Project Periodic Report

Grant Agreement number:	**602152**
Project acronym:	**PRONIA**
Project title:	**Personalised Prognostic Tools for Early Psychosis Management**
Funding Scheme:	**Collaborative Project**
Date of latest version of Annex I against which the assessment will be made:	**07-05-2018**
Periodic report:	**1st ☐ 2nd ☐ 3rd ☒ 4th ☐**
Period covered:	**From 01-10-2016 to 31-03-2018**

Name, title and organisation of the scientific representative of the project's coordinator:

Prof. Dr. Nikolaos Koutsouleris – LUDWIG-MAXIMILIANS-UNIVERSITAET MUENCHEN (LMU)

Department of Psychiatry and Psychotherapy

Nußbaumstr. 7

80336 Munich, Germany

Phone: +49 89 4400 55885

Fax: +49 89 4400 54749

E-mail: Nikolaos.Koutsouleris@med.uni-muenchen.de

Project website address: www.pronia.eu

Table of Content

Declaration by the scientific representative of the project coordinator

I, as scientific representative of the coordinator of this project and in line with the obligations as stated in Article II.2.3 of the Grant Agreement declare that:

The attached periodic report represents an accurate description of the work carried out in this project for this reporting period;

The project (tick as appropriate)[1]:

- ☐ has fully achieved its objectives and technical goals for the period;
- ☒ has achieved most of its objectives and technical goals for the period with relatively minor deviations;
- ☐ has failed to achieve critical objectives and/or is not at all on schedule.

The public website, if applicable

- ☒ is up to date
- ☐ is not up to date

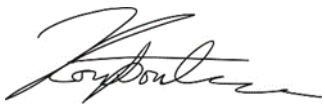
To my best knowledge, the financial statements which are being submitted as part of this report are in line with the actual work carried out and are consistent with the report on the resources used for the project (section 4) and if applicable with the certificate on financial statement.

All beneficiaries, in particular non-profit public bodies, secondary and higher education establishments, research organisations and SMEs, have declared to have verified their legal status. Any changes have been reported under section 2.4 (Project Management) in accordance with Article II.3.f of the Grant Agreement.

Name of scientific representative of the Coordinator: Prof. Dr. med. Nikolaos Koutsouleris

Date: 12/06/2018

Signature of scientific representative of the Coordinator:



[1] If either of these boxes below is ticked, the report should reflect these and any remedial actions taken.

Section 1 - Publishable summary

PRONIA

Logo:

Project title: Personalised Prognostic Tools for Early Psychosis Management

Website: www.pronia.eu

Contractors involved (PRONIA consortium):Prof. Dr. Nikolaos Koutsouleris – LUDWIG-MAXIMILIANS-UNIVERSITAET MUENCHEN (LMU)

Department of Psychiatry and Psychotherapy

Nußbaumstr. 7

80336 Munich, Germany

Phone: +49 89 4400 55885

Fax: +49 89 4400 54749

Other partners and team leaders:

02 UNIVERSITAET BASEL (UNIBAS/UPK) – Stefan Borgwardt

03 KLINIKUM DER UNIVERSITAET ZU KOELN (UKK) – Marlene Rosen

04 THE UNIVERSITY OF BIRMINGHAM (UoB) – Rachel Upthegrove

05 TURUN YLIOPISTO (UTU) – Raimo Salokangas

06 UNIVERSITA DEGLI STUDI DI UDINE (Uni Udine) – Franco Fabbro

07 UNIVERSITY OF MELBOURNE (UoM) – Christos Pantelis

08 DYNAMIC EVOLUTION GMBH (DynEV) – Michael Jovicevic

09 GABO:MI GESELLSCHAFT FUR ABLAUFORGANISATION:MILLIARIUM MBH & CO KG (GABO:mi) – Birgit Fuchs - TERMINATED

10 General Electric Deutschland Holding GmbH (GE GRC) – Dirk Bequé

11 GE HEALTHCARE GMBH (GE HC) – Timo Schirmer

12 UNIVERSITA DEGLI STUDI DI MILAN (UMIL) – Paolo Brambilla

13 ARTTIC (ART) – Martin Dietz

14 WESTFAELISCHE WILHELMS-UNIVERSITAET MUENSTER (WWU) - Rebekka Lencer (see Amendment 3)

15 UNIVERSITA DEGLI STUDI DI BARI ALDO MORO (UNIBA) - Alessandro Bertolino (see Amendment 3)

16 HEINRICH-HEINE-UNIVERSITAET DUESSELDORF (UDUS) - Eva Meisenzahl (see Amendment 3)

17 UNIVERSITAETSKLINIKUM BONN (ubk) - Markus Nöthen (see Amendment 3)

18 MAX-PLANCK-GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV (MPG) - Matthias Mann (see Amendment 3)

1.1 Summary description of project context and objectives

Reliable and accessible tools for an individualised early recognition of psychoses are increasingly at the centre stage of international mental healthcare policy and research. This is due to the fact that affective and non-affective psychoses typically commence in the most productive and critical period of life – late adolescence and early adulthood – frequently entailing long-term disability and increased mortality in the affected patients. These factors drive the 6.3% of the global

burden of disease caused by these disorders and account for the €207 billion per year in Europe alone spent due to direct and indirect costs. As such, psychoses rank among the top most expensive brain-related disorders, with a similar health-economic impact as cardiovascular diseases.

A significant component in the reduction of this burden is to provide preventive treatments to those at highest risk of developing these devastating illnesses in the future. However, these targeted interventions will only become regularly and widely available if quantifiable biological markers of the emerging disease and its impact on a patient's long-term functional outcome will become part of the normal diagnostic workflows in clinical psychiatry. The characterisation of these markers could also have a significant impact on the development of novel preventive interventions by e.g. providing new biological surrogate markers and quantification methods to the drug discovery pipelines of the pharmaceutical industry.

Objectives: Therefore, the multi-centre collaborative project PRONIA ("Personalised Prognostic Tools for Early Psychosis Management") has been awarded with €6,000,000 FP7 grant of the European Commission to first (i) explore the utility of routine brain imaging and complementary data in predicting different mental health-related outcomes in persons with at-risk and early stages of psychoses, and (ii) evaluate which predictive signatures in these data generalise well across different mental health services and pathways to care. Secondly, based on the knowledge gained through this biomarker validation process, PRONIA will implement new multi-modal risk quantification tools and embed them into an eHealth prototype providing telepsychiatric services along with biomarker-based risk stratification tools to accurately predict mental health-related disability in young help-seeking persons. More specifically, PRONIA will:

1) Optimise the consortium's candidate imaging markers for a **clinically reliable prediction and staging of psychoses** by augmenting them with complementary patient data and generalising them across mental health services based on cross-centre, multi-modal pattern recognition (COMPARE).
2) **Develop & validate new surrogate markers** for an individualised risk quantification by analysing brain imaging and complementary data with COMPARE to:
   a) **Predict poor functional outcomes such as social and vocational exclusion** in patients with recent-onset psychosis and depression as well as help-seekers with different at-risk mental states for psychosis,
   b) **Model the impact of concomitant psychiatric conditions** on predictions in objectives 1 and 2a that relate e.g. to the presence of depression, risk-conferring personality traits and substance abuse,
   c) **Monitor disease progression and remission** across the at-risk and early stages of psychoses to dynamically refine predictions in 1, 2a & 2b by combining serial MRI scanning, neuropsychological, psychometric and metabolic assessments.
3) **Disseminate & commercially exploit** these surrogate markers by delivering **cybernetic prognostic services** to health services, research institutions and the biopharmaceutical industry through a **telemedicine-based** European company.

The third objective expresses PRONIA's goal to realise licensing, commercialisation, and sustained engineering strategies of these biomarker-based early recognition services through broadly available telemedicine applications. This will provide psychosis risk profiling tools to diverse target groups in the healthcare markets, including care-givers, the pharmaceutical industry and research institutions. By disseminating objective risk quantification, PRONIA's products will provide firm diagnostic grounds for preventive therapy, improving outcomes and reducing costs. Thus, they will offer a unique selling proposition to the mental health sectors in Europe and beyond.

1.2 Work performed since the beginning of the project and the main results achieved so far

WP1 started developing standardised clinical examination workflows in 07/2013. Until 12/2013 these workflows were fully digitalised in the PRONIA portal allowing for centralised data acquisition and management (see WP9). Mainly due to the local heterogeneity of ethical and data security approval processes, recruitment start was deferred to a locally varying degree, ranging from the 15/02/2014 (LMU) to the 29/07/2014 (University of Birmingham). In addition, three further psychiatric university centres joined the PRONIA consortium as beneficiaries and recruitment centres (Münster, Bari and Düsseldorf; recruitment started between 01/10/2016 and 01/08/2017). Up to the end of the third reporting period, the whole consortium screened 6829 participants. This resulted in 1809 participants included in the study with an overall attrition rate of 19,4% because of non-participation, drop-outs or exclusions. 1485 of the 1809 individuals (82%) who consented to participate completed the baseline examination (T0), 783 (43.3%) the 9-month T1 visit and 439 (24.3%) the final 18-month visit T2. Importantly, in the *2nd funding period*, the original QC process consisting of

monthly telephone case conferences, quarterly FAQ conferences, and biannual inter-rater reliability testing, was significantly extended by implementing, testing, and applying a multi-stage semi-automated data quality ascertainment and correction procedure. This procedure regularly extracts all clinical and neurocognitive (see WP4) data from the portal, checks these data automatically against a canon of programmed data quality and completeness rules and returns a feedback spreadsheet to each PRONIA rater team. Based on this spreadsheet, the respective team members correct / complete their data entry and return the commented spreadsheet to UKK where their annotations are fed back into the database QC system. Thus, in an iterative process all phenotypic data are cleaned and prepared for the first major analysis phase to start in the 3rd funding period. In the *3rd funding period*, WP1 maintained quality control and assurance of assessments and related procedures (central web conferences and local face-to-face trainings, monthly case conferences with an independent expert, FAQ database, on-demand consultancy service, folder for introduction of new raters and different inter-rater reliability checks revealing very satisfying results). Computer-aided data quality control was optimised, and data-cleaning procedure consolidated facilitating the release of a first data set for analyses (baseline data of the discovery sample (all cases recruited until May 1st 2016). For follow-up data (up to T1) of the discovery sample, quality check will be completed by 15th of May 2018.Furthermore, by deliverable 01.04 a first important step could be done towards generating prognostic models that encompass a concise set of clinical variables which conjointly maximise prediction accuracy and cross-centre generalisation capacity, while minimising the number of variables needed for prediction and hence the overall duration and burden of clinical assessment.

WP2 together with WP11 implemented a certification framework for the fully modular and object-based re-engineering and extension of the project's machine learning platform NeuroMiner (NM). Conceptualisation of NeuroMiner2 (NM2) and its interfaces started in October 2013 paralleled by an in-depth documentation of the program's functions and capabilities. Subsequently, the programming of NM2 modules commenced in 01/2014. Given the rapidly broadening accessibility of machine learning algorithms in the field, we decided at the beginning of *the 2nd funding period* to deviate from the original NM2 development plan and instead merged the new NM2 functions with our original NM framework. Using this enhanced and more user-friendly program, we conducted several pilot analyses on the growing PRONIA database and summarize related findings in the WP2 report. We are currently bug-fixing the application and improving its broad usability further. In the beginning of the *3rd funding period*, WP2 released a beta version of NM as open source easy-to-use predictive analytics framework for clinically-oriented neuroscientists. Based on the feedback of the 2nd PRONIA summer school which was organized by WP2 (June 2017), numerous improvements concerning the usability of the toolbox as well as new functionalities were added to the application. A final release of the NM1.0 is expected for end of June 2018. In addition to software development, WP2 heavily focused on producing the first scientific publications PRONIA by testing the feasibility and generalizability of (1) individualized functional outcome prediction in CHR and ROD patients, and (2) individualized prediction of disease transition in the CHR for psychosis. The former analysis has been submitted to a high-impact journal and is currently under revision. In the final funding period, WP2 will take up certification efforts again in close collaboration with WP11 in order to prepare for the commercialisation phase of the consortium's prognostic tools.

During the *1st funding period*, **WP3** defined a harmonised MRI framework for standardised cross-site management of MRI protocols and data. A software tool (PI2T) was developed for anonymization and conversion of MRI images and for uploading the MRI data to the PRONIA database server. WP3 also conducted an MRI calibration study with 10 healthy volunteers along with an MRI phantom at each of the involved scanners by using the established MRI acquisition protocol (see D3.1). *In the 2nd funding period*, WP3 established MRI pre-processing and quality control (QC) algorithms, which have been merged into a unified WP3 processing system. More specifically, we integrated all pipelines for structural, functional and diffusion tensor image processing into a single workbench. The server hosting the WP3 processing master has been set up and tested. Starting in December 2016 we have regularly applied these automatic procedures to our phantom and human MRI measurements to monitor the quality of the MRI data and generate multi-modal brain descriptors for the analysis phase in the 3rd funding period (see D3.2). Three phantom QC reports have been circulated within the consortium and potential corrective actions have been initiated. *In the 3rd funding period*, the WP3 team applied the well-established and modality-specific pre-processing and QC pipelines to all available neuroimaging data. All T0 data has been pre-processed and controlled for quality with this pipeline and processing of T1 images is already ongoing. Preliminary modality-specific MRI predictors including QC measures have been released to the consortium and analyses are currently ongoing with data from the discovery sample (N = 768). The aim of these analyses is to test whether MRI markers (see D3.3 for details) can predict the temporal course of clinical variables such as functioning or depressive symptoms. A first paper has already been submitted and we hope to complete the other analyses by the end of this year. Novel and more sophisticated pre-processing pipelines are currently under construction and soon ready to be applied.

In the 1st funding period **WP4** developed a computerized neurocognitive test battery, PCB for use in the PRONIA study. *In the 2nd funding period*, WP4 further improved PCB and developed together with LMU an automated QC procedure for the regular QC of PRONIA's neurocognitive datasets. Furthermore, pilot analyses were conducted to assess baseline differences between the 4 study groups covering a host of neurocognitive domains. These preliminary findings mainly showed that persons with clinical high-risk state for psychosis – irrespective of the subsequent outcome – rank between patients with first-episode psychosis and patients with first-episode depression in terms of verbal memory, cognitive flexibility, social cognitive and visual memory performance. In the *3rd funding period*, WP4 carried out the normative study on the neurocognitive data collected within PRONIA on the healthy controls group and obtained the norms for the tests composing the PCB. Additionally, machine learning regression and classification analyses were conducted by mean of NeuroMiner with the aim of deriving the neurocognitive predictors of poor functional and social outcomes that may accompany the risk of psychosis. Particularly, a first set of regression analyses explored the relation between target neurocognitive features and the participants' everyday functioning, independently by the presence or absence of any psychiatric condition. Secondly, classification analyses were carried out to build reliable models for classifying patients with clinical high-risk state for psychosis with respect to those with first-episode depression. While the regressions models, with the cognitive scores as predicting features, could reveal significant correlations between observed and predicted functioning scores, unexpectedly the classification analyses did not allow a reliable separation between patients with clinical high-risk state for psychosis or with first-episode depression. An additional study explored the relation between schizotypy traits and neurocognitive skills across patients at high-risk state for psychosis, with first-episode psychosis, first-episode depression and healthy controls. This study pointed out different patterns of schizotypy traits across the three patients' groups that were coherent with their psychopathological status, as well as different associations between clinical group, schizotypy traits and cognitive measures. More study proposals concerning the environmental risk factors (Discrimination, Immigration, Quality of Life, effect of alcohol, cannabis, and tobacco), psychopharmacological treatment, gender differences, structural and functional brain connectivity and neurocognitive measures, at baseline and longitudinally, across the four clinical groups, were approved by the PRONIA Steering Committee and will be carried out by the WP4 in the upcoming period.

In the *1st funding period*, **WP5** implemented the blood sampling SOPs, including (1) the provision of standardised blood collection kits, (2) local workflows, (3) shipment and storage at the Helmholtz centre. Based on this blood sampling infrastructure, wereceived and processed samples from 893 study participants until the end of the *2nd funding period*. As soon as the quality control procedure of the discovery sample is finalized (December 2016), WP05 will start with genotyping the processed blood samples using the PsychChip Illumina technology. In the *3rd funding period*, 1,700 samples will be genome-wide genotyped using Illumina's Infinium Global Screening (GSA) Array-24 BeadChip version 2 + Psych content (GSA). The GSA includes > 650,000 markers and offers an unparalleled genomic coverage and imputation performance. The Psych content comprises 50,000 variants associated with common psychiatric disorders such as schizophrenia, bipolar disorder, and autism spectrum disorders. After an extensive quality control, a genome-wide association study (GWAS) will be performed. Furthermore, polygenic risk scores (PRS) for each individual will be calculated. The PRS will be calculated using data from the worldwide largest genome-wide association studies for schizophrenia, bipolar disorder, and depression performed within the Psychiatric Genomics Consortium. These data will then be transferred to the PRONIA portal, thus making them available to the consortium for multivariate pattern analyses.

In the *3rd funding period*, **WP6** has been exploring structured data fusion through the use of data tensors. Tensor-based analysis and decomposition is advantageous in the context of data completion (missing data), data compression and the fusion of information from different modalities. Recently, WP6 has been developing a test-bed using tensor decomposition and fusion with data arising from image-based modalities in particular – although we do plan to extend analysis to neuro-cognitive information also.

During the *2nd funding period*, **WP7** tested semi-supervised pattern analysis pipelines in a preliminary database of 365 subjects examined at LMU and UKK. This cross-sectional database contained phenotypic baseline data on the severity of prodromal symptoms (SIPS) and information on the premorbid adjustment as well as resting-state (rs)-fMRI images, which were processed with the rs-fMRI connectivity pipeline of WP3. The pilot analyses of WP7 showed that subgrouping patients based on gender, age, and body-mass-index prior to using supervised machine learning methods significantly increased the single-subject diagnostic separability of our study cohorts. In the *3rd funding period*, WP7 investigated subgroups derived from the SIPS psychosis symptom measure, created a toolbox to subgroup whole-brain sMRI images, investigated sMRI site effects, developed a hybrid semi-supervised learning pipeline in collaboration with WP2, and tested sensitisation effects related to subgroup, diagnostic, and prognostic semi-supervised classifications. Results demonstrated diagnostic specificity of SIPS subgroups, proof-of-concept evidence for whole-brain sMRI subgrouping, effective elimination of site effects from sMRI images for the purposes of

subgrouping, sensitisation of analyses using age and sex, and preliminary evidence of semi-supervised sensitisation using sMRI subgroups. The proof-of-concept analysis, including the creation of the sMRI clustering toolbox, was published.

In the 2nd funding period, **WP8** tested different strategies to correct for centre effects in the rs-fMRI data. First, different geometric distortion correction methods employing the b0 field map scans were compared with respect to their effects on functional brain connectivity matrices. We are currently investigating which of these correction methods optimizes classification accuracy across different supervised machine learning task. Second, we analysed the data from the PRONIA calibration study using Generalization Theory[2] which generated image maps with most reproducible structural and functional brain descriptors across sites. We used these maps to threshold the subjects' data prior to training machine learning models and found that limiting the brain descriptors to the set most reproducible set across sites boosted the multi-site generalizability of our classification models (see also WP2). In the *3rd funding period*, WP8 explored the use of one-class machine learning methods as algorithmic basis for outlier detection systems. More specifically, one-class support vector machines were trained and validated on the quality control indices produced in WP3. The algorithms showed good capacity in detecting MRI artifacts when they were trained in a scanner-specific manner. In the final funding period these methods will be integrated into the analytical pipelines of the consortium to explore how outliers affect predictions and hoe outlier information can be utilized to develop more robust prognostic tools.

In the 2nd funding period, **WP9** finished the web-based PRONIA@home interface which enables the telepsychiatric evaluation of help-seeking individuals and made it available to recruitment efforts of the consortium. Furthermore, the PRONIA App was finalized which enables the on-the-fly digitalization and centralization of phenotypic data recorded during rating of our study participants. Finally, WP9 set up a plan on how to integrate the consortium's prognostic NeuroMiner modules into the PRONIA portal infrastructure, thus preparing the grounds for the implementation of a web-based prototype for PRONIA's prognostic tools. In the *3rd funding period*, WP9 built a preliminary Multi-Agent System (MAS) program that automatically applies surrogate makers generated by NeuroMiner to compute individualized outcome estimates for given patient. The MAS fully integrates web-based assessment tools developed during the 2nd funding period and provides data quality assurance and prediction result reports upon request from the case manager. The program is ready for implementation in the PRONIA portal web-based environment.

In the 1st funding period, **WP10** developed a new strategy for the external validation of PRONIA's prognostic prototype after the projected funding of the Melbourne site was cut by 60%, by implementing (1) a synergistic external validation study for PRONIA and PSYSCAN in Melbourne, and (2) initiating a broader international collaboration (HARMONY) with the North American Prodromal Longitudinal Study (NAPLS). *In the 2nd funding period*, the joint NIMH proposal was awarded to the four consortia (HARMONY grant) and PRONIA received a subcontract to set up a database infrastructure for fully anonymised cloud-based data sharing with the other NIMH-HARMONY partners. In the *3rd funding period*, the HARMONY collaboration proceeded successfully with several external validation and joint analysis project being pursued by the consortium partners. Furthermore, despite the reduced funding available to the Melbourne site, the PRONIA-PSYSCAN recruitment infrastructure became operational in the beginning of 2018 and study enrolment has developed positively in the recent months.

In the 1st funding period, **WP11** established strategies to monitor the consortium's freedom-to-operate (f-t-o), to manage IP rights within PRONIA and with respect to patent authorities, and to develop a commercialisation roadmap for the project. A first patent application (PCT/EP2014/002154) was filed on 5 August 2014 and is presently pending. Furthermore, WP11 provided support to WP2 & 3 to guide software development toward certification in the sense of the Medical Device Directive 93/43/EC. In the *2nd funding period*, the filed patent was revised and entered the nationalisation phase in the US. We are currently discussing the optimal way for achieving certification status of our prognostic tools given the limited personal resources available for this task. In the 3rd funding period, efforts toward a business plan for a PRONIA spinoff have gained traction based on the positive results of the first predictive analyses

[2] Generalizability Theory: Overview. NOREEN M. WEBB AND RICHARD J. SHAVELSON. Volume 2, pp. 717 – 719. In Encyclopedia of Statistics in Behavioral Science. ISBN-13: 978-0-470-86080-9. Editors Brian S. Everitt & David C. Howell

conducted by WP2. A business consultant was hired to develop a commercialisation strategy together with the Coordinator and the WP11 leader. This strategy will be discussed and finalized during a commercialisation workshop in October 2018. Furthermore, given the availability of a first set of predictive tools and a prototype of the MAS, we will now re-activate the certification work in order to be ready for commercialisation at the end of the next funding period.

In the *2nd reporting period*, **WP12** prepared and submitted the 2nd Amendment to the PRONIA Grant Agreement (currently under evaluation by the European Commission). Project meetings and regular phone conferences were carried out to discuss actual progress; :milliarium, a web-based project management and communication tool, serves as a central platform for important documents and the documentation of the work progress. The project website was continuously maintained and updated and the 1st PRONIA Summer School was organized for internal and external young researchers. In the 3rd *reporting period*, **WP12** prepared and submitted the 3rd Amendment to the PRONIA Grant Agreement (currently under evaluation by the European Commission). Project meetings and regular phone conferences were carried out to discuss actual progress; :milliarium, a web-based project management and communication tool, serving as a central platform for important documents and the documentation of the work progress, was kept up to date. The project website was continuously maintained and updated and the 1st PRONIA Symposium and the 2nd PRONIA Summer School was organized for internal and external young researchers.

1.3 The expected final results and their potential impact and use (including the socio-economic impact and the wider societal implications of the project so far)

Accumulating evidence emphasises the potential large-scale benefits of early intervention in the high-risk states of psychoses. However, the development and provision of early intervention to vulnerable populations critically depends on prognostic instruments that enable healthcare professionals to reliably identify persons at risk of chronic disability according to their objective risk profiles. PRONIA will develop, validate and disseminate such a prognostic system on the basis of neuroimaging and complementary data to **(i) individually forecast severe mental disorders** prior to disease onset, **(ii) individually predict chronic disability** across early disease stages, and **(iii) provide the evidence-based diagnostic grounds** for the individualised selection and adjustment of preventive treatment strategies across heterogeneous healthcare settings. In the 1st project period, PRONIA laid the grounds to provide these results by already recruiting and following a substantial number of study participants within the state-of-the-art digitalised and secure data acquisition infrastructure of its SME partners.

In the 2nd funding period, PRONIA largely expanded its multi-modal database and decided to split it into a discovery sample consisting of participants recruited until the 1st of May 2016 and a validation cohort recruited after this date, which will be extended through collaborations within the HARMONY project and other associated partners. Furthermore, the consortium implemented a set of centralised semi-automated quality control workflows, which will make sure that the findings generated by PRONIA will be based on high-quality data. Importantly, as foreseen by the DoW, we started with pilot machine learning analyses using phenotypic information as well as structural and functional MRI data, thus probing the diagnostic separability of our study groups and the predictability of social functioning in the CHR cohort. We evaluated the impact of site-related heterogeneity including scanner and population effects and devised strategies that demonstrated significant potential to overcome these effects and generate reproducible multi-site predictors. These strategies were implemented at different stages of our data processing streams, ranging from the pre-processing of brain-related information, through the selection of highly reproducible predictive features to the attenuation of site-related effects in the machine learning procedures. We could show that a major objective of the project, the single-subject prediction of functional outcomes in persons with a clinical high-risk state for psychosis, can be achieved within the 3rd funding period. Hence, the project has made significant steps toward implementing the envisaged **bioinformatics processing and prediction framework** which will enable us in the next funding period to develop and validate generalizable biomarkers and predictive models for an improved personalized management of the early stages of affective and non-affective psychoses. In addition, we will make significant parts of this framework available as the open source platform NeuroMiner to foster predictive analytics in the field of clinical neurosciences.

In the 3rd funding period, we delivered on these aims, by finalizing a first full-fledged analysis concerning the individualized prediction of social and role functioning using structural MRI and clinical data. This analysis also provided the first empirical report suggesting considerable increases in prognostic certainty brought about by sequentially combining different data domains within a prognostic algorithm. The respective manuscript has been submitted to a key scientific journal and a revision is currently considered for publication. This paper illustrates that the project has achieved considerable infrastructural progress across multiple domains, consisting of (1) the finalization and continuous usage of a semi-automated quality ascertainment system which guarantee the highest levels of clinical data quality and integrity, (2) the extension and refinement of the image processing routines hosted by the consortium's image processing server (WP3), (3) the establishment of a normative neurocognitive database to be used by researchers analysing neuropsychological information (WP4), (4) the implementation of powerful outlier detection systems for robust prediction, and (5) the development of a prototypic prognostic tool which can be access via the internet (WP9). These achievements have allowed us to make the PRONIA discovery database available for a large array of analyses. These analyses are currently performed by different cross-WP task forces in the consortium, which involve the study of (1) environmental mediators of risk for poor clinical and functional outcomes, (2) personality and schizotypal traits as modulators of risk, and (3) strategies for efficiently combining heterogeneous high-dimensional data within machine learning pipelines, as well as (4) strategies for subtyping patients and factorizing bio-behavioural dimensions of mental illness, with the aim to inform the generation of more powerful predictive systems. In the final funding period, PRONIA will leverage these parallel analyses streams, evaluate the addition genetic and proteomic information (WP5) and integrate the most promising predictive algorithms satisfying the project's objectives into the prognostic prototype developed by WP9. Finally, these algorithms will be prioritized for certification and a clinical impact analysis which will be part of the commercialisation efforts of the consortium (WP11)