The EUropean Network of National Schizophrenia Networks Studying Gene-Environment Interactions (EU-GEI) – Incidence and First-Episode Case-Control Programme

Charlotte Gayer-Anderson[1], Hannah E Jongsma[2,3], Marta Di Forti[4], Diego Quattrone[4], Eva Velthorst[5,6,7], Lieuwe de Haan[7], Jean-Paul Selten[8,9], Andrei Szöke[10,11,12], Pierre-Michel Llorca[13], Andrea Tortelli[14], Celso Arango[15], Julio Bobes[16], Miguel Bernardo[17], Julio Sanjuán[18], José Luis Santos[19], Manuel Arrojo[20], Mara Parellada[15], Ilaria Tarricone[21], Domenico Berardi[21], Mirella Ruggeri[22], Antonio Lasalvia[22,23], Laura Ferraro[24], Caterina La Cascia[24], Daniele La Barbera[24], Paulo Rossi Menezes[25], Cristina Marta Del-Ben[26], **EU-GEI WP2 Group[††]**, Bart P Rutten[9], Jim van Os[9,27,28], Peter B Jones[3,29], Robin M Murray[28], James B Kirkbride[2], Craig Morgan[1]*

Correspondence to: Craig Morgan, Department of Health Services and Population Research, Institute of Psychiatry, Psychology & Neuroscience, King's College London, 16 De Crespigny Park, London, SE5 8AF, England (email: craig.morgan@kcl.ac.uk; telephone: +44 (0)20 7848 0351)

Author affiliations:

[1]Department of Health Services and Population Research, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, England

[2]PsyLife Group, Division of Psychiatry, University College London, London, England

[3]Department of Psychiatry, University of Cambridge, Cambridge, England

[4]Social, Genetic, and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, England

[5]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, USA

[6]Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, New York, USA

[7]Early Psychosis Section, Department of Psychiatry, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands

[8]GGZ Rivierduinen, Institute for Mental Health, Leiden, the Netherlands

[9]Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University Medical Centre, Maastricht, the Netherlands

[10]Institut National de la Santé et de la Recherche Médicale, U955, Equipe 15 Neuro-Psychiatrie Translationnelle, Créteil, France

[11]AP-HP, Pôle de Psychiatrie des Hôpitaux Universitaires Henri Mondor, Créteil, France

[12]Fondation FondaMental, Créteil, France

[13]EA 7280 Npsydo, Université Clermont Auvergne, Clermont-Ferrand, France.

[14]Establissement Public de Santé, Maison Blanche, Paris, France

[15]Department of Child and Adolescent Psychiatry, Hospital General Universitario Gregorio Marañón, School of Medicine, Universidad Complutense, ISGM, CIBERSAM, Madrid, Spain

[16]Department of Medicine, Psychiatry Area, School of Medicine, Universidad de Oviedo, Centro de Investigación Biomédica en Red de Salud Mental, Oviedo, Spain

[17]Barcelona Clinic Schizophrenia Unit, Neuroscience Institute, Hospital Clinic, Department of Medicine, University of Barcelona, Institut d'Investigacions Biomèdiques, August Pi I Sunyer, Centro de Investigación Biomédica en Red de Salud Mental, Barcelona, Spain

[18]Department of Psychiatry, School of Medicine, Universidad de Valencia, Centro de Investigación Biomédica en Red de Salud Mental, Valencia, Spain

[19]Department of Psychiatry, Hospital "Virgen de la Luz," Cuenca, Spain

[20]Department of Psychiatry, Psychiatric Genetic Group, Instituto de Investigación Sanitaria de Santiago de Compostela, Complejo Hospitalario Universitario de Santiago de Compostela, Spain

[21]Department of Biomedical and NeuroMotor Sciences, Psychiatry Unit, Alma Mater Studiorium Università di Bologna, Bologna, Italy

[22]Section of Psychiatry, Department of Neuroscience, Biomedicine and Movement, University of Verona, Verona, Italy

[23]Section of Psychiatry, Azienda Ospedaliera Universitaria Integrata di Verona, Verona, Italy

[24]Department of Experimental Biomedicine and Clinical Neuroscience, Section of Psychiatry, University of Palermo, Palermo, Italy

[25]Department of Preventive Medicine, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil

[26]Division of Psychiatry, Department of Neuroscience and Behaviour, Ribeirão Preto Medical School, Universidade de São Paulo, São Paulo, Brazil

[27]Department of Psychiatry, Brain Center Rudolf Magnus, Utrecht University Medical Centre, Utrecht, the Netherlands

[28]Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, England

[29]CAMEO Early Intervention Service, Cambridgeshire and Peterborough National Health Service Foundation Trust, Cambridge, England

[††] **For a full list of collaborators, see Supporting Document**

Abstract word count: 198

Main body word count: 3,583

ABSTRACT

Purpose: The EUropean Network of National Schizophrenia Networks Studying Gene-Environment Interactions (EU-GEI) study contains an unparalleled wealth of comprehensive data that allows for testing hypotheses about i) variations in incidence within and between countries, including by urbanicity and minority ethnic groups; and ii) the role of multiple environmental and genetic risk factors, and their interactions, in the development of psychotic disorders.

Methods: Between 2010 and 2015, we identified 2,774 with incident cases of psychotic disorders during 12.9 million person-years at risks, across 17 sites in six countries (UK, the Netherlands, France, Spain, Italy, and Brazil). Of the 2,774 incident cases, 1,130 cases were assessed in detail and form the case sample for case-control analyses. Across all sites, 1,497 controls were recruited and assessed. We collected data on an extensive range of exposures and outcomes, including demographic, clinical (e.g., premorbid adjustment), social (e.g., childhood and adult adversity, cannabis use, migration, discrimination), cognitive (e.g., IQ, facial affect processing, attributional biases), and biological (DNA via blood sample/cheek swab). We describe the methodology of the study and some descriptive results, including representativeness of the cohort.

Conclusions: This resource constitutes the largest and most extensive incidence and case-control study of psychosis ever conducted.

INTRODUCTION

The lifetime prevalence of psychotic disorders is around three per cent [1]. The associated individual, familial, social, and economic costs are vast. Psychotic disorders cause considerable distress to sufferers and their families and often lead to marked social dysfunction and exclusion. The economic costs are huge: in Europe, an estimated €94 billion per year [2], of which over half is due to the indirect costs of unemployment, lost productivity, and informal care, etc. [3]. The World Health Organisation estimated that in Western countries, the treatment and care of patients with a psychotic disorder range from 1.6% to 2.6% of total healthcare expenditures [4]. Further, individuals with a psychotic disorder are far more likely to have a physical health problem [5] and to die younger, by as much as 20 years on average, compared with the general population [6].

Our knowledge of the distribution and determinants of psychotic disorders has increased in recent years. The incidence varies by area (e.g., higher in some urban versus rural areas) [7,8] and social group (e.g., higher in some minority ethnic groups) [9,10] and, in addition to well-established genetic and neurodevelopmental risk factors [11,12], there is now substantial evidence implicating several environmental risk factors [13], such as childhood adversity [14,15] and cannabis use [16]. Pooled relative risks for these risk factors range between two and four, and population attributable risk fractions range between 20 and 35% [17,18]. Further, there is accumulating evidence that these myriad risk factors interact in complex ways to increase risk of psychosis via effects on the dopaminergic system, dysregulation of which may be the biological process that underpins the formation of psychotic experiences.

However, there remain many gaps, inconsistencies, and unanswered questions, and recent work hints at different patterns of risk in different settings. For example, recent evidence has failed to show a universal association between city living and psychosis [19,20]. To further add to this conundrum, Colodro-Conde and colleagues [21] found that the high prevalence of psychosis in some

urban areas may be due to gene-environment selection, such that individuals with higher genetic loading for psychosis live in more densely populated areas. More generally, this points to a major limitation to our current knowledge of psychotic disorders: we know that environments affect onset and outcomes, but research so far has been conducted – with some important exceptions – in a remarkably small number of settings (i.e., select centres in the US, UK, and Australasia). Combined, these points emphasise the need for research in more diverse contexts to examine more nuanced hypotheses on the complex interplay between biology and environments in the aetiology of psychotic disorders.

Our knowledge of psychotic disorders is limited, in part, because of heterogeneity in methods, which limits our ability to compare findings across populations [22]. For example, differences in study design (i.e. case-register, versus cohort-based designs, versus first-contact studies), the age structures of populations at risk, case-identification procedures, diagnostic criteria, definitions and measurement of environmental factors, and analytic strategies, has made cross-country comparisons difficult and likely obscured important clues to aetiology [23]. The only large-scale international comparative studies conducted to date are the World Health Organisation's multi-country projects of the 1970s and 1980s, which compared the incidence and clinical and social characteristics of treated cases of psychoses from twelve diverse settings in ten countries using a standardised procedure for case identification and data collection [24]. However, since this landmark programme, there have been far-reaching economic and social changes (e.g. migration patterns, cannabis availability and use, and distribution of social risks) with conceivable impacts on the social epidemiology and aetiology of the psychoses. Moreover, studies of environment-gene interactions in psychotic disorders are rare and have typically involved small samples, with limited phenotyping and limited assessment of environmental factors [21,25,26].

The EU-GEI programme was established to address these gaps and limitations [27]. EU-GEI is a multi-

national research collaboration that was funded for five years (1 May 2010 – 30 April 2015). It consisted of 11 Work Packages [see **Supplementary Table S1**]. This paper profiles the incidence and case-control programme of work (Work Package 2), which comprises the largest multi-site study of psychotic disorders ever conducted. In this paper, we describe the objectives and main aspects of the study.

Objectives

The overall goal of the present work package was to investigate the role of multiple environmental and genetic risk factors, and their interactions, in the development of psychotic disorders. Specifically, our aims were: 1) to investigate the impact of hypothesized environmental exposures, measured at individual and area levels, on: a) risk of psychotic disorders; and b) high rates of disorder in urban areas and in migrant and minority ethnic groups; and 2) to examine hypothesized a) gene x environment interactions (GxE); and b) environment x environment interactions (ExE) across the life course.

METHODS

Study design

The data resource comprises a multi-site population-based incidence and case-control sample of cases with a first episode of psychosis (International Classification of Diseases [ICD]-10 diagnoses F20-29 and F30-33) and controls drawn from tightly defined catchment areas in 17 sites in six countries (England, the Netherlands, France, Spain, Italy, and Brazil; see **Figure 1**). The sites were purposefully selected to include a mix of urban and rural areas, with varying proportions from minority ethnic groups (see **Table 1**).

[Insert **Fig 1** Map of EU-GEI settings for the Incidence and Case-Control Work Package]

Sample

Recruitment and data collection were conducted over a five-year period between 2010 and 2015 (**Table 1**). We also added data from the Veneto region, Italy, collected as part of an earlier study (the Psychosis Incident Cohort Outcome Study [PICOS]; 2005-2007), but with sufficiently similar methods to be pooled with that collected for this study. The incidence sample comprised 2,774 individuals with a first episode of psychosis. Of these, 1,519 were approached, and 1,130 were consented and assessed (41% of the total incidence sample). Reasons for non-participation among cases who were approached were: refusal to participate, language barriers, and exclusion after consenting as they did not meet the age inclusion criteria. In addition, 1,497 controls were recruited and assessed.

[Insert Table 1]

Statistical power

Our sample of 1,130 cases and 1,497 controls has high statistical power to test our primary study hypotheses, even after accounting for missing data and for the current necessity of restricting genetic analyses to individuals of non-African ancestry. For example, in a restricted sample of cases 1,031 and 1,438 controls, we have greater than 80% power to detect an interaction odds ratio of 1.2 at $p <= 0.05$, assuming an odds ratio of 2.0 for an environmental exposure and of 1.2 for each unit increase in polygenic score (assuming N (0,1) distribution).

Case ascertainment and recruitment

All cases presenting to one of the 17 participating centres in 6 countries with a suspected first episode of psychosis were potentially eligible for inclusion in the study. The inclusion criteria for cases were: a) presence of at least one positive psychotic symptom for at least one day duration or two negative psychotic symptoms (for at least six months duration) within the time-frame of the study; b) aged between 18 and 64 years (inclusive); and c) resident within a clearly defined

catchment area at the time of their first presentation. Residence was defined as a minimum of a one night stay at a residential address within the catchment areas. Exclusion criteria were: a) previous contact with specialist mental health services for psychotic symptoms outside of the study period at each site; b) evidence of psychotic symptoms precipitated by an organic cause (ICD-10: F09); c) transient psychotic symptoms resulting from acute intoxication (F1X.5); d) severe learning disabilities, defined by an IQ less than 50 or diagnosis of intellectual disability (F70 to F79); and, for the case-control part only, e) insufficient fluency of the primary language at each site to complete assessments.

Case identification procedures involved teams of researchers regularly screening both general adult and specialist mental health services (both in- and out-patient). The screening process involved researchers regularly liaising with clinical staff and checking clinical records to identity potential cases. The researchers only included those individuals who they could be sure met the criteria based on the symptoms reported in the clinical notes. Potential cases were then approached when considered appropriate by clinical staff and informed consent sought.

Control recruitment

Inclusion criteria for controls were: a) aged between 18 and 64 years; b) resident within a clearly defined catchment area at thetime of consent into the study; c) sufficient command of the primary language at each site to complete assessments; and d) no current or past psychotic disorder. To select a population-based sample of controls broadly representative of local populations in relation to age, gender, and ethnicity, a mixture of random and quota sampling was used. Quotas for control recruitment were based on the most accurate local demographic data available. Quotas were then filled using a variety of recruitment methods, including through: 1) random sampling from lists of all postal addresses (e.g., in London); 2) stratified random sampling via GP lists (e.g., in London and Cambridge) from randomly selected surgeries; and 3) *ad hoc* approaches (e.g., internet and

newspaper adverts, leaflets at local stations, shops, and job centers). In some sites (e.g., London), some groups (e.g., black African and black Caribbean) were oversampled to enable subsequent sub-group analyses. To deal with this in subsequent analyses, weights were generated, based on the most accurate local demographic data available, to minimize any resulting bias in estimating the prevalence of exposures among controls.

Individuals who agreed to take part were screened for a history of psychosis. Those who reported previous or current treatment for psychosis were excluded. Those who responded positively to any question in the screening instrument, indicating a possible psychotic experience, were interviewed further with standardised interviews to assess symptoms and to establish presence or otherwise of a psychotic disorder. On this basis, no potential controls were found to have a past or current psychotic disorder.

Data contents

We collected data on an extensive range of exposures and outcomes across multiple domains using previously validated questionnaires, tasks, and procedures: demographic, clinical, social, psychological, cognitive, and biological (**Table 2**). All environmental exposures and cognitive and psychological tests were measured using previously validated questionnaires and tasks.

Genetic risk was assessed both indirectly, using a familial liability score for psychosis [28], and directly, using DNA extracted from two 9ml non-fasting venous blood samples and/or via saliva samples (Oragene). Samples were genotyped using custom Illumina HumanCoreExome-24 BeadChip genotyping arrays containing probes for 570,038 genetic variants (Illumina Inc., San Diego, CA). Genotype data were called using the GenomeStudio package, transferred into PLINK format for further analysis, and underwent quality control based on genotype variants and samples.

[Insert Table 2]

Quality assurance and control

Prior to and during data collection, annual multi-site meetings were arranged to bring together principal investigators and core researchers to ensure that standardised procedures were being implemented, to provide training, to discuss issues with data collection, and to conduct inter-rater reliability exercises. The study was designed to ensure comparable procedures and methods across settings, with some local adaptation to allow for variations in healthcare provision and health service contact points. The primary deviation from protocol was in the Veneto region, Italy, where data were derived from a previous study which used comparable methods [29], but had a lower upper-age limit of 54.

Training of researchers who were responsible for administrating the assessments was performed at the outset and throughout the study. This was organised by a technical working committee of the overall EU-GEI study (Work Package 11). An online resource was made available with taped interviews, samples of recordings, and written summaries for staff training purposes. Inter-rater reliability was assessed annually. Researchers were required to attain and maintain a minimum threshold of correct ratings before being allowed to administer the core assessments. Sufficient levels of inter-rater reliability for the core measurements, ranging from 0.70-0.91, were achieved, and are shown in **Table 3**.

[Insert Table 3]

Data management

Data was collected on paper and, for some cognitive tasks (e.g., the White Noise task), on laptops and securely stored at each of the participating centres, and was entered locally using an encrypted

web-based system, using commercial software (4D) that was adapted specifically for EU-GEI purposes. Data was entered once with field codes restricted to logical values where possible, to minimise data entry errors. Blood or saliva samples were taken at approved clinical research facilities by an experienced researcher, were fully anonymized and identified by bar code, and sent to the Institute of Psychological Medicine and Clinical Neurology at Cardiff University for genotyping. The data resource has undergone a rigorous period of validation checks and cleaning by a small number of experienced researchers. This has involved checks of missing data and corroboration of these against the paper files at each of the 17 sites.

Ethical approval

All participants who agreed to take part in the study provided informed, written consent following full explanation of the study. Ethical approval for the study was provided by relevant research ethics committees in each of the study sites [30].

RESULTS

Sample representativeness and characteristics

There were similar proportions from minority ethnic groups among consented and non-consented cases (43% vs. 40%). However, the proportion of men and the proportions in younger age groups were higher among consented, compared with non-consented, cases (men: 62% vs. 57%; aged 18 to 34 years: 69% vs. 60%) (**Table 4a**). Compared with the general population, controls were more likely to belong to a minority ethnic group (controls: 28%, population at-risk: 23%) and were younger (aged between 18 to 34 years, controls: 56%, population at-risk: 38%) (**Table 4b**). The greater proportion of controls who were from ethnic minority reflects oversampling in some sites (e.g., London) to enable subsequent sub-group analyses.

Cases were younger than controls (median age of cases was 29 years (interquartile range [IQR] 22-

37), and controls 33 years old (IQR 26-47)). Compared with controls, a greater proportion of the cases were men (62% vs. 49%), migrants (28% vs. 22%), and left school without any qualifications (16% versus 6%); a smaller proportion was of white ethnicity (63% versus 73%) (see **Supplementary Table S2**).

[Insert Table 4]

DISCUSSION

This study was conducted in a diverse range of settings across Europe and one setting in Brazil, selected to ensure a mix of urban and rural areas with large migrant and ethnic minority populations. This maximises its applicability to and importance for public health initiatives, with potential implications for both prevention and intervention, particularly among minority ethnic groups, and in urban areas, and in relation to cannabis and other substance use and developmental adversity. Our primary hypotheses centre on examining variations in incidence and symptoms, environmental risk factors, and the interplay between environment and genetic factors in the development of psychotic disorders.

Incidence and symptoms

We have already published findings of the overall variations in incidence of psychoses by site [30]. Our findings suggest marked geographical differences in the incidence of psychotic disorders, with around an 8-fold variation among study sites after accounting for age, sex, and minority ethnic status. At an area level, initial analyses suggest that some of this variation may be related to the proportion of owner-occupied homes in an area (a tentative proxy for social cohesion or socioeconomic deprivation), i.e., areas with more owner-occupied homes had, on average, lower rates of psychotic disorder. Analyses of variations in incidence by ethnic group are ongoing. Analyses

of symptom data on incident cases, collated using the OPCRIT, have examined the validity of a transdiagnostic dimensional structure of psychopathology and, in doing so, have challenged the common binary categorisation of psychoses into non-affective and affective disorders [31]. Our findings suggest that a bifactor model of psychopathology, comprising one general factor and five dimensions (positive, negative, manic, disorganised, and depressive symptoms), best represents the structure of symptoms among those with a psychotic disorder. We further found: compared with majority populations, cases in minority ethnic groups scored higher on the positive psychotic symptom dimension; and, compared with rural areas, cases in urban areas scored higher on the general symptom dimension.

Environmental risk

The initial focus of analyses of our case-control data resource are the associations and population impact of putative environmental risk factors, including childhood adversity and abuse, adult adversity, discrimination, and cannabis use. In analyses of cannabis use data, for example, we found that, compared with those who did not use cannabis, the odds of psychosis were: (1) around 3 times higher among those who used cannabis daily; (2) around 2 times higher among those who spent more than 20 Euros a week on cannabis; and (3) around 50% higher among those who used cannabis high in THC [32]. In addition, we found variations in population attributable fractions for daily cannabis use on psychosis [32], with population attributable fractions (i.e., the proportion of psychosis, assuming causality, attributable to daily use) ranging from 1% (in Puy-de-Dôme, France) to 44% (in Amsterdam). Similar analyses examining childhood and adult adversity are ongoing, focusing on type, severity, and age of exposure (Morgan et al., in preparation).

These analyses will be further extended to examine environment-environment and gene-environment interactions and to more clearly elucidate the pathogenic processes underpinning observed variations in incidence across study sites [30] and high rates of psychotic disorders in urban

areas [7,8], and in migrant and ethnic minority groups [9,10].

Strengths and weaknesses

To our knowledge, this is the most extensive multi-site incidence and case-control study of first-episode psychosis ever conducted, with comprehensive data on a variety of environmental, psychological, and genetic risk factors. The primary strength of the EU-GEI study is its potential to provide ground-breaking and important information about the development of psychoses, by investigating the complex interrelationships between candidate environmental, psychological, and biological (genetic) factors and psychotic disorders, including the mechanisms through which they increase risk. In addition, given that our study was carried out in major urban and rural sites with heterogeneous populations suggests that our external validity may extend to other centres with similar population profiles. The combined incidence and case-control methodology allows for precise identification of, and ability to account for, any potential selection biases amongst the recruited and assessed cases. The richness of the exposure information available will allow for more nuanced analyses and a more fine-grained understanding of their impact on psychotic disorder than has been possible to date. Importantly, the inclusion of only cases with a first episode of psychosis (rather than individuals with long-standing disorder) allows inferences to be made about causal connections and processes.

The primary limitation of this data resource relates to case identification. As in all previous studies, we relied on first contact with mental health services as a proxy for first onset. While it is likely most individuals who develop a psychotic disorder do present to services, at least in sites with well-developed public health systems, some who do not present will be missed and this may introduce selection biases. Any rate estimates should therefore be considered as treated incidence. Further, variations in referral procedures of patients with psychosis from primary to secondary mental health care settings and in the organization of secondary mental health care services across catchment

areas may have influenced the identification of cases, and may explain some of the variation in estimates of incidence across study sites and countries. For example, unlike in other settings, patients in Madrid are not constrained to using mental health services in their residential catchment areas [33]. However, as highlighted by Jongsma et al. (2018) [30], the divergences in service provision and cultural context are unlikely to fully explain the 8-fold variation in incidence across sites.

There are also several limitations that are inherent to case-control designs. First, while substantial efforts were made at the outset to reduce the potential biases in the identification of cases (e.g., recruitment of participants from a number of sources using a variety of methods, including inpatient wards and community teams) and controls (e.g., use of mixture of random and quota sampling), we were not entirely successful; our cases are not fully representative of the sample identified in the incidence study, and our controls not of the population-at-risk. For example, reliance in some sites on recruitment of controls through ad hoc methods, such as newspaper advertisements, may have biased samples. Interpretations of estimated effects (odds ratios) should be considered with this in mind.

Second, there is the potential for both recall and observer bias. In order to minimise these, and validate environmental exposures, several steps were taken. For core environmental exposures (e.g. childhood adversity and cannabis use), we used extensive, well-validated measures, that drew on life course methods to anchor memories and improve recall. All researchers administering these assessments went through intensive training, with regular top-ups. Further, where possible, we drew on corroborative sources of information in the assessment of exposure to childhood and adulthood adversity (e.g. clinical records, interviews with siblings of a subsample of cases (n = 272)).

Third, measurement of exposure occurred after onset of disorder, making causal inferences problematic. To establish the temporal ordering of exposure and outcome, we carefully established

the date of onset of disorder and, for measures of exposures in childhood and adulthood, ensured that all assessments related to the period pre-onset.

Finally, given the large battery of tests and interviews conducted with our participants, data were missing for some assessments, particularly towards the end of the study battery. Where appropriate, a standardised procedure for multiple imputation will be used to minimise the loss of precision or selection biases which may otherwise be introduced in complete case analyses.

Data resource access

The EU-GEI WP2 principal investigators [contact: craig.morgan@kcl.ac.uk] welcome formal requests for access to the data, biological samples, and/or collaborative projects. Researchers will be required to complete an EU-GEI WP2 data interest form to state their intended hypotheses and analysis plan, which will be reviewed by the PIs to determine whether the proposal can be addressed by this data resource, does not duplicate on-going or completed analyses with this dataset, and lies within the scope of current ethical approvals. More information about the study can be found on the study website (http://www.eu-gei.eu/).

Conflict of Interest: On behalf of all authors, the corresponding author states that there is no conflict of interest.

REFERENCES

1. Perälä J, Suvisaari J, Saarni SI, Kuoppasalmi K, Isometsä E, Pirkola S, Partonen T, Tuulio-Henriksson A, Hintikka J, Kieseppä T, et al. (2007) Lifetime prevalence of psychotic and bipolar I disorders in a general population. Arch Gen Psych 64 (1):19-28

2. Olesen J, Gustavsson A, Svensson M, Wittchen HU, Jönsson B, CDBE2010 Study Group, European Brain Council (2012) The economic cost of brain disorders in Europe. Eur J Neurol 19 (1):155-162

3. Chong HY, Teoh SL, Wu DB-C, Kotirum S, Chiou C-F, Chaiyakunapruk N (2016) Global economic burden of schizophrenia: a systematic review. Neuropsychiatr Dis Treat 12:357

4. Barbato A (1998) Nations for mental health: Schizophrenia and public health. Geneva: World Health Organization Division of Mental Health and Prevention of Substance Abuse (MSA)

5. De Hert M, Cohen D, Bobes J, Cetkovich-Bakmas M, Leucht S, Ndetei DM, Newcomer JW, Uwakwe R, Asai I, Möller HJ, et al. (2011) Physical illness in patients with severe mental disorders. II. Barriers to care, monitoring and treatment guidelines, plus recommendations at the system and individual level. World Psychiatry 10 (2):138-151

6. Laursen TM, Munk-Olsen T, Vestergaard M (2012) Life expectancy and cardiovascular mortality in persons with schizophrenia. Curr Opin Psychiatry 25 (2):83-88

7. Vassos E, Pedersen CB, Murray RM, Collier DA, Lewis CM (2012) Meta-analysis of the association of urbanicity with schizophrenia. Schizophr Bull 38 (6):1118-1123

8. Kirkbride JB, Keyes KM, Susser E (2018) City Living and Psychotic Disorders—Implications of Global Heterogeneity for Theory Development. JAMA Psychiatry

9. Kirkbride JB, Hameed Y, Ioannidis K, Ankireddypalli G, Crane CM, Nasir M, Kabacs N, Metastasio A, Jenkins O, Espandian A, et al. (2017) Ethnic minority status, age-at-immigration and psychosis risk in rural environments: evidence from the SEPEA study. Schizophr Bull 43 (6):1251-1261

10. Bourque F, van der Ven E, Malla A (2011) A meta-analysis of the risk for psychotic disorders among first-and second-generation immigrants. Psychol Med 41 (5):897-910

11. Esterberg ML, Trotman HD, Holtzman C, Compton MT, Walker EF (2010) The impact of a family history of psychosis on age-at-onset and positive and negative symptoms of schizophrenia: a meta-analysis. Schizophr Res 120 (1-3):121-130

12. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014) Biological insights from 108 schizophrenia-associated genetic loci. Nature 511:421-427

13. Belbasis L, Köhler C, Stefanis N, Stubbs B, van Os J, Vieta E, Seeman M, Arango C, Carvalho A, Evangelou E (2018) Risk factors and peripheral biomarkers for schizophrenia spectrum disorders: an umbrella review of meta-analyses. Acta Psychiatr Scand 137 (2):88-97

14. Varese F, Smeets F, Drukker M, Lieverse R, Lataster T, Viechtbauer W, Read J, van Os J, Bentall RP (2012) Childhood adversities increase the risk of psychosis: a meta-analysis of patient-control, prospective-and cross-sectional cohort studies. Schizophr Bull 38 (4):661-671

15. Morgan C, Gayer-Anderson C (2016) Childhood adversities and psychosis: evidence, challenges, implications. World Psychiatry 15 (2):93-102

16. Colizzi M, Murray R (2018) Cannabis and psychosis: what do we know and what should we do? Br J Psychiatry 212 (4):195-196

17. Di Forti M, Marconi A, Carra E, Fraietta S, Trotta A, Bonomo M, Bianconi F, Gardner-Sood P, O'Connor J, Russo M, et al. (2015) Proportion of patients in south London with first-episode psychosis attributable to use of high potency cannabis: a case-control study. Lancet Psychiatry 2 (3):233-238

18. Wicks S, Hjern A, Gunnell D, Lewis G, Dalman C (2005) Social adversity in childhood and the risk of developing psychosis: a national cohort study. Am J Psychiatry 162 (9):1652-1657

19. DeVylder JE, Kelleher I, Lalane M, Oh H, Link BG, Koyanagi A (2018) Association of urbanicity with psychosis in low-and middle-income countries. JAMA Psychiatry 75 (7):678-686

20. Fett A-K, Lemmers-Jansen I, Krabbendam L (2019) Psychosis and urbanicity–a review of the recent literature from epidemiology to neurourbanism. Curr Opin Psychiatry 32(3):232-241

21. Colodro-Conde L, Couvy-Duchesne B, Whitfield JB, Streit F, Gordon S, Kemper KE, Yengo L, Zheng Z, Trzaskowski M, De Zeeuw EL, et al. (2018) Association between population density and genetic risk for schizophrenia. JAMA Psychiatry 75 (9):901-910

22. McGrath J, Saha S, Welham J, El Saadi O, MacCauley C, Chant D (2004) A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. BMC Med 2 (1):13

23. McGrath JJ (2003) Invited commentary: Gaining traction on the epidemiologic landscape of schizophrenia. Am J Epidemiol 158 (4):301-304

24. Jablensky A, Sartorius N, Ernberg G, Anker M, Korten A, Cooper JE, Day R, Bertelsen A (1992) Schizophrenia: manifestations, incidence and course in different cultures. A World Health Organization Ten-Country Study. Psychol Med Supplement 20:1-97

25. Hartz SM, Horton AC, Hancock DB, Baker TB, Caporaso NE, Chen L-S, Hokanson JE, Lutz SM, Marazita ML, McNeil DW, et al. (2018) Genetic correlation between smoking behaviors and schizophrenia. Schizophr Res 194:86-90

26. Sariaslan A, Fazel S, D'onofrio B, Långström N, Larsson H, Bergen S, Kuja-Halkola R, Lichtenstein P (2016) Schizophrenia and subsequent neighborhood deprivation: revisiting the social drift hypothesis using population, twin and molecular genetic data. Translational Psychiatry 6 (5):e796

27. European Network of National Networks studying Gene-Environment Interactions in Schizophrenia (EU-GEI) (2014) Identifying gene-environment interactions in schizophrenia: contemporary challenges for integrated, large-scale investigations. Schizophr Bull 40 (4):729-736

28. Verdoux H, Van Os J, Sham PC, Jones PB, Gilvarry K, Murray R (1996) Does familiality predispose to both emergence and persistence of psychosis?: a follow-up study. Br J Psychiatry 168 (5):620-626

29. Lasalvia A, Bonetto C, Tosato S, Zanatta G, Cristofalo D, Salazzari D, Lazzarotto L, Bertani M, Bissoli S, De Santi K, et al. (2014) First-contact incidence of psychosis in north-eastern Italy: influence of age, gender, immigration and socioeconomic deprivation. Br J Psychiatry 205 (2):127-134

30. Jongsma HE, Gayer-Anderson C, Lasalvia A, Quattrone D, Mulè A, Szöke A, Selten J-P, Turner C, Arango C, Tarricone I, et al. (2018) Treated incidence of psychotic disorders in the multinational EU-GEI study. JAMA Psychiatry 75 (1):36-46

31. Quattrone D, Di Forti M, Gayer-Anderson C, Ferraro L, Jongsma HE, Tripoli G, La Cascia C, La Barbera D, Tarricone I, Berardi D, et al. (2019) Transdiagnostic dimensions of psychopathology at first episode psychosis: findings from the multinational EU-GEI study. Psychol Med 49 (8):1378-1391

32. Di Forti M, Quattrone D, Freeman TP, Jongsma HE, Tripoli G, Gayer-Anderson C, Quigley H, Rodriguez V, Ferraro L, La Cascia C, et al. (2019). The contribution of cannabis use to variation in the incidence of psychotic disorder across Europe: the EUGEI case-control study. Lancet Psychiatry 6 (5):427-436

33. Boletín Oficial del Estado. Law 6/2009, November 16; Freedom of Choice in the Health of the Community of Madrid [in Spanish]. https://www.boe.es/buscar/pdf/2010/BOE-A-2010-2187-consolidado.pdf. Published November 2009. Accessed December 2018.