjects with schizophrenia in the region that synaptic alteration was suggested by in-vivo testing. These parameters might serve as a normative for diagnosis of disorders characterized with impairment of higher cognitive functions.

WP 4

WP4 incorporates association analysis of pharmacogenetic loci and treatment response in different aspects. The methodology of high-salt extraction and genotyping are established, and previously described (Miller et al. 1988). Genotyping for MTHFR, ZNF804A and CNR1 will be performed by real-time PCR on 7500 Real-Time PCR Systems (Applied Biosystems, Life Technologies, USA) using TaqMan[®] SNP Genotyping Assays for each SNP, as follows: MTHFR C677T (rs1801133) TaqMan[®] SNP Genotyping Assay ID: C_1202883_20; MTHFR A1298G (rs1801131) TaqMan[®] SNP Genotyping Assay ID: C_850486_20; ZNF804A (rs1344706) TaqMan[®] SNP Genotyping Assay ID: C_2834835_10; CNR1 (rs7766029) TaqMan[®] SNP Genotyping Assay ID: C_28979971_20; CNR1 (rs12720071) TaqMan[®] SNP Genotyping Assay ID: C_30749291_10; NDUFV2 (NADH dehydrogenase (ubiquinone) flavoprotein 2) (rs2032161) TaqMan[®] SNP Genotyping Assay ID: C_11901400_10, Genotyping for HSPA1B will be performed by PCR-RFLP on GeneAmp PCR System 9600 (Applied Biosystems, Life Technologies, USA), as described previously (Kowalczyk et al. 2014).

EXPECTED OUTCOMES AND IMPLICATIONS

With WP1 we expect to identify psychopathological and neurocognitive deficits that differentiate healthy con-

trols from patients with FEP, and to identify their changes in relation to clinical and functional outcome over time. With WP2 we want to explore the deficits in psychosis further, by including functional measures in the analysis - we expect to identify specific deficits in brain perfusion in resting state and under activation and response to stress using salivatory cortisol that differentiate healthy controls from patients with FEP, and identify their changes in relation to clinical and functional outcome over time. The identification of a specific set of deficits that can act as state or trait markers can be used in the clinical practice as an additional diagnostic tool in the first episode of psychosis, especially considering the fact that FEP can result from different underlying disorders, and for prediction of treatment outcome (relapse, remission...) which can be then used to help tailoring individualized treatment options. By adding the results from WP3 and WP4 as well, we expect to develop an innovative multidisciplinary approach to identify multisystem biomarkers based on data gathered from clinical, functional, structural and genetic level in vivo as well as on neuronal level in post-mortal tissue. The unique contribution of this research includes possible clinical and scientific implications.

Limitations

First, we selected a consecutive sample of patients with FEP and the convenient sample from the healthy control population. Therefore the study has the increased risk of the sample bias and lower probability of representativeness for the targeted populations. Second, although the naturalistic design may mirror the real-life situation better, the number of confounders may be greater than in a controlled study design, such as the influence of different medications or other treatment options and will be included in the analyses. Third, although all subjects in our study were young, and treated only up to three weeks with antipsychotics, we cannot exclude the effects of medication on some of the studies features (for example BFV).

Acknowledgements:

This research was supported by the grant of the Croatian Science Foundation No UIP-2014-09-1245 Biomarkers in schizophrenia – integration of complementary methods in longitudinal follow up of first episode psychosis patients. The authors declare no further conflicts of interest. The work of Ana Hladnik and Sanja Darmopil was supported by the European Union through the European Regional Development Fund, Operational Programme Competitiveness and Cohesion, grant agreement No. KK.01.1.1.01.0007, CoRE - Neuro. The work of Aleksandar Savic was supported by the grant of the Croatian Science Foundation IP-2014-09-2979 Multimodal approach to treatment and long-term assessment of depressive disorder using magnetic resonance imaging.

Conflict of interest: None to declare.

Contribution of individual authors:

Martina Rojnic Kuzman conceived the main study concept.

Martina Rojnic Kuzman, Porin Makaric, Dina Bosnjak Kuharic, Ivana Kekin & Linda Rossini Gajsak designed the study protocol.

Martina Rojnic Kuzman & Porin Makaric wrote the first draft and the final version of the manuscript.

Zarko Bajic wrote statistical methods.

Dina Bosnjak Kuharic, Ivana Kekin, Linda Rossini Gajsak, Marina Boban, Nada Bozina, Tamara Bozina, Mirela Celic Ruzic, Sanja Darmopil, Igor Filipcic, Lana Ganoci, Ana Hladnik, Zoran Madzarac, Branko Malojcic, Alma Mihaljevic Peles, Daniel J. Mueller, Drazenka Ostojic, Zdravko Petanjek, Ratimir Petrovic, Zeljka Vogrinc, Aleksandar Savic, Ante Silic, Marina Sagud, Maja Zivkovic & Zarko Bajic gave critical comments and revised the drafts of the study.

All authors approved the final version of the manuscript.

References

1. Aaslid R, Markwalder T-M & Normes H: Noninvasive transcranial Doppler ultrasound recording of flow velocity in basal cerebral arteries. *Journal of neurosurgery* 1982; 57:769-74
2. Åberg K, Adkins DE, Bukszár J, Webb BT, Caroff SN, Miller DD et al.: Genomewide association study of movement-related adverse antipsychotic effects. *Biological psychiatry* 2010; 67:279-82
3. Addington D, Addington J & Schissel B: A depression rating scale for schizophrenics. *Schizophrenia research* 1990; 3:247-51
4. Addington D, Addington J & Maticka-Tyndale E: Assessing depression in schizophrenia: the Calgary Depression Scale. *The British Journal of Psychiatry*, 1993
5. American Psychiatric Association: *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013
6. an der Heiden W & Häfner H: The epidemiology of onset and course of schizophrenia. *European archives of psychiatry and clinical neuroscience* 2000; 250:292-303
7. Ashton MC, Lee K & Goldberg LR: The IPIP-HEXACO scales: An alternative, public-domain measure of the personality constructs in the HEXACO model. *Personality and Individual Differences* 2007; 42:1515-26
8. Boban M, Črnac P, Junaković A & Malojčić B: Hemodynamic monitoring of middle cerebral arteries during cognitive tasks performance. *Psychiatry and clinical neurosciences* 2014a; 68:795-803
9. Boban M, Črnac P, Junaković A, Garami Z & Malojčić B: Blood flow velocity changes in anterior cerebral arteries during cognitive tasks performance. *Brain and cognition* 2014b; 84:26-33
10. Buss AH & Perry M: The aggression questionnaire. *Journal of personality and social psychology* 1992; 63:452

11. Chiu SK, Collier CP, Clark AF & Wynn-Edwards KE: Salivary cortisol on ROCHE Elecsys immunoassay system: pilot biological variation studies. *Clinical biochemistry* 2003; 36:211-4
12. Cotton S, Mackinnon A, Gleeson J, Hides L, Herrman H, Filia K et al.: The First-Episode Psychosis Outcome Study: Preliminary Data On the Long-Term Follow-Up of First-Episode Psychosis Patients Who Were Treated At the Early Psychosis Prevention and Intervention Centre Between 1998 and 2000. *Schizophrenia Bulletin* 2017; 43(Suppl 1): S134
13. Dedovic K, Renwick R, Mahani NK, Engert V, Lupien SJ & Pruessner JC: The Montreal Imaging Stress Task: using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain. *Journal of Psychiatry and Neuroscience* 2005; 30:319
14. Dickerson SS & Kemeny ME: Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological bulletin* 2004; 130:355
15. Dressendorfer RA, Kirschbaum C, Rohde W, Stahl F & Strasburger CJ: Synthesis of a cortisol-biotin conjugate and evaluation as a tracer in an immunoassay for salivary cortisol measurement. *J Steroid Biochem Mol Biol* 1992; 43:683-92
16. Dubois B, Slachevsky A, Litvan I & Pillon B: The FAB: a Frontal Assessment Battery at bedside. *Neurology* 2000; 55:1621-6
17. Emde RN, Osofsky JD & Butterfield PM: *The IFEEL pictures: A new instrument for interpreting emotions*. International Universities Press, Inc. 1993
18. Emsley R, Oosthuizen P, Koen L, Niehaus D, & Martinez L: Comparison of treatment response in second-episode versus first-episode schizophrenia. *Journal of clinical psychopharmacology* 2013; 33:80-83
19. Fastenau PS, Denburg NL & Hufford BJ: Adult norms for the Rey-Osterrieth Complex Figure Test and for supplemental recognition and matching trials from the Extended Complex Figure Test. *Clin Neuropsychol* 1999; 13:30-47
20. Folstein MF, Folstein SE & McHugh PR: "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12:189-98
21. Freedman M, Leach L, Kaplan E, Winocur G, Shulman K & Delis DC: *Clock Drawing: A Neuropsychological Analysis*. Oxford University Press, 1994
22. Fusar-Poli P, Smieskova R, Kempton M, Ho B, Andreasen N & Borgwardt S: Progressive brain changes in schizophrenia related to antipsychotic treatment? A meta-analysis of longitudinal MRI studies. *Neuroscience & Biobehavioral Reviews* 2013; 37:1680-91
23. Fusar-Poli P, Papanastasiou E, Stahl D, Rocchetti M, Carpenter W, Shergill S et al.: Treatments of negative symptoms in schizophrenia: meta-analysis of 168 randomized placebo-controlled trials. *Schizophrenia Bulletin* 2014; 41:892-9
24. Glausier JR & Lewis DA: Dendritic spine pathology in schizophrenia. *Neuroscience* 2013; 251: 90-107
25. Golden CJ: Identification of brain disorders by the Stroop Color and Word Test. *J Clin Psychol* 1976; 32:654-8
26. Gur RC, Sara R, Hagendoorn M, Marom O, Hughett P, Macy L et al.: A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of neuroscience methods* 2002; 115:137-43
27. Hanrahan K, McCarthy AM, Kleiber C, Lutgendorf S & Tsalikian E: Strategies for salivary cortisol collection and analysis in research with children. *Applied nursing research: ANR* 2006; 19:95-101
28. Holmes TH & Rahe RH: The social readjustment rating scale. *Journal of psychosomatic research* 1967; 11:213-8
29. Hutt ML: The Kohs block-design tests. A revision for clinical practice. *Journal of Applied Psychology* 1932; 16:298
30. Juraeva D, Haenisch B, Zapatka M, Frank J, Witt SH, Mühleisen TW et al.: Integrated pathway-based approach identifies association between genomic regions at CTCF and CACNB2 and schizophrenia. *PLoS genetics* 2014; 10:e1004345
31. Kapucu OL, Nobili F, Varrone A, Booij J, Vander Borgh T, Nägren K et al.: EANM procedure guideline for brain perfusion SPECT using 99mTc-labelled radiopharmaceuticals, version 2. *Eur J Nucl Med Mol Imaging* 2009; 36:2093-102
32. Kahn RS, Fleischhacker WW, Boter H, Davidson M, Vergouwe Y, Keet IP et al.: Effectiveness of antipsychotic drugs in first-episode schizophrenia and schizophreniform disorder: an open randomised clinical trial. *The Lancet* 2008; 371:1085-97
33. Kay SR, Fiszbein A & Opfer LA: The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia bulletin* 1987; 13:261
34. Kirschbaum C, Pirke K-M & Hellhammer DH: The 'Trier Social Stress Test' – a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology* 1993; 28:76-81
35. Knecht S, Deppe M, Ringelstein E-B, Wirtz M, Lohmann H, Dräger B et al.: Reproducibility of functional transcranial Doppler sonography in determining hemispheric language lateralization. *Stroke* 1998; 29:1155-9
36. Knecht S, Dräger B, Deppe M, Bobe L, Lohmann H, Flöel A et al.: Handedness and hemispheric language dominance in healthy humans. *Brain* 2000; 123:2512-8
37. Kowalczyk M, Owczarek A, Suchanek R, Paul-Samojedny M, Fila-Danilow A, Borkowska P et al.: Heat shock protein 70 gene polymorphisms are associated with paranoid schizophrenia in the Polish population. *Cell Stress Chaperones* 2014; 19:205-15
38. Lepage M, Bodnar M & Bowie CR: Neurocognition: clinical and functional outcomes in schizophrenia. *The Canadian Journal of Psychiatry* 2014; 59:5-12
39. Levine SZ, Rabinowitz J, Faries D, Lawson AH & Ascher-Svanum H: Treatment response trajectories and antipsychotic medications: examination of up to 18 months of treatment in the CATIE chronic schizophrenia trial. *Schizophrenia Research* 2012; 137:141-6
40. Lewis DA: Inhibitory neurons in human cortical circuits: substrate for cognitive dysfunction in schizophrenia. *Current opinion in neurobiology* 2014; 26:22-6
41. Lichtenberger EO & Kaufman AS: *Essentials of WAIS-IV assessment*. John Wiley & Sons, 2009
42. Marušić A, Roskar S & Roškar M: *Questionnaire on suicide ideation and behaviour*. 2007
43. McClay JL, Adkins DE, Åberg K, Stroup S, Perkins DO, Vladimirov VI et al.: Genome-wide pharmacogenomic

- analysis of response to treatment with antipsychotics. *Molecular psychiatry* 2011; 16:76
44. Mesholam-Gately RJ, Giuliano AJ, Goff KP, Faraone SV & Seidman LJ: Neurocognition in first-episode schizophrenia: a meta-analytic review. *Neuropsychology* 2009; 23:315
 45. Miller R, Ream G, McCormack J, Gunduz-Bruce H, Sevy S, & Robinson D: A prospective study of cannabis use as a risk factor for non-adherence and treatment dropout in first-episode schizophrenia. *Schizophrenia research* 2009; 113:138-144
 46. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; 16:1215
 47. Nicodemus KK, Elvevåg B, Foltz PW, Rosenstein M, Diaz-Asper C & Weinberger DR: Category fluency, latent semantic analysis and schizophrenia: a candidate gene approach. *Cortex* 2014; 55:182-91
 48. Norbeck JS: Modification of life event questionnaires for use with female respondents. *Research in Nursing & Health* 1984; 7:61-71
 49. Oldfield RC: The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971; 9:97-113
 50. Parker G, Tupling H & Brown L: A parental bonding instrument. *Psychology and Psychotherapy: Theory, Research and Practice* 1979; 52:1-10
 51. Patton JH & Stanford MS: Factor structure of the Barratt impulsiveness scale. *Journal of clinical psychology* 1995; 51:768-74
 52. Petanjek Z, Judas M, Kostović I, Uylings HB: Lifespan alterations of basal dendritic trees of pyramidal neurons in the human prefrontal cortex: a layer-specific pattern. *Cereb Cortex* 2008; 18:915-29
 53. Petanjek Z, Berger B, Esclapez M: Origins of cortical GABAergic neurons in the cynomolgus monkey. *Cereb Cortex* 2009; 19:249-62
 54. Radua J, Borgwardt S, Crescini A, Mataix-Cols D, Meyer-Lindenberg A, McGuire P et al.: Multimodal meta-analysis of structural and functional brain changes in first episode psychosis and the effects of antipsychotic medication. *Neuroscience & Biobehavioral Reviews* 2012; 36:2325-33
 55. Riecher-Rössler A, Rybakowski J, Pflueger M, Beyrau R, Kahn R, Malik P et al.: Hyperprolactinemia in anti-psychotic-naive patients with first-episode psychosis. *Psychological medicine* 2013; 43:2571-82
 56. Ripke S, Neale BM, Corvin A, Walters JT, Farh K-H, Holmans PA et al.: Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; 511:421
 57. Rosen C, Marvin R, Reilly JL, Deleon O, Harris MS, Keedy SK, Solari H, Weiden P, Sweeney JA: Phenomenology of first-episode psychosis in schizophrenia, bipolar disorder, and unipolar depression: a comparative analysis. *Clin Schizophr Relat Psychoses* 2012; 6:145-51
 58. Rosenberg MSatas-i & Press. U. Rosenberg Self-Esteem Scale 1965
 59. Rush AJ, Gullion CM, Basco MR, Jarrett RB & Trivedi MH: The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychological medicine* 1996; 26:477-86
 60. Schmidt M: Rey auditory verbal learning test: A handbook. Western Psychological Services Los Angeles, CA. 1996
 61. Sedmak D, Hrvoj-Mihić B, Džaja D, Habek N, Uylings HB & Petanjek Z: Biphasic dendritic growth of dorsolateral prefrontal cortex associative neurons and early cognitive development. *Croat Med J* 2018; 59:189-202
 62. Slotema CW, Blom JD, Niemantsverdriet M, Deen M & Sommer IE: Comorbid Diagnosis of Psychotic Disorders in Borderline Personality Disorder: Prevalence and Influence on Outcome. *Frontiers in psychiatry* 2018; 9:84
 63. Szirmai I, Amrein I, Pálvölgyi L, Debreczeni R & Kamondi A: Correlation between blood flow velocity in the middle cerebral artery and EEG during cognitive effort. *Cognitive Brain Research* 2005; 24:33-40
 64. Tessier A, Boyer L, Husky M, Baylé F, Llorca PM & Misdrachi D: Medication adherence in schizophrenia: The role of insight, therapeutic alliance and perceived trauma associated with psychiatric care. *Psychiatry research* 2017; 257:315-321
 65. Tombaugh TN: Trail Making Test A and B: normative data stratified by age and education. *Arch Clin Neuropsychol* 2004; 19:203-14
 66. von Reutern G-M, Kaps M & Büdingen HJ: Ultraschall-diagnostik der hirnersorgenden Arterien: Dopplersonographie und Farbduplexsonographie der extra-und intrakraniellen Arterien; 39 Tabellen. Thieme, 2000
 67. Wechsler D: Wechsler memory scale. 1945
 68. Weiden PJ, Mackell JA & McDonnell DD: Obesity as a risk factor for antipsychotic noncompliance. *Schizophr Res* 2004; 66:51-7
 69. Whoqol Group: Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological medicine* 1998; 28:551-8
 70. World Health Organization: The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. World Health Organization. 1992
 71. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2013; 310:2191-2194
 72. Young RC, Biggs JT, Ziegler VE & Meyer DA: A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry*. 1978; 133:429-35

Correspondence:

Martina Rojnic Kuzman, MD, PhD
 Department of Psychiatry, Zagreb University Hospital Centre
 Kispaticeva 12, 10 000 Zagreb, Croatia
 E-mail: mrojnic@gmail.com