In the last 2 decades, several neuroimaging studies investigated brain abnormalities associated with the early stages of psychosis in the hope that these could aid the prediction of onset and clinical outcome. Despite advancements in the field, neuroimaging has yet to deliver. This is in part explained by the use of univariate analytical techniques, small samples and lack of statistical power, lack of external validation of potential biomarkers, and lack of integration of nonimaging measures (eg, genetic, clinical, cognitive data). PSYSCAN is an international, longitudinal, multicenter study on the early stages of psychosis which uses machine learning techniques to analyze imaging, clinical, cognitive, and biological data with the aim of facilitating the prediction of psychosis onset and outcome. In this article, we provide an overview of the PSYSCAN protocol and we discuss benefits and methodological challenges of large multicenter studies that employ neuroimaging measures.

Keywords: psychosis/first episode of psychosis/clinical high risk of psychosis/PSYSCAN/neuroimaging/MRI/machine learning/prediction
Introduction

Neuroimaging provides a powerful, noninvasive method to unveil some of the neurobiological mechanisms underlying serious psychiatric disorders such as schizophrenia. In the last 2 decades, researchers have tried to identify brain abnormalities that could aid the prediction of psychosis onset and clinical outcomes in the early stages of psychosis, so that patients can be offered different forms of treatment according to their individual needs. For example, despite several advancements, one of the key challenges in the management of individuals at clinical high risk of psychosis (CHR-P) is that it is not currently possible to identify the subgroups that will subsequently transition to psychosis or that will develop other mental health disorders. Stratification of these subgroups would allow potentially preventative interventions to be selectively offered to these individuals. This is an important task given that one-size-fits-all therapeutic approaches are not particularly effective to prevent the onset of psychosis in this population. Similar limitations are observed in the clinical management of patients who have already experienced a first episode of psychosis (FEP): it is difficult to reliably predict if and when patients will suffer a relapse solely on clinical presentation. If course of illness could be determined early on, targeted intervention could potentially prevent future hospitalizations. Another important challenge in the management and treatment of people with psychosis is that available antipsychotic medications are partially or not effective in about one-third of patients. It is currently not possible to predict which patients will go on to show a poor response to treatment. Identifying these patients would promote earlier access to effective medications for treatment resistant psychosis, such as clozapine, as well as using psychotherapy more effectively. Although some neuroimaging studies have reported findings that may differentiate patients with distinct clinical outcomes, these have yet to be externally validated and translated into tools that can be used in clinical practice.

In this context, clinically valid and reliable neuroimaging biomarkers of psychosis onset and outcome have yet to be identified, as the results from structural, functional, and neurochemical magnetic resonance imaging (MRI) studies are mixed. This might in part be explained by (1) differences in the sociodemographic (eg, age, ethnicity, migration, socioeconomic status) and clinical features of the samples studied, which often reflect inter-site differences in catchment area populations, type of early detection and intervention services involved, and potential sampling biases (particularly in CHR-P individuals); (2) the use of relatively small samples, which might result in type I and type II errors and therefore limited generalizability of the findings; (3) heterogeneous image preprocessing protocols; and lastly (4) the use of different inclusion criteria in different studies, leading to substantial heterogeneity in symptom severity and comorbidities.

To date, the majority of neuroimaging studies investigating abnormalities in the CHR-P and FEP populations have employed univariate analytical methods that allow statistical inferences at the group, rather than the individual level. Although univariate approaches are suited to detect focal abnormalities at a group level, differences in brain anatomy and functioning in the early stages of psychosis appear to be relatively subtle and widespread. Univariate approaches also involve multiple testing and the subsequent correction for multiple comparisons, which may be too conservative and not sensitive enough to detect alterations that are expressed at a network level rather than in a few distinct brain areas. The application of multivariate data-driven approaches, such as machine learning, allows inferences to be made at the individual level, and therefore carry greater translational potential for application in clinical practice.

In addition, multivariate approaches consider multiple voxels simultaneously and between-voxel correlations, rather than each voxel independently, and might therefore be better suited to detecting abnormalities at a network level, rather than focally.

Machine learning has been employed in the field of mental health to make predictions on a number of neurological and psychiatric conditions, including Alzheimer’s disease, depression, anxiety disorders, eating disorders, and psychosis. For critical reviews, please see Orrù et al, Gifford et al, Vieira et al, Arbabshirani et al and Dwyer et al. In the context of psychosis, machine learning has been used to investigate different stages of illness ranging from psychosis risk, first episode of psychosis to established schizophrenia. Studies have employed structural and functional MRI data but also nonimaging data to make predictions on broadly 3 areas: diagnosis, prognosis, and response to treatment. Aiding diagnosis classification is clinically helpful for some psychiatric conditions, such as anxiety disorders, or prodromal stages, where there is diagnostic uncertainty, while for other conditions, such as established schizophrenia, prediction of prognosis or response to treatment might be clinically more meaningful.

Studies on the early stages of psychosis, including both FEP and CHR-P individuals have generally shown accuracies above 75%. However, a recent study challenged the potential of machine learning for detecting changes in the early stages of psychosis. Using relatively large datasets of FEP patients, Vieira et al reported lower classification accuracies than previous studies (between 50% and 70%), but also poor generalizability of models to other sites. While the initial machine learning studies suggest that this approach holds some promise, they have involved relatively small groups of patients.
This is important as the reliability of machine learning is directly affected by sample size, and overall accuracies are seen to decrease with sample size. This suggests that results from previous studies, including studies on patients with psychosis, may not be generalizable, and so must be interpreted with caution. This also suggests that overfitting may be taking place; which is when a machine learning model is fitted to noise in the data rather than to an underlying pattern of interest. In this context, overfit models might give very high accuracies on the training data but will not generalize to new data. On the other hand, a negative correlation between sample size and accuracy might also be a sign of publication bias rather than overfitting. A further limitation of multivariate studies in the field to date is that in most cases, the findings have not been validated in an independent dataset. Finally, each lab tends to use its own pre-processing techniques and machine learning analytical approaches, adding further complexity/heterogeneity when comparing results from different studies.

Structural and functional brain changes are not the only objective measures that can be used to aid prediction of clinical outcome in psychosis. Indeed, evidence suggests that psychosis is associated with genetic changes as well as alterations in cognitive functioning. Therefore, integrating neuroimaging, biological, clinical, and cognitive data may facilitate the multimodal prediction of psychosis onset and clinical outcomes. In this context, multivariate analysis approaches such as machine learning, which can take into account simultaneously different clinically meaningful measures, have the potential of generating valid, clinically relevant, and usable prediction models.

PSYSCAN (http://psyscan.eu) is a research program funded as part of the European Funding 7th Framework Programme that was designed to address the methodological issues described above, with the goal of translating findings from neuroimaging, genetics, clinical and cognitive measures from individuals in the early phase of psychosis (ie, CHR-P and FEP) into mainstream clinical practice.

**PSYSCAN: Translating Neuroimaging Findings From Research into Clinical Practice**

PSYSCAN is an international, longitudinal, multicenter study on the early stages of psychosis (ie, CHR-P and FEP stages) involving partners from the United Kingdom (London and Edinburgh), the Netherlands (Amsterdam, Maastricht, and Utrecht), Spain (Madrid and Santander), Denmark (Glostrup/Copenhagen), Germany (Marburg and Heidelberg), Ireland (Galway), Israel (Tel Aviv), Austria (Vienna), Switzerland (Zurich), Australia (Melbourne), Italy (Naples), plus affiliate sites in China (Hong Kong), Canada (Toronto), South Korea (Seoul), and Brazil (Sao Paulo). This consortium aims to recruit a large sample of CHR-P and FEP participants and to collect a number of multimodal imaging measures (ie, structural, resting state functional MRI, and diffusion tensor imaging data), which will be integrated with psychopathological, sociodemographic, genetic, metabolomic, proteomic, immunological and cognitive data with the aim of improving outcome prediction. Data are being collected at first presentation and again at a number of follow-up timepoints, with the same instruments and methodological procedures being used at each site. A healthy control group is also being recruited and will serve as a comparative/control group for the CHR-P as well as for the FEP cohort.

Figure 1 and table 1 show the design of the PSYSCAN study. Supplementary material 1 provides a detailed overview of the sociodemographic information, clinical, and cognitive measures collected at the different follow-up timepoints.

**Methodological Considerations in Multimodal Multicenter Studies**

Multicenter studies provide a means of acquiring data from relatively large samples of subjects, representing different geographical areas. However, the involvement of several sites also introduces methodological challenges, particularly in controlling for the effects of site differences when acquiring imaging data. Table 2 lists some potential benefits and challenges that can arise when performing multicenter studies involving imaging acquisition.

**Use of a Common Imaging Acquisition Protocol Across Sites**

One of the major issues in multicenter neuroimaging studies are the effects of intersite variations in scanner make, model, and field strength. To minimize such effects, all sites in the PSYSCAN consortium used 3T scanners, adopted a common image acquisition protocol and underwent a site qualification procedure to ensure that the standard acquisition protocol could be implemented locally (supplementary materials 2). The site qualification was led by IXICO (https://ixico.com), the industrial partner in this project. The acquisition protocol includes published pulse-sequence design (ie, Alzheimer’s Disease Neuroimaging Initiative ADNI-2 T1 and ADNI-3 FLAIR) and study-specific Diffusion Tensor Imaging (DTI) and resting state functional MRI sequences (supplementary materials 2).

**Healthy Traveling Subjects Scanned at Different Sites**

In the PSYSCAN study, although all sites are using scanners with the same field strength and harmonized acquisition parameters (supplementary materials 2), additional variability can still arise through the use of scanners that differ in model, manufacturer or specific features. The study has therefore included a travelling
subject study, an approach that has been previously adopted in other multicenter studies. Six healthy individuals have been scanned on 2 separate occasions using the PSYSCAN neuroimaging protocol, at 6 sites within the consortium. The data from these subjects will be used to quantify within- and between-scanner heterogeneity, to identify the determinants of this variance, to quantify the effects on the data, and to develop post-hoc calibration methods to attenuate these effects.

Recruitment at Different Sites

A further methodological challenge for multicenter studies is to ensure reliability of nonimaging data collection across multiple sites. Work-Package 5 (WP5) is responsible for collecting data from (1) CHR-P individuals, (2) FEP patients, and (3) healthy controls, in a naturalistic prospective design. Standardized and harmonized psychopathological, demographic, cognitive, and genetic measures are collected at baseline and during follow-up assessments. Researchers at each site completed both face-to-face and online training (http://psyscan.eu) on the instruments being used to screen and assess participants, including, but not limited to, the Positive and Negative Syndrome Scale, a revised version of the Comprehensive Assessment of an At Risk Mental State which allows the additional scoring of the Structured Interview of Psychosis-Risk Syndromes and the Schizophrenia Proneness Instrument for Adults. To further ensure reliability across sites, particularly in the assessment of CHR-P participants, teleconferences to discuss all included cases take place every 1–2 months. Centralized monitoring and yearly on-site monitoring visits are also conducted to ensure that the protocols are being followed correctly, and to address any local issues related to subject assessment and follow-up. Cognitive function is assessed on iPads using tests derived from the CANTAB cognitive battery (ie paired associates learning, spatial span task, processing speed, emotion recognition task). These are brief computerized tests participants perform using an iPad, with a total assessment time of around 20 minutes. Twenty minutes was chosen as a feasible time frame for clinical practice. The use of nonverbal rather than verbal tasks facilitates the use of the assessment in subjects with a wide range of native languages. Using a computerized assessment permits vocal instructions in the subject’s native language to be embedded within the test (translation was done from English to the 10 other languages part of the consortium using a forward/backward translation method, and a check of the vocal instructions by a native speaker).
Table 1. Aims and Expected Outcomes From PSYSCAN Work-Packages (WP)

<table>
<thead>
<tr>
<th>Work-Package</th>
<th>Aim</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP1. Management of the PSYSCAN project</td>
<td>To ensure (a) scientific dialogue and communication across the consortium and (b) efficient administration and reporting in accordance with EC guidelines and requirements</td>
<td>Overall organization and coordination of the PSYSCAN project and consortium</td>
</tr>
<tr>
<td>WP2. Merged legacy datasets</td>
<td>To collect, organize, analyze, and report on existing datasets collected by the consortium partners over the last 20 years from (a) patients with psychosis, (b) patients with genetic vulnerability for psychosis (22q11.2 deletion syndrome), (c) subjects at clinical high risk of psychosis (CHR-P), and (d) healthy controls</td>
<td>Development of new methods Data analysis on merged existing datasets New methods testing and validation</td>
</tr>
<tr>
<td>WP3. Development of software for data analysis</td>
<td>To design, code, assemble, document, and test new specialized software modules for machine learning, connectivity and network analysis</td>
<td>Development of new methods for the analysis or imaging, clinical, cognitive, and biological data Application of machine learning techniques on legacy and prospective data</td>
</tr>
<tr>
<td>WP4. Data management</td>
<td>Responsible for (a) the technical aspects of data management of both WP2 and WP5, (b) development of quality control protocols for imaging data collected as part of WP5</td>
<td>Large database with neuroimaging, demographic, cognitive and clinical data collected as part of WP5</td>
</tr>
<tr>
<td>WP5. Naturalistic prospective study</td>
<td>To collect new, homogenized data from (a) CHR-P individuals, (b) patients with first episode psychosis (FEP), and (c) healthy controls in a naturalistic prospective study comprising around 1000 subjects in total. Standardized and harmonized measures of neuroimaging, clinical, cognitive, biological, and genetic variables are collected at baseline and at follow-up to determine clinical and functional outcomes.</td>
<td>New longitudinal neuroimaging, cognitive, clinical, and biological measures collected from ~700 subjects (CHR-P, FEP, and HC)</td>
</tr>
<tr>
<td>WP6. Dissemination</td>
<td>To disseminate the activities and results of PSYSCAN, including the development of a website (psyscan.eu), annual stakeholder workshops for consultation and dissemination, production of leaflets, social media engagement (including Facebook and Twitter), publication of articles in scientific journals, and the organization of a final PSYSCAN conference</td>
<td>PSYSCAN website, publications, workshops, and conference</td>
</tr>
</tbody>
</table>

Table 2. Methodological Considerations in Multicenter Studies Involving Neuroimaging

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Challenge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of a common imaging acquisition protocol across sites</td>
<td>Site qualification is resource intensive and time consuming</td>
</tr>
<tr>
<td></td>
<td>Standardized and on-going quality check to ensure that there are no changes to the MRI protocol is resource intensive and time consuming</td>
</tr>
<tr>
<td></td>
<td>Compromise on acquisition protocols used, as not all sites have access to latest technology</td>
</tr>
<tr>
<td>Recruitment at different sites</td>
<td>Logistically and resource intensive</td>
</tr>
<tr>
<td></td>
<td>Under-recruitment at single sites can be problematic in future analyses, particularly for neuroimaging analyses</td>
</tr>
<tr>
<td>Healthy traveling subjects scanned at different sites</td>
<td>Logistically intensive, potentially expensive</td>
</tr>
<tr>
<td></td>
<td>Jet lag can influence state of alertness and affect functional MRI</td>
</tr>
<tr>
<td>Use of legacy data</td>
<td>Due to the small group, generalizability can be questionable</td>
</tr>
<tr>
<td></td>
<td>Data transfer and organization can be logistically and resource intensive</td>
</tr>
<tr>
<td>Combining imaging and nonimaging data</td>
<td>Data handling difficulties such as partial information, missing data, lack of common measurements</td>
</tr>
<tr>
<td></td>
<td>New methods for the integration of different data type have to be developed</td>
</tr>
<tr>
<td></td>
<td>High level of technical and statistical expertise needed to handle increased data complexity</td>
</tr>
</tbody>
</table>
Use of Legacy Data

In addition to the new data being acquired in WP5, existing data from previous studies have been collated from more than 3000 subjects, using 16 datasets from 9 of the consortium partners. These include structural MRI, DTI, and fMRI data, as well as nonimaging data (basic sociodemographic and clinical data), from patients with psychosis, patients with genetic vulnerability for psychosis (22q11.2 deletion syndrome), CHR-P individuals and healthy controls. These legacy data are being used to facilitate the development of novel machine learning algorithms that will be applied to the new datasets from WP5. These algorithms will initially be validated by dividing the samples into discovery and validation subsamples, and more definitively by testing the algorithms on analogous, independent datasets from other research consortia, through the Harmonization of At Risk Multisite Observational Networks for Youth (HARMONY) collaboration that includes the Personalised Prognostic Tool for early Psychosis Management (PRONIA—another EU-FP7 program) and 2 National Institute of Mental Health (NIMH)—funded programs, the North American Prodrome Longitudinal Study (NAPLS), and the Philadelphia Neurodevelopmental Cohort (PNC).

Combining Imaging and Nonimaging Data

One of the core aims of the project is to integrate neuroimaging, clinical, cognitive, and peripheral biomarker data to facilitate the prediction of psychosis onset, clinical and functional outcomes. Therefore, alongside clinical and cognitive data, blood samples (including whole blood, serum, and plasma) are collected for analyses of genomic, proteomic, metabolomic, and immune markers at baseline, 6 and 12 months in the CHR-P and HC cohorts and at baseline and 12 months in the FEP cohort. In particular, DNA will be extracted from whole blood for a GWAS analysis to allow the polygenic risk score for schizophrenia to be determined for each individual. A broad range of proteomic and metabolomic and immunological markers will be examined which can be readily determined from the frozen serum and plasma samples. Current markers in the literature have highlighted CFI and C6 proteins, reduced levels of essential polyunsaturated fatty acids and increased levels of IL-6, however, this is a rapidly developing field so we will plan to undertake both an exploratory and hypothesis-led approach based on the most recent findings at the time of analysis.

To date, most predictive algorithms in psychosis have used data from a single modality, such as MRI data or clinical data. The combination of imaging and nonimaging data may result in a more accurate model with higher predictive power compared with that of previous prediction tools. In particular, if different risk estimation tools are used in the context of a sequential and stepped assessment to enrich the risk prediction. With both legacy and newly collected data, supervised machine learning approaches will be used to predict clinically meaningful outcomes (eg, psychosis onset, social and role functioning, changes in symptom scores, and treatment response) from both neuroimaging and nonneuroimaging data. Similarly, unsupervised machine learning approaches will be used to identify subgroups of patients and investigate their clinical outcomes. These subgroups could potentially then be used to further inform supervised learning, for instance, by stratifying the subjects before making predictions.

Future Directions

Key steps for future progress in the field include the validation of prognostic and predictive algorithms in independent datasets from other projects (ie, external validation). Wide collaborations with other consortia such as HARMONY will provide the opportunity to initiate such endeavours. This independent validation of prognostic/predictive algorithms is critical to the ultimate identification of measures that can reliably predict psychosis onset in CHR-P individuals or clinical outcomes in those with a first episode of psychosis. For these measures to be embedded in day-to-day clinical practice they ideally should be collectable using methods that (1) are widely available, (2) do not require an excessive amount of patient or clinician time, and that (3) have a reasonable cost. This applies to some of the potentially useful measures in psychosis, such as MRI scanning or a blood sample for whole genome sequencing, whereas others may require technology that is relatively inaccessible (eg, positron emission tomography—PET-scanning) or analyses that are currently relatively expensive (eg, proteomics). It is therefore more likely that complex, multimodal, risk estimation algorithms would enter clinical routine only in a stepped risk assessment framework, in line with previous successful examples of clinical medicine (eg, cardiovascular and pulmonary).

Overall, PSYSCAN and similar large cohort studies are purposely designed to significantly contribute to the bench-to-bed approach and aim to develop clinically usable tools to predict psychotic illness onset and course, differential diagnosis, treatment response, and functional outcome, with practical implications for individualized treatment.

Supplementary Material

Supplementary data are available at Schizophrenia Bulletin online.

Funding

The PSYSCAN Project is supported by grant agreement no. 603196 under the European Union’s Seventh Framework Programme.
Acknowledgments

We would like to thank all participants who took part in the study. Conflict of Interest: S.G. received honoraria, advisory board, or consulting fees from the following companies: Gedeon-Richter, Janssen Pharmaceuticals, Janssen-Cilag Polska Sp. z o.o., Otsuka, Pierre Fabre and Sunovion Pharmaceauticals. B.G. is the leader of a Lundbeck Foundation Centre of Excellence for Clinical Intervention and Neuropsychiatric Schizophrenia Research (CINS), which is partially financed by an independent grant from the Lundbeck Foundation based on international review and partially financed by the Mental Health Services in the Capital Region of Denmark, the University of Copenhagen, and other foundations. Her group has also received a research grant from Lundbeck A/S for another independent investigator initiated study. All grants are the property of the Mental Health Services in the Capital Region of Denmark and administrated by them. G.S. is president of the Austrian Society of Neuropsychopharmacology and Biological Psychiatry, which is partially financed by the support from pharmaceutical companies. G.S. received consulting fees and/or honoraria for speeches within the last 3 years from Angelini, AOP Orphan, Alkermes, Janssen, Lundbeck, Pfizer. PFP received advisory board fees and research funds from Lundbeck.

References


Author/s:
Tognin, S; van Hell, HH; Merritt, K; Winter-van Rossum, I; Bossong, MG; Kempton, MJ;
Modinos, G; Fusar-Poli, P; Mechelli, A; Dazzan, P; Maat, A; de Haan, L; Crespo-Facorro, B;
Glenthoj, B; Lawrie, SM; McDonald, C; Gruber, O; van Amelsvoort, T; Arango, C; Kircher, T;
Nelson, B; Galderisi, S; Bressan, R; Kwon, JS; Weiser, M; Mizrahi, R; Sachs, G; Maatz, A;
Kahn, R; McGuire, P

Title:
Towards Precision Medicine in Psychosis: Benefits and Challenges of Multimodal Multicenter
Studies-PSYSCAN: Translating Neuroimaging Findings From Research into Clinical Practice

Date:
2020-03-01

Citation:
Tognin, S., van Hell, H. H., Merritt, K., Winter-van Rossum, I., Bossong, M. G., Kempton, M.
J., Modinos, G., Fusar-Poli, P., Mechelli, A., Dazzan, P., Maat, A., de Haan, L., Crespo-
Facorro, B., Glenthoj, B., Lawrie, S. M., McDonald, C., Gruber, O., van Amelsvoort, T.,
Arango, C., .... McGuire, P. (2020). Towards Precision Medicine in Psychosis: Benefits and
Challenges of Multimodal Multicenter Studies-PSYSCAN: Translating Neuroimaging Findings
From Research into Clinical Practice. SCHIZOPHRENIA BULLETIN, 46 (2), pp.432-441.
https://doi.org/10.1093/schbul/sbz067.

Persistent Link:
http://hdl.handle.net/11343/244642

File Description:
published version

License:
CC BY-NC