Epidemiology and Genetics of Common Mental Disorders in the general population: the PEGASUS-Murcia project

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-004035
Article Type:	Protocol
Date Submitted by the Author:	15-Sep-2013
Complete List of Authors:	Navarro-Mateu, Fernando; Unidad Docencia, Investigación y Formación en Salud Mental (UDIF-SM), Subdirección Gral. Salud Mental; CIBER de Epidemiología y Salud Pública (CIBERESP), Tormo, Mª José; Consejería de Sanidad, Servicio de Epidemiología; CIBER de Epidemiología y Salud Pública (CIBERESP), Vilagut, Gemma; IMIM-Institut Hospital del Mar d'Investigacions Médiques, ; CIBER de Epidemiología y Salud Pública (CIBERESP), Alonso, Jordi; IMIM-Institut Hospital del Mar d'Investigacions Médiques, ; CIBER de Epidemiología y Salud Pública (CIBERESP), Ruiz-Merino, Guadalupe; Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia, Escámez, Teresa; IMIB BIOBANC-MUR. Biobanco-HUVA-AECC-FFIS, ; Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia, Salmerón, Diego; Consejería de Sanidad, Servicio de Epidemiología; CIBER de Epidemiología y Salud Pública (CIBERESP), Júdez, Javier; Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia, ; IMIB BIOBANC-MUR. Biobanco-HUVA-AECC-FFIS, Martínez, Salvador; Instituto de Neurociencias. UMH-CSIC, Navarro, Carmen; Consejería de Sanidad, Servicio de Epidemiología; CIBER de Epidemiología y Salud Pública (CIBERESP),
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Epidemiology, Genetics and genomics, Public health
Keywords:	MENTAL HEALTH, EPIDEMIOLOGY, GENETICS

STUDY PROTOCOL

Title: **Epidemiology and Genetics of Common Mental Disorders in the general population: the PEGASUS-Murcia project**

Authors: F. Navarro-Mateu[1,2,*], M.J. Tormo[2,3,4], G. Vilagut[2,5], J. Alonso[2,5,6], G. Ruíz-Merino[7], T. Escámez[7,8], D. Salmerón[2,3,4], J. Júdez[7,8], S. Martínez[9], C. Navarro[2,3,4]

[1] Unidad de Docencia, Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud. Murcia. Spain.
[2] CIBER de Epidemiología y Salud Pública (CIBERESP). Spain.
[3] Servicio de Epidemiología. Consejería de Sanidad y Política Social. Murcia. Spain.
[4] Departamento de Ciencias Sociosanitarias. Universidad de Murcia. Spain.
[5] IMIM-Institut Hospital del Mar d´Investigacions Médiques. Barcelona. Spain.
[6] Departamento de Salud y Ciencias Experimentales. Universidad Pompeu Fabra. Barcelona. Spain.
[7] Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia. Spain.
[8] IMIB BIOBANC-MUR. Biobanco-HUVA-AECC-FFIS. Murcia. Spain.
[9] Instituto de Neurociencias. UMH-CSIC. Alicante. Spain.

* **Corresponding author:** Fernando Navarro-Mateu, MD, PhD. Unidad de Docencia, Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud. c/ Lorca, nº 58. 30120-El Palmar (Murcia). Spain. Email: fernando.navarro@carm.es

Emails of all coauthors

FNM: fernando.navarro@carm.es
MJT: mjose.tormo@carm.es
GV: gvilagut@imim.es
JA: jalonso@imim.es
GRM: guadalupe.ruiz@carm.es
TE: teresa.escamez@carm.es
DS: diego.salmeron@carm.es
JJ: javier.judez@ffis.es
SM: smartinez@umh.es
CN: carmen.navarro@carm.es

KEYWORDS: **Cross-sectional survey, mental disorders, prevalence, gene-environmental interactions, genome, epigenome, transcriptome.**

Word count (excluding title page, abstract, references, figures and tables)**: 5 498**

ABSTRACT (296 words)

Background: Multidisciplinary collaboration between clinicians, epidemiologists, neurogeneticists and statisticians on research projects has been encouraged to improve our knowledge of the complex mechanisms underlying the etiology and burden of mental disorders. The PEGASUS-Murcia project was designed to assess the prevalence of common mental disorders, to identify risk and protective factors and it also included the collection of biological samples to study gene-environmental interactions in the context of the World Mental Health Survey Initiative.

Methods and Analysis: The PEGASUS-Murcia project is a new cross-sectional face-to-face interview survey based on a representative sample of non-institutionalized adults in the Region of Murcia (Mediterranean Southeast, Spain). Trained lay interviewers used the latest version for use in Spain of the computer-assisted personal interview (CAPI) of the Composite International Diagnostic Interview (CIDI 3.0), specifically adapted for the project. Two biological samples of buccal mucosal epithelium were collected from each interviewed participant, one for DNA extraction for genomic and epigenomic analyses and the other to obtain mRNA for gene expression quantification. Several quality control procedures were implemented to assure the highest reliability and validity of the data. This paper describes the rationale, sampling methods and questionnaire content as well as the laboratory methodology.

Ethics and dissemination: Informed consent was obtained from all participants and the protocol was approved by a Regional Ethics Research Committee. Results will be disseminated in peer reviewed publications and presented at national and international conferences.

Discussion: Cross-sectional studies which combine detailed personal information with biological data offer new and exciting opportunities to study gene-environmental interactions in the etiology of common mental disorders in representative samples of the general population. A collaborative multidisciplinary research approach offers the potential to advance our knowledge of the underlying complex interactions and this opens the field for further innovative study designs in psychiatric epidemiology.

KEYWORDS: Cross-sectional survey, mental disorders, prevalence, gene-environmental interactions, genome, epigenome, transcriptome.

ARTICLE SUMMARY

Article focus

- Study protocol of the PEGASUS-Murcia project, a new cross-sectional face-to-face interview survey based on a representative sample of non-institutionalized adults in the Region of Murcia (Mediterranean Southeast, Spain).
- The first objective is to estimate the prevalence of the most common mental disorders in general population, analyzing the association with sociodemographic factors, quality of life, treatment, use of services, unmet need and quality of care received and comparing the results with those obtained from Spain, Europe and other non-European countries.
- The second objective it to study the genetic, epigenetic and transcriptomic influences associated with mental disorders.

Key messages

- The study of the complex interactions between environmental and genetic risk and protective factors involved in mental disorders is better approached by a multidisciplinary research team.

Strengths and limitation of this study

- The major strength of this protocol is the assessment of environmental and genetic factors not only associated to mental disorder but also with positive mental health in a representative sample of the general population by a multidisciplinary research team.
- The limitation of this protocol is that its cross-sectional design which, while it allows association studies and the generation of new hypotheses, limits the possible causal interpretation of the findings.

BACKGROUND

The World Mental Health (WMH) Survey Initiative is a WHO (World Health Organization) initiative specifically designed to carry out epidemiological surveys in a representative number of countries in all major regions of the world.[1-3] All previous WMH surveys have used or are currently using the same diagnostic interview, the WHO Composite International Diagnostic Interview (WMH-CIDI, hereafter referred to as CIDI), a fully-structured research diagnostic interview questionnaire designed to be used by trained lay interviewers without clinical experience. This initiative has generated an enormous body of comparative cross-national data on the epidemiology of mental disorders all over the world.[3-7] As part of it, the European Study of the Epidemiology of Mental Disorders (ESEMeD) project was designed to collect data from representative samples of the adult population in six European countries: Belgium, France, Germany, Italy, the Netherlands and Spain.[2,8,9] It has also generated a large number of scientific papers on the most prevalent mental health disorders (mood, anxiety, and alcohol abuse) in Europe.[10-17] There is a general consensus on the importance of the ESEMeD project in terms of improving scientific knowledge of the epidemiology of mental disorders in Europe.[1,2,9]

Genes and environment factors in the etiology of mental disorders.

Despite decades of intensive research, it remains difficult to identify specific genes and to characterize those environmental factors primarily responsible for mental disorders.[18-22] The concept of genes and environmental factors as independent causes of mental disorders has been replaced by one of complex interactions between them. These Gene-Environment (GxE) interactions imply a genetic predisposition of some subjects to be expressed differently depending on the environment to which they are exposed.[23,24] For example, the important role of environmental factors, especially stressful life events (SLEs), is now widely accepted. Exposure to various SLEs (work or physical problems, assault, natural disasters, etc.), separately or cumulatively over the life of an individual, increases the risk of depression although in only a proportion of those exposed.[25,26] These data suggest the existence of genetic differences which might explain individual variation in the sensitivity of people to the depressogenic effects of

SLEs. On the other hand, the serotonin transporter (*SERT* or *5HTT*) gene, a key regulator of serotonergic neurotransmission and one of the most studied genetic polymorphisms in relation to affective disorders,[27] has been associated with depression,[28,29] neuroticism[30] and posttraumatic stress disorder (PTSD).[31] However, these findings have not always been replicated.[32-34]

These inconsistencies may be explained by, at least, three different factors. Firstly, in adults, higher levels of neuroticism are associated with an increased risk of depression,[35] anxiety[36] and PTSD after exposure to a traumatic even[37] and are a powerful predictor of comorbidity between depression and anxiety.[38] Neuroticism includes those personality traits that represent how some people perceive the world around them as threatening or stressful. In addition, some personality traits also influence the individual tendency to be potentially exposed to stressful environments. Predisposed individuals may tend to choose environments prone to having a high risk of exposure to stressful events. Specifically, this scenario, known as GxE correlation, may mediate the relationship between neuroticism and specific SLEs.[39] Secondly, the genetic factors influencing the level of neuroticism, including the *5-HTTLPR* polymorphism, are shared by persons having anxious-depressive spectrum disorders.[38,40] Lastly, GxE interactions have been described involving *5-HTTLPR* and depression,[41] anxiety[42] and PTSD.[43] Despite all of the above evidence, genetic association and GxE interaction studies do not usually analyze or control for the level of neuroticism in the relationship between *5-HTTLPR*, SLEs and anxious-depressive spectrum disorders.

However, the question arising in this context is how environmental and genetic factors interact to produce a mental disorder.[21,44] In recent years, increasing interest in epigenetic factors described in other human diseases has focused on its role in mental disorders.[45] The study of the epigenome, changes in gene expression by modulating the accessibility of information that occurs without modifying the DNA sequence, suggests that, although inheritable, these changes are not necessarily stable over the life span of individuals and can be modified under some environmental stimuli that modulate the activity of the enzymes involved, opening new prospects for developing therapeutic approaches based on epigenetic mechanisms.[46] Epigenetic mechanisms have been associated with different mental disorders including depression,[47]

PTSD,[48] schizophrenia,[49,50] autism,[49] bipolar disorder[50] and alcohol dependence.[51] In fact, epigenetic regulation of the glucocorticoid receptor signaling in neurons has been recently shown to be the mechanism underlying GxE interactions to explain risk and resilience of PTSD after SLE in childhood.[52]

In order to integrate all these findings and create new opportunities and challenges offered by the GxE interaction scenarios in the field of mental disorders, a multidisciplinary collaboration between clinicians, epidemiologists, geneticists and statisticians offers greater opportunities.[20,23,53] One of the proposed mechanisms for this collaboration includes carrying out community psychiatric surveys and this has been facilitated by the possibility of obtaining DNA and/or mRNA from peripheral tissues (blood, saliva or buccal cells). Population-based surveys offer several advantages over other study designs to contribute to the clarification of the GxE interactions in mental disorders.[44,54,55] Firstly, current knowledge of genes as risk factors is based almost exclusively on clinical and non-representative population samples. Secondly, the distribution of the gene polymorphisms of interest in the general population has not been well investigated. Thirdly, this type of study can provide samples for future case-control studies and can be the bases for future longitudinal ones. Finally, hypotheses generated from epidemiologic surveys may contribute to test new basic studies and can be considered as a complementary strategy to translational research.

Pegasus-Murcia Project

Spain actively participated in the ESEMeD Project with a representative sample of the adult general Spanish population (n=5473) and the results have been published in national and international journals.[56-61] However, the sample size within most of the Autonomous Communities in Spain was too small to be able to achieve accurate and precise estimates at the Regional level where Health Care policies are decided. Moreover, several differences between the Autonomous Communities in Spain in important aspects related to mental health such as socioeconomic[62] and territorial inequalities in health care supply and in long-term care, access to and use of health care facilities,[63], premature deaths due to alcohol consumption[64] and the prevalence of psychological distress[65] have recently been described.

Murcia is one of the 17 Autonomous Communities of Spain. It is a located in the southeast of the country on the Mediterranean coast, with a population of 1,424,063 inhabitants at the time of the survey (INE 2008, National Statistical Institute of Spain), almost a third of them (30.7%) living in the capital.

The PEGASUS-Murcia ("Psychiatric Enquiry to General Population in Southeast Spain-Murcia") project has been designed in order to obtain regional data of the prevalence, burden and care of a representative sample of the general adult population of Murcia to allow planning of new regional mental health policies and to compare the results with the national sample of Spain, Europe and all other countries participating in the WMH Survey Initiative. The project also constitutes a unique opportunity to initiate a biological bank of a well-studied representative sample of the general population.

Objectives

The PEGASUS-Murcia project is a multi-purpose, observational, cross-sectional, comparative study of the non-institutionalized general population of Murcia Region whose objective is to improve knowledge about common psychiatric disorders in two main areas. The first is the epidemiology of mental disorders and protective and risk factors in the general population of Murcia. The specific objectives are: i) to estimate the one-month, 12-month and lifetime prevalence of the most common mental disorders, specifically, mood and anxiety disorders, in the general population of Murcia; ii) to assess the independent association of mood and anxiety disorders with sociodemographic factors (gender, age, education and urban/rural location) and selected risk factors (family history, childhood experiences, religion, partnership status and sexual problems, among others); iii) to assess the quality of life of persons with the most common psychiatric disorders and to analyze how other variables (physical medical conditions and sociodemographic factors) may influence this outcome; iv) to assess the treatment for these disorders and to evaluate the unmet need and the quality of care received; and v) to compare our results with those obtained from Spain, Europe and other non-European countries, including the United States.

The second objective is the genetic, epigenetic and transcriptomic influences associated with mental disorders. Its specific aims include: i) the estimation of the distribution of different candidate genes in the general population and their association with different psychiatric disorders; ii) the identification of sensitive alleles underlying potential GxE interactions and the study of epigenetic mechanisms involved, specifically, DNA methylation; and iii) the analysis of gene expression alterations through transcriptomic assays.

METHODS AND ANALYSIS

Study Design

The project is a cross-sectional face-to-face interview survey based on a representative sample of the adult and non-institutionalized general population of the Murcia Region. Those who complete the interview will be invited to provide two biological samples from their oral mucous membranes. The target population is defined as persons aged 18 or older residing in Murcia, not living in institutions and with an active health card (defined as persons included in PERSAN, a regional registry that contains all residents with a Health Card which is periodically up-dated. Exclusion criteria are: i) Confirmed irretrievable contact errors (e.g. telephone number and/or address); ii) Institutionalized individuals (e.g. in prison, in a hospital or in another institution) or those living outside the Autonomous Community during the survey field work; and iii) individuals not able to understand the Spanish language or not able to conduct the questionnaire due to his/her physical or mental condition.

Sampling plan

The geographical area of the survey is the Murcia Region and a two-stage, stratified sampling design has been used. The primary sampling unit is the Primary Health Centre and the second is the individual. The sampling frame has been PERSAN, the regional health care population database in Murcia. Primary Health Centres have been grouped into nine strata, the current Health Care Areas in Murcia Region. The initial sample size was 4,500 adult individuals divided into the nine Health Care Areas with proportionate allocation. A representative sample of two centres has been chosen in each health area, without individual participant replacement.

Selection probability for each centre was known a priori and it was proportional to the size of the centre (% of adult individuals registered in the centre) and the proportion of adult individuals in the centre whose place of residence was rural, semi-urban or urban. Within each of the two selected Health Centers, a stratified random sample procedure, performed for each combination of gender, age group (18-24, 25-34, 35-49, 50-64 and 65+) and type of residence (rural, semi-urban and urban), constitutes a stratum and individuals have been selected using simple random sampling.

For each Health Care Area, the sample size of each stratum has been selected such that individuals in with the same demographic characteristics had equal probability of being selected independently of the selected center. If a high number of those fulfilling the exclusion criteria in one area is reached, a fixed number of additional individuals will be released (subsequent releases), according to the number of interviews completed in the area and following the same selection procedure within each of the centers as the ones used to select the initial release (no new center will be selected for these releases). Any replacement of those persons who do not want to collaborate or who do not meet the non-eligibility criteria is not allowed.

Survey procedures and data control

Those selected will receive no financial incentive to participate and there will be no individual replacement procedure. Questions are asked by trained lay interviewers using the computer-assisted personal interview (CAPI) that was programmed centrally using the Blaise software system. This is an interviewing application developed by Statistics Netherlands (Herleen, the Netherlands) and designed to ease the handling of elaborate skip and complex randomization patterns and to facilitate data entry, allow the elaboration of some questions and direct the interviewer through the questioning sequence.

Periodically, the completed interviews will be submitted to the central project Data Center (Regional Mental Health Service, Murcia-Spain) for checking and storage following a predetermined security procedure. All raw data will be transferred to the Hospital del Mar Medical Research Institute (IMIM) and the Department of Health Care Policy at Harvard University, coordinating centers of the ESEMeD and WMH Survey Initiative projects

respectively, via secure websites. The database has been declared to the Spanish Data Protection Agency.

A survey firm has been contracted to undertake the fieldwork and, in order to ensure the quality of the survey, several strategies are being implemented: i) a one week training course for all interviewers by WHO certified trainers on the original protocol and use of the CAPI version of the CIDI; ii) development of a written manual to standardize the interviewing procedure and all scientific and administrative elements that could affect comparability of data; iii) regular meetings with the survey firm to ensure adherence to the protocol and to deal with any difficulty that may have arisen; and iv) data quality analysis to detect any inconsistencies and/or incomplete data.

The survey firm has been provided with sufficient data to allow contact with each of the individuals of the selected sample and only after 10 unsuccessful attempts the person will be considered to be uncontactable or after confirmation that the selected person does not live at that address and new contact information is unavailable. Several methods will be used to improve the participation of those selected: i) an informative flyer providing general information related to the project and giving notice of future contact will be sent by conventional post together with an invitation letter signed by a person from the Health Care Authority; ii) a phone call to invite them to participate in the interview process and to offer them the possibility to do the interview either at home or in their Primary Care Center; iii) Several informative sessions for the healthcare personnel of the Primary Care Centers will be organized to facilitate their collaboration should the participants ask them about the project; iv) During the period when the interviews will take place, some official posters will be put in public centers to inform people about the project; v) All interviewers will be provided with an official identification and have been trained on how to explain the institutional nature of the research project.

The survey questionnaire

The questionnaire used in the PEGASUS-Murcia project is a revised version of the CIDI which, together with diagnostic information on the most common mental disorders, also includes

specific information on the severity of the disorders, symptoms, disability, quality of life, use of services and medication and several risk factors.

The Composite International Diagnostic Interview (CIDI)

The CIDI is a comprehensive, highly-structured interview specifically designed by the World Health Organization (WHO) for the purpose of ascertaining diagnoses of mental illnesses based on the WHO International Classification of Disease (ICD-10) and not exclusively on DSM definitions and criteria. This objective is particularly important for cross-national comparative research of the epidemiology of mental illnesses throughout the entire world [66]. It comprises nearly 5000 questions divided into 42 sections (Table 1) and these, in turn, are grouped into two main parts: diagnostic and other. The first includes the clinical part of the interview with an introductory screening section and 22 diagnostic sections that assess different psychiatric conditions. The second includes various non-clinical sections which assess utilization of services, use of psychotropic drugs, degree of functioning in several aspects, chronic physical conditions, risk factors, social networks, caregiver burden and socio-demographic variables.

Table 1: Description of the adapted version of the World Health Organization -Composite International Diagnostic Interview (WHO-CIDI) used in the PEGASUS-Murcia project

Sections	Module	Number of Items	Rules for administration *
Household Listing	Methodological	5	All respondents
Screening (SCR)	Screening	51	All respondents
Minimental State Examination	Risk Factors		If older than 60 years old
Quality/Lie subscale	Functioning and physical Disorder	24	Random assignment to the beginning of the questionnaire or at the end
Depression	Mood disorder	189	Screening questions (SCR)
Mania	Mood disorder	95	Screening questions (SCR)
Panic Disorder	Anxiety	106	Screening questions (SCR)
Specific Phobia	Anxiety	143	Screening questions (SCR)
Social Phobia	Anxiety	85	Screening questions (SCR)
Agoraphobia	Anxiety	84	Screening questions (SCR)
General Anxiety Disorder	Anxiety	116	Screening questions (SCR)
Suicidality	Other Diagnostic	46	All respondents
Use of Services	Treatment	243	All respondents
Group of Questions (Tobacco and physical exercise)	Risk/Protective Factors	22 to 32	All respondents
Pharmacoepidemiology	Treatment	241	All respondents
Substances	Substance abuse	182	Long path
Post-Traumatic Stress Disorder	Anxiety	464 to 491	Long path
Chronic Conditions	Functioning and physical	201	Long path

	Disorder		
30 Days Functions	Functioning and physical Disorder	75	Long path
30 Days Symptoms	Functioning and physical Disorder	75	Long path
Eating Disorders	Other Diagnostic	80	50% of Long path
Obsessive-Compulsive Disorder	Anxiety	124	33% of Long path
CAPE[‡]	Psychosis	42 to 84	All respondents
CFQ[β]	Risk Factors	25	All respondents
SLE[€]	Risk Factors	13 to 39	All respondents
Neuroticism and Extroversion subscales[¥]	Risk/Protective Factors	12	All respondents
Resilience Scale	Protective Factors	25	All respondents
Employment	Socio-demographics	121	Long path
Finances	Socio-demographics	21	Long path
Marriage	Socio-demographics	91	All respondents
Partner violence	Risk Factors	2 to 15	All respondents
Children	Socio-demographics	44	Long path
Social Networks	Risk/Protective Factors	16	All respondents
Adult Demographics	Socio-demographics	68	Long path
Child Demographics	Socio-demographics	34	Long path
Demographic Short	Socio-demographics	25-36	Long path
Childhood	Risk/Protective Factors	110	Long path
Attention Hyperactivity	Childhood	90	Long path and Screening
Oppositional Defiant	Childhood	46	Long path and Screening
Conduct Disorder	Childhood	54	Long path
Separation Anxiety Disorder	Childhood	86	Screening questions (SCR)
Family Burden	Risk Factors	40	Long path
Quality/ Lie subscale	Functioning and physical Disorder	26	Random assignment to the beginning of the questionnaire or at the end
Respondent Contacts	Methodological	19	All respondents
Interviewer Observation	Methodological	14	All respondents

[$] EQ-5D: European Quality of Life Scale; [&] SF-12 v2: Short Form 12 Health Questionnaire; [†] Lie subscale of the abbreviated version of the Eysenck Personality Questionnaire (EPQR-A); [‡] CAPE: Community Assessment of Psychic Experiences; [β] CFQ: Cognitive Failure Questionnaire; [€] SLE: Stressful Life Events; [¥] Neuroticism and Extroversion subscales of the abbreviate version of the Eysenck Personality Questionnaire (EPQR-A)

* Long Path inclusion criteria: a) all individuals that could be considered as "high risk individuals", because they had positively answered a number of specific questions related to mood and anxiety disorders, and b) a random subsample (25%) of the respondents without symptoms ("low risk individuals"). The remaining 75% of respondents without screening symptoms not randomly selected for the long path followed the Short Path of the questionnaire

The most recent version of the CIDI (version 3.0) is the end result of a number of international studies and adaptations made since 2000 when it was first used in WMH surveys. It was first created in English and has been translated into more than 30 different languages using the standard WHO protocol with a rigorous process of adaptation.[67,68] Several clinical reappraisal studies have been carried out and the concordance of the CIDI version 3.0 has been evaluated in different subgroups of WMH surveys using the Structured Clinical Interview for DSM-IV (SCID) as the clinical gold standard and a moderate to excellent concordance has been found for most mental disorders.[69,70] CIDI is available in two formats: the paper form or PAPI (Paper and

Pencil Interviewing) and the computerized form or CAPI (Computer Assisted Personal Interviewing), designed to ease the handling of elaborate skip and complex randomization patterns and to facilitate data entry with a resulting reduction in interview time and errors in data collection and recording. The original Spanish CAPI version used in Spain had not been updated since it was used in the context of the ESEMeD project almost ten years ago. Since then, all improvements in the questionnaire have only been added to the CIDI Latin American (LA) v20.0 version. However, due to linguistic and cultural differences in Spanish-speaking populations, this CAPI version had to be culturally adapted for use in Spain by our research team and this process is fully described elsewhere.[71]

To further shorten the length of the questionnaire, some sections were not selected for the purposes of this project. These include Intermittent Explosive Disorder, Personality I and II, Neurasthenia and Pre-Menstrual and Gambling sections. Some others were substituted by other questions or questionnaires, e.g. the Tobacco Use section was simplified using some questions obtained from the Spanish National Health Survey and the Psychosis section with the CAPE instrument (Community Assessment of Psychic Experiences), both described below.

Other study instruments

Several other instruments were added to the original CIDI for the specific purposes of the PEGASUS-Murcia project. These include the Spanish version of different questionnaires: i) Mini-Mental State Examination for interviewees older than 60 years old;[72,73] ii) the Cognitive Failure Questionnaire (CFQ);[74,75] iii) the Neuroticism, Extroversion and Lie subscales of the abbreviated version of the Eysenck Personality Questionnaire (EPQR-A);[76-78] iv) the Resilience Scale;[79,80] v) the Community Assessment of Psychic Experiences (CAPE)[81] to measure attenuated psychotic symptoms in the general population instead of the Psychosis section of the CIDI, as the latest is only used as a screening instrument in the detection of psychosis. Those who positively answer two items of the positive dimension with a score equal or superior to 3, have been hospitalized for psychiatric reasons and/or have received psychotropic medication during the last year will be evaluated by a clinic psychiatrist with the module C (Psychotic Disorders) of the SCID (Structured Clinical Interview for DSM Disorders) ; vi) a brief list of 12

stressful life events in the last 12 months was included by the combination of a List of Threatening Experiences (LTE)[82,83] and the emotional and life-changing impact of each event;[84] vii) the European Quality of Life Scale (EuroQol 5D)[85] and the Short Form 12 Health Questionnaire (SF-12 v2);[86] viii) an ad-hoc questionnaire of partner violence obtained from the Spanish National Health Survey and from the regional mental health clinical guidelines;[87] and, finally, ix) some questions related to tobacco use and physical exercises from the Spanish National Health Survey.

Questionnaire pathways

In order to optimize the duration of the interview, the WMH questionnaire was divided into two parts with questions in Part 1 administered to all respondents and those in Part 2 only to a subsample of individuals who followed the long path of the interview. Part 2 of the interview includes detailed information about a wide range of aspects related to the primary disorders and also to mental disorders of secondary interest (Table 1). The inclusion criteria for the long path are: a) all individuals that could be considered as "high risk individuals" because they positively answer a number of specific questions related to mood and anxiety disorders and b) a random subsample (25%) of the respondent without symptoms ("low risk individuals"). The remaining 75% of respondents without screening symptoms not randomly selected for the long path followed the short path. All these pathways are automatically made by the computer without any intervention of the interviewer. In this short itinerary, the sections omitted were substituted by a specific section that included those questions needed to calculate some demographic indicators. Moreover, two sections were only used in a percentage of the long path itinerary, Eating Disorders (50 %) and Obsessive-Compulsive Disorder (33%).

Quality control procedures

Data quality will be controlled in a number of ways to ensure that the predetermined protocol has been followed achieving the greatest reliability and validity and these quality control procedures will be organized and supervised by members of the coordinating centers. The principal investigator will reviewed all responses to open-ended questions to check if narratives excludes a clinical diagnosis of mental disorders, i.e., whether symptoms were due to a physical

illness. All these procedures will be verified by the coordinating centers and the final document included several aspects, for example, sample releases, the duration of the interviews and the proportion of positive responses to selected screening questions. Local members of the research team will be responsible for verifying the informed consent forms and the quality checking following computerized protocols. These procedures are similar to those implemented in the ESEMeD project and are fully described elsewhere.[8] Briefly, they consist of checks of individual pieces of information from the interviewees, for example, completion status, consistency across the questionnaire, questionnaire itinerary and length of the interview, and from the interviewers, number of disorders screened positively, verification of a random selection of almost 1% of interviews completed by a telephone contact to confirm the interview and some aspects related to it such as place, approximate duration and identification of the interviewer.

Laboratory Methods

On completion of the interview, two biological samples will be obtained from each interviewee, one for DNA extraction for genomic and epigenomic analysis and the other one to obtain mRNA for gene expression quantification (transcriptomic assays). These samples will be taken using swabs compatible with molecular amplification techniques, as they do not interfere with the amplification process (FLOQSwabs Flocked Swabs, Copan Flock Technologies srl).

Samples for DNA extraction will be collected in sterile 1.5 ml tubes. Those to be used for RNA extraction will be harvested in dark sterile tubes containing RNA protect cell Reagent (QIAGEN, Hilden, Germany), which provides immediate stabilization of RNA. Cells will be thus stabilized at room temperature and can then be stored or transported at ambient temperature prior to RNA purification. Tubes will be labeled with tags (14C.B. 40X40 type) with a specific code for each sample and will be packaged and sent to BIOBANC-Mur (the biobank for biomedical research network of the Region of Murcia, RD09/0076/00065, as a partner of the Spanish National Biobanks Network; IMIB: Instituto Murciano de Investigación Biosanitaria) according to current Spanish legislation and following the regulations of the International Air Transport Association (IATA) on biological sample shipping.

Those sample accepted by BIOBANC-Mur will be registered using a specific biobanking software (bio-e-bank, VITROSOFT, SL), as part of a Laboratory Integrated Management System (LIMS). The nucleic acid extraction will be performed automatically (QIAcube system; QIAGEN, Hilden, Germany) to minimize variability due to manual handling using QIAamp DNA Blood Mini Kit and RNeasyPlus Mini Kit (QIAGEN, Hilden, Germany) for DNA and RNA extraction, respectively. The synthesis of complementary DNA (cDNA) from mRNA for expression studies will be developed for all samples by reverse transcription using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). All processes will be performed according to the manufacturer's instructions.

DNA and RNA will be quantitated by measuring absorbance at 260/280 nm using a spectrophotometer.[88-90] Between 260 nm and 230 nm (A260/230) absorbance is commonly used as a secondary indicator of nucleic acid purity[91-93]. The integrity of DNA will be visualized by electrophoresis on 1% agarose gel (migration for 1 hour at 100 V) using 100 ng of total DNA and a 23 kb DNA ladder (Lambda DNA/HindIII Marker (Thermo Fisher Scientific) as DNA marker. All mRNA samples will be transformed into cDNA.

Specially trained technicians will be used to monitor the specimen collection by donors and to perform sample manipulations in order to minimize variability of results and to obtain the optimal quality of nucleic acids for this and future studies. The processed biospecimens (150 µl of DNA and 80 µl of cDNA) will be stored in 750 µl microtubes in an ultra-freezer at -80 °C located in BIOBANC-Mur.

Statistical methods

The expected response-rate (RR) has been set to a minimum of 65%, based on a previous regional community survey which included the donation of blood samples [94,95]. The response rate will be calculated based on the proportion of people interviewed and was defined as the number of completed interviews divided by the total number of cases minus the number of non-eligible cases.

Weighting procedures

Given that the interview is divided into two parts and only a portion of the sample will be selected for the second part, two types of weightings are considered to estimate population parameters. The first is to weight for the probability of selection for each Health Care Area, Health Center and demographic stratum and the second is for the random skips included in the questionnaire. The method designed is described in Box 1.

BOX 1: Weighting procedures

First weighting procedure:

Step 1) For each Healthcare Area h, health centre c and demographic stratum (sex, age group and type of residence), all individuals have sampling weight $w_s=1/p_{hc}p^1_{hcsgr}$, where p_{hc} is the probability that the centre c was selected, $p^1_{hcsgr} = n_{hcsgr} / N_{hcsgr}$ and n_{hcsgr} is the sample size for the demographic stratum with N_{hcsgr} individuals registered in the sampling frame.

Step 2) Non-response weight (w_{nr}): if p^*_{hcsgr} is the proportion of eligible persons that is actually interviewed in the Healthcare Area h, centre c, sex s, age group g and type of residence r, the non-response weight of the persons in the Healthcare Area h, centre c, sex s, age group g and type of residence r is $w_{nr} = 1/ p^*_{hcsgr}$.

Step 3) Unadjusted weight (w_{unadj}): it was calculated as the product of sampling weight by non-response weight: $w_{unadj} = w_s w_{nr}$.

Step 4) Post-stratification weight (w_{ps}): data on population of the region of Murcia by sex, age and Healthcare Area were provided by the CREM (Centro Regional de Estadística de Murcia; Padrón 2010 (http://www.carm.es/econet/sicrem/PU_padron/). The population for the age group 18-24 has been estimated as the population for the age group 18-19 plus the population for the age group 20-24. The population for the age group 18-19 has been estimated as the population for the age group 15-19 times the proportion of population aged 18-19 in the age group 15-19 in Murcia: 0.4116 for males and 0.4165 for females. A post-stratification weight was created to ensure that the joint distribution of the post-stratifying variables Healthcare Area, sex and age group matches the known population joint distribution of Murcia.

Step 5) Adjusted weight (w_{adj}): the adjusted weight of an individual in the Healthcare Area h, centre c, sex s, age group g and type of residence r is $w_{adj} = w_{unadj} w_{psk}$..

Step 6) Normalized weight: $w_{norm} = w_{adj}\, n / \sum_{i=1}^{n} w_{adj_i}$.

Step 7) Trimmed weight (w_{trim}) : trim the normalized weight obtained from step 6. The upper and lower 5% were trimmed to the mean of each tail.

Step 8) Normalized trimmed weight: $w = w_{trim}\, n / \sum_{i=1}^{n} w_{trim_i}$.

Second weighting procedure:

To take into account the random skips in the CIDI questionnaire applied to define the long path we calculated the skip pattern weights. Only a portion of the sample completed the second part (Part 2) of the survey. The probability of inclusion into Part 2 is based on the presence or absence of disorder symptoms as defined in the interview schedule. Again, different steps will be followed:

Step 1) Part 2 selection weight (w_{p2s})**:** each individual i in the sample that accepted to respond the first part of the survey were selected into Part 2 with probability π_i where π_i = 1 for high risk individuals of having mental disorders and π_i = 0.25 for the rest. Then the Part 2 selection weight of individual i is $w_{p2s} = 1/ \pi_i$.

Step 2) Unadjusted part 2 weight ($w_{p2unadj}$): the product of w_{trim} (Part 1) and the Part 2 selection weights.

Step 3) Part 2 post-stratification weight (w_{p2psk}): similar to the previous post-stratification procedure, a post-stratification weight was created to ensure that the joint distribution of the variables Healthcare Area, sex and age group in Part 2 match the known population distribution of Murcia.

Step 4) Part 2 adjusted weight (w_{p2adj}): the adjusted weight of an individual i in the Healthcare Area h, centre c, sex s, age group g and type of residence r is $w_{p2adj} = w_{p2unadj} w_{p2psk}$.

Step 5) Part 2 Normalized weight: $w_{P2norm} = w_{P2adj}\, n / \sum_{i=1}^{n} w_{P2adj_i}$.

Analysis of the data and forthcoming research projects

There are three data analysis centres in the project: Harvard University (Boston, USA), IMIM (Barcelona, Spain) and the Regional Centers of Epidemiology and Mental Health (Murcia, Spain). Harvard will supervise all quality procedures and provides consultancy in many aspects of the analysis, including the sampling design, the weighting procedures and the verification of the CIDI diagnostic algorithms. All the analyses will be performed using SAS™ and SPSS programs.

Related to this research project, several other lines of research with different designs are being developed, for example, case-control studies and meta-analyses. An example of the former is a case-control study of the GxE interactions, involving *5-HTTLPR* polymorphisms, designed to analyse the impact of an earthquake in the mental health of the population exposed have been recently been granted. Cases will be those people with a diagnostic of affective and/or anxiety disorder exposed to the earthquake attended in the Mental Health Care center and controls will be obtained from those exposed to the earthquake that are going to be interviewed in the PEGASUS-Murcia project and without a diagnosis of any affective and/or anxiety disorder. Recently, our research team has published a meta-analysis of the relationship between *5-HTTLPR* polymorphism and PTSD.[34]

ETHICS AND DISSEMINATION

Eligible individuals will be asked to sign two independent informed consents to participate, the first one to be interviewed, including the possibility of future new contacts and the second to provide the biological samples but only those who had already completed the questionnaire. Name and contact information will be stored separately from any information provided as part of the study questionnaire. The protocol was approved by the Clinical Research Ethics Committee of the University Hospital *Virgen de la Arrixaca* of Murcia and the database of personal information was registered with the National Data Protection Agency. Data from PEGASUS-Murcia project will be included in the WMH Cross National Sample for

international comparisons. The study findings will be submitted to peer-reviewed journals for publication, and presented at national and international scientific meetings.

DISCUSSION

The epidemiology of mental illnesses is a fascinating but highly complex area of research. This complexity is primarily due to the wide range of factors, environmental and genetic, which combines to produce a recognized psychiatric disorder. Previous epidemiological research has resulted in the production of a great amount of data but it has been difficult to make cross-national comparisons due to methodological variability. The WMH Survey Initiative aimed to address this issue by using an international standardized protocol, allowing comparisons of the most common mental disorders and their associated factors throughout the world. Using this study design, it therefore offers the opportunity for new surveys to be performed in the context of an international collaborative initiative and the possibility to adapt the questionnaire according to the specific aims of the research being undertaken. The PEGASUS-Murcia project can be considered as an example of how the latter has been successfully achieved. It is a cross-sectional study designed to assess the prevalence of the most frequent mental disorders and their correlates in a representative sample of the general population of Murcia. Its primary strengths are: i) the fact that it was specifically adapted to assess factors not only associated with mental disorders but also with positive mental health in a representative sample of the general population; ii) its context focused on regional needs where healthcare decisions are taken regarding resource allocation and mental health planning; iii) the collection of biological samples not only for DNA analysis but also for mRNA; iv) all the information collected in our study, including biological samples, can be correlated with past and future health events because all Spanish population had free access to the Healthcare System at the time of its inception and were thus registered and provided with a unique identification number and therefore; v) finally, the inclusion of a multidisciplinary research team is in accordance with the international consensus regarding the need for interdisciplinary collaboration between clinicians, epidemiologists and neuroscience researchers to increase their combined efforts to study the

complex gene-gene and gene-environmental interactions underlying mental health disorders.[23,55,96,97]

Concerns have been expressed about the cost-effectiveness of psychiatric epidemiological surveys, such as World Mental Health 2000 (WMH-2000) projects,[98] an example being the rationale for starting a new psychiatric epidemiological survey in the Autonomous Community of Murcia if Spain had already participated in the ESEMeD project. However, there are several reasons to justify this regional initiative. Firstly, public health and healthcare agencies usually allocate mental health resources, including human, based on data from national epidemiologic surveys,[99] such as that provided by the Spanish participation in the ESEMeD Project. As previously mentioned, the involvement of the Region of Murcia in the Spanish ESEMeD survey did not allow evaluation of specific regional data. Nowadays, the main responsibility for planning and management of Healthcare resources in Spain lies with the Autonomous Communities and differences exist between them in terms of accessibility, amount of health-care resources and political decision-making.[62-65] Devolution of this responsibility to Murcia occurred in December 2001.

Secondly, the inclusion of biological data in a well-designed multidisciplinary epidemiological study offers great advantages in terms of a more global understanding of mental disorders. These are complex illnesses of the brain where social, familial, psychological and biological elements interact throughout the entire life of a person to influence his/her risk of developing a mental health disorder. To extend our understanding of the physiopathology and epidemiology of the more common ones (mood and anxiety), it is necessary to identify genetic loci and polymorphic alleles and their distribution in the healthy and affected population whose function in determining risk for, and protection against, these conditions probably depends on gene-gene and GxE interactions. The collection of genetic material from representative samples from the general population, well described using international diagnostic instruments such as CIDI, offers new and different possibilities to evaluate candidate genes in non-biased samples and to describe their distribution in the general population that may contribute to clarification of the complexity of mental disorders.

Thirdly, our project involving a multidisciplinary research team gives new opportunities to develop different study designs that can move from descriptive to analytical epidemiology. For example, this representative sample constitutes a good source of controls for future case-control studies, where cases will be provided from the public health care clinics, and can be the starting point for future cohort studies. Our project was designed to allow for all these possibilities.

Limitations of the study

Currently, the main limitations of the PEGASUS-Murcia project are related to: i) the cross-sectional design which, while it allows association studies, limits the possible causal interpretation of the findings. However, these findings may provide new hypotheses and enable the design of new studies; ii) not all interviewees will provide biological samples and this may affect the representativeness of some mental disorders in future analyses. To determine if this will result in selection bias, we will analyze whether there are distinguishing characteristics between donors and non-donors in the distribution of mental disorders and other characteristics of the participants; iii) the population stratification in our study which will be used for future genetic association analyses is performed by using the stated ancestral origin by participants[100] instead of using genetic markers; and iv) biological samples will be obtained from oral mucosal scrapings and not from brain neurons. However, this is a general situation given the ethical issues and difficulties in obtaining neural tissues and, in any case, gene expression does not appear to be specific to neural tissue, at least in some genes that have ubiquitous expression, for example, 5-HTTLPR.[101-104]

Conclusions and Future Directions

The PEGASUS-Murcia project is a sound bases for multidisciplinary collaborative mental health research studies which will provide not only a huge amount of epidemiological information but will also offer exiting opportunities to clarify the complex interactions between genetic and environmental factors which result in a range of mental health disorders.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

FNM, MJT, GV, JA, TE, SM and CN conceived the design and supervised the whole process of the study. GV, JA and FNM have coordinated the project with the International Consortium of Psychiatric Epidemiology (ICPE). MJT, JA and CN are coordinating the epidemiologic aspects. TE, JJ and SM are responsible for the genetic aspects. MJT, DS and GV were responsible for the sampling methods. GV, GRM and DS are responsible of the implementation of the qualitative procedures and the statistical analyses. All authors read and approved the final manuscript.

Acknowledgments: We thank Carlos Giribert Muñoz, Deputy Director of Mental Health and Psychiatric Services of Murcia for his support in developing the PEGASUS-Murcia project; Inés Morán-Sánchez, Mª Luisa Pujalte and Ascensión Garriga for their contribution to the initial phases of this project; Monica Ballesta Ruíz for her contribution to the sample selection procedure; David Martínez Martínez for his contribution to the management of the software; all the collaborators from the BIOBANC-Mur (R. Martínez Marín, B. Veas-Pérez de Tudela López, A. Parra Montoya, E. Sánchez Baeza); Pedro J. Bernal for his collaboration with the PERSAN database; and, finally, to Mike Tobin for his helpful discussions and contribution during the English translation of the document.

The PEGASUS-Murcia project was supported by the Regional Health Authorities of Murcia ("Servicio Murciano de Salud and Consejería de Sanidad y Política Social") (Decreto nº 455/2009), the "Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia" (Nº Expedientes: CM0829 I and FFIDS/EMER09/14) and the "Ayudas para proyectos de Investigación en Salud –ISCIII- del Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica" (PI12/00809). None of the funders had any role in the design, analysis, interpretation of results, or preparation of this paper. The PEGASUS-Murcia project is carried out in conjunction with the WHOWMH Survey Initiative. We thank the WMH Coordinating Center staff at Harvard and Michigan Universities and, especially Professor Ron Kessler, for their assistance with the instrumentation, fieldwork and data analysis. These activities were supported by the United States National Institute of Mental Health (R01MH070884), the John D. and Catherine T. MacArthur Foundation, the Pfizer Foundation, the U.S. Public Health Service (R13-MH066849, R01-MH069864, and R01 DA016558), the Fogarty International Center (FIRCA R03- TW006481), the Pan American Health Organization, the Eli Lilly & Company Foundation, Ortho-McNeil Pharmaceutical, Inc., GlaxoSmithKline, Bristol-Myers Squibb and Shire. A complete list of WMH publications can be found at http://www.hcp.med.harvard.edu/wmh/.

Figure 1: Flow chart of the PEGASUS-Murcia project

† The response rate is defined as: ***(completed interviews) / (total released respondent sample cases –respondent nonsample cases).***

‡ **High risk individuals**: those who positively answer a number of specific questions related to mood and anxiety disorders in the screening section. **Low risk individuals**: those without symptoms related to mood and anxiety disorders in the screening section.

¥ **Long Path inclusion criteria**: a) all high risk individuals and b) a random subsample of 25% of the low risk individuals. The remaining 75% of respondents without screening symptoms not randomly selected for the long path will follow the **Short Path** of the questionnaire

References

1. Kessler RC. The global burden of anxiety and mood disorders: putting the European Study of the Epidemiology of Mental Disorders (ESEMeD) findings into perspective. *J Clin Psychiatry*. 2007;68 Suppl 2:10-9.:10-19.
2. Alonso J, Lepine JP. Overview of key data from the European Study of the Epidemiology of Mental Disorders (ESEMeD). *J Clin Psychiatry*. 2007;68 Suppl 2:3-9.
3. Kessler RC, Aguilar-Gaxiola S, Alonso J et al. The WHO World Mental Health (WMH) Surveys. *Psychiatrie (Stuttg)*. 2009;6:5-9.
4. Kessler RC. The World Health Organization International Consortium in Psychiatric Epidemiology (ICPE): initial work and future directions -- the NAPE Lecture 1998. Nordic Association for Psychiatric Epidemiology. *Acta Psychiatr Scand*. 1999;99:2-9.
5. Kessler RC, Angermeyer M, ANTHONY JC et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*. 2007;6:168-176.
6. de Graaf R, Kessler RC, Fayyad J et al. The prevalence and effects of adult attention-deficit/hyperactivity disorder (ADHD) on the performance of workers: results from the WHO World Mental Health Survey Initiative. *Occup Environ Med*. 2008;65:835-842.
7. Scott KM, von Korff M, Alonso J et al. Mental-physical co-morbidity and its relationship with disability: results from the World Mental Health Surveys. *Psychol Med*. 2009;39:33-43.
8. Alonso J, Angermeyer MC, Bernert S et al. Sampling and methods of the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;8-20.
9. Nutt DJ, Kessler RC, Alonso J et al. Consensus statement on the benefit to the community of ESEMeD (European Study of the Epidemiology of Mental Disorders) survey data on depression and anxiety. *J Clin Psychiatry*. 2007;68:42-48.

10. Alonso J, Angermeyer MC, Bernert S et al. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;21-27.

11. Alonso J, Angermeyer MC, Bernert S et al. 12-Month comorbidity patterns and associated factors in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;28-37.

12. Alonso J, Angermeyer MC, Bernert S et al. Disability and quality of life impact of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand*. 2004;420:38-46.

13. Alonso J, Angermeyer MC, Bernert S et al. Use of mental health services in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;47-54.

14. Alonso J, Buron A, Bruffaerts R et al. Association of perceived stigma and mood and anxiety disorders: results from the World Mental Health Surveys. *Acta Psychiatr Scand*. 2008;118:305-314.

15. Alonso J, Codony M, Kovess V et al. Population level of unmet need for mental healthcare in Europe. *Br J Psychiatry*. 2007;190:299-306.

16. Codony M, Alonso J, Almansa J et al. Perceived need for mental health care and service use among adults in Western Europe: results of the ESEMeD project. *Psychiatr Serv*. 2009;60:1051-1058.

17. Sevilla-Dedieu C, Kovess-Masfety V, Gilbert F et al. Mental health care and out-of-pocket expenditures in Europe: results from the ESEMeD project. *J Ment Health Policy Econ*. 2011;14:95-105.

18. Hyman SE. The millennium of mind, brain, and behavior. *Arch Gen Psychiatry*. 2000;57:88-89.

19. Kubota T, Miyake K, Hirasawa T. Epigenetic understanding of gene-environment interactions in psychiatric disorders: a new concept of clinical genetics. *Clin Epigenetics*. 2012;4:1.

20. Merikangas KR, Zhang H, Avenevoli S et al. Longitudinal trajectories of depression and anxiety in a prospective community study: the Zurich Cohort Study. *Arch Gen Psychiatry*. 2003;60:993-1000.

21. Toyokawa S, Uddin M, Koenen KC et al. How does the social environment 'get into the mind'? Epigenetics at the intersection of social and psychiatric epidemiology. *Soc Sci Med*. 2012;74:67-74.

22. Tsuang M, Bar J, Stone W et al. Gene-environment interactions in mental disorders. *World Psychiatry*. 2004;3:73-83.

23. Caspi A, Moffitt TE. Gene-environment interactions in psychiatry: joining forces with neuroscience. *Nat Rev Neurosci*. 2006;7:583-590.

24. Caspi A, Hariri AR, Holmes A et al. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *Am J Psychiatry*. 2010;167:509-527.

25. Koenen KC, Fu QJ, Ertel K et al. Common genetic liability to major depression and posttraumatic stress disorder in men. *J Affect Disord*. 2008;105:109-115.

26. Norris FH, Friedman MJ, Watson PJ et al. 60,000 disaster victims speak: Part I. An empirical review of the empirical literature, 1981-2001. *Psychiatry*. 2002;65:207-239.

27. Caspi A, Hariri AR, Holmes A et al. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *Am J Psychiatry*. 2010;167:509-527.

28. Clarke H, Flint J, Attwood AS et al. Association of the 5- HTTLPR genotype and unipolar depression: a meta-analysis. *Psychol Med*. 2010;40:1767-1778.

29. Karg K, Burmeister M, Shedden K et al. The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. *Arch Gen Psychiatry*. 2011;68:444-454.

30. Takano A, Arakawa R, Hayashi M et al. Relationship between neuroticism personality trait and serotonin transporter binding. *Biol Psychiatry*. 2007;62:588-592.

31. Koenen K, Kilpatrick D, Acierno R et al. Serotonin transporter gene and social support moderate post-traumatic stress disorder in hurricane exposed adults. *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics*. 2006;141B:742-743.

32. Risch N, Herrell R, Lehner T et al. Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis. *JAMA*. 2009;301:2462-2471.

33. Munafo MR, Durrant C, Lewis G et al. Gene X environment interactions at the serotonin transporter locus. *Biol Psychiatry*. 2009;65:211-219.

34. Navarro-Mateu F, Escámez T, Koenen KC et al. Meta-Analyses of the *5-HTTLPR* Polymorphisms and Post-Traumatic Stress Disorder. *PLOS One*. 2013;8:e66227.

35. Kendler KS, Neale MC, Kessler RC et al. A longitudinal twin study of personality and major depression in women. *Arch Gen Psychiatry*. 1993;50:853-862.

36. Kotov R, Gamez W, Schmidt F et al. Linking "big" personality traits to anxiety, depressive, and substance use disorders: a meta-analysis. *Psychol Bull*. 2010;136:768-821.

37. Wu D, Yin H, Xu S et al. Risk factors for posttraumatic stress reactions among Chinese students following exposure to a snowstorm disaster. *BMC Public Health*. 2011;11:96.

38. Hettema JM, Neale MC, Myers JM et al. A population-based twin study of the relationship between neuroticism and internalizing disorders. *Am J Psychiatry*. 2006;163:857-864.

39. Ploubidis GB, Frangou S. Neuroticism and psychological distress: to what extent is their association due to person-environment correlation? *European Psychiatry*. 2011;26:1-5.

40. Skelton K, Ressler KJ, Norrholm SD et al. PTSD and gene variants: new pathways and new thinking. *Neuropharmacology*. 2012;62:628-637.

41. Surtees PG, Wainwright NW, Willis-Owen SA et al. Social adversity, the serotonin transporter (5-HTTLPR) polymorphism and major depressive disorder. *Biol Psychiatry*. 2006;59:224-229.

42. Stein MB, Schork NJ, Gelernter J. Gene-by-environment (serotonin transporter and childhood maltreatment) interaction for anxiety sensitivity, an intermediate phenotype for anxiety disorders. *Neuropsychopharmacology*. 2008;33:312-319.

43. Koenen KC, Aiello AE, Bakshis E et al. Modification of the association between serotonin transporter genotype and risk of posttraumatic stress disorder in adults by county-level social environment. *Am J Epidemiol*. 2009;169:704-711.

44. Galea S, Uddin M, Koenen K. The urban environment and mental disorders: Epigenetic links. *Epigenetics*. 2011;6:400-404.

45. Esteller M. The necessity of a human epigenome project. *Carcinogenesis*. 2006;27:1121-1125.

46. Narayan P, Dragunow M. Pharmacology of epigenetics in brain disorders. *Br J Pharmacol*. 2010;159:285-303.

47. Olsson CA, Foley DL, Parkinson-Bates M et al. Prospects for epigenetic research within cohort studies of psychological disorder: a pilot investigation of a peripheral cell marker of epigenetic risk for depression. *Biol Psychol*. 2010;83:159-165.

48. Koenen KC, Uddin M, Chang SC et al. SLC6A4 methylation modifies the effect of the number of traumatic events on risk for posttraumatic stress disorder. *Depress Anxiety*. 2011;28:639-647.

49. Smith CL, Bolton A, Nguyen G. Genomic and Epigenomic Instability, Fragile Sites, Schizophrenia and Autism. *Current Genomics*. 2010;11:447-469.

50. Dempster EL, Pidsley R, Schalkwyk LC et al. Disease-associated epigenetic changes in monozygotic twins discordant for schizophrenia and bipolar disorder. *Hum Mol Genet*. 2011;20:4786-4796.

51. Park BY, Lee BC, Jung KH et al. Epigenetic changes of serotonin transporter in the patients with alcohol dependence: methylation of an serotonin transporter promoter CpG island. *Psychiatry Investig*. 2011;8:130-133.

52. Klengel T, Mehta D, Anacker C et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nat Neurosci*. 2013;16:33-41.

53. van Os J, Kenis G, Rutten BP. The environment and schizophrenia. *Nature*. 2010;468:203-212.

54. Merikangas KR, Risch N. Will the Genomics Revolution Revolutionize Psychiatry? *American Journal of Psychiatry*. 2003;160:625-635.

55. Weissman MM, Brown AS, Talati A. Translational epidemiology in psychiatry: linking population to clinical and basic sciences. *Arch Gen Psychiatry*. 2011;68:600-608.

56. Autonell J, Vila F, Pinto-Meza A et al. [One year prevalence of mental disorders comorbidity and associated socio-demographic risk factors in the general population of Spain. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:4-11.

57. Haro JM, Palacin C, Vilagut G et al. [Prevalence of mental disorders and associated factors: results from the ESEMeD-Spain study]. *Med Clin (Barc)*. 2006;126:445-451.

58. Codony M, Alonso J, Almansa J et al. [Mental health care use in the Spanish general populations. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:21-28.

59. Codony M, Alonso J, Almansa J et al. [Psychotropic medications use in Spain. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:29-36.

60. Gabilondo A, Rojas-Farreras S, Rodriguez A et al. Use of primary and specialized mental health care for a major depressive episode in Spain by ESEMeD respondents. *Psychiatr Serv*. 2011;62:152-161.

61. Perales J, Olaya B, Fernandez A et al. Association of childhood adversities with the first onset of mental disorders in Spain: results from the ESEMeD project. *Soc Psychiatry Psychiatr Epidemiol*. 2013;48:371-384.

62. Borrell C, Peiro R, Ramon N et al. [Socioeconomic inequalities and health plans in the Autonomous Communities of Spain]. *Gac Sanit*. 2005;19:277-285.

63. Gonzalez B, Urbanos RM, Ortega P. [Public and private supply of health services by autonomous communities in Spain]. *Gac Sanit*. 2004;18 Suppl 1:82-89.

64. Fierro I, Yanez JL, Alvarez FJ. [Premature death and potential years of life lost due to alcohol consumption in Spain and the different autonomous communities in 2004]. *Aten Primaria*. 2010;42:95-101.

65. Ricci-Cabello I, Ruiz-Perez I, Plazaola-Castano J et al. [Mental disease, existence of diagnostic, use of psychotropic medication: differences by autonomous communities under the national health survey 2006]. *Rev Esp Salud Publica*. 2010;84:29-41.

66. Kessler RC, Ustun TB. The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res*. 2004;13:93-121.

67. Alonso J, Ferrer M, Romera B et al. The European Study of the Epidemiology of Mental Disorders (ESEMeD/MHEDEA 2000) project: rationale and methods. *Int J Methods Psychiatr Res*. 2002;11:55-67.

68. Harkness J, Pennell B, Villar A et al. Translation Procedures and Translation Assessment in the World Mental Health Survey Initiative. In: Kessler RC, Bedirhan Üstün T, eds. *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders*.2008; New York: Cambridge University Press. World Health Organization.

69. Haro JM, Arbabzadeh-Bouchez S, Bugha T et al. Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health Surveys. *Int J Methods Psychiatr Res*. 2006;15:167-180.

70. Kessler RC, Abelson J, Demler O et al. Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMHCIDI). *Int J Methods Psychiatr Res*. 2004;13:122-139.

71. Navarro-Mateu F, Moran-Sanchez I, Alonso J et al. Cultural adaptation of the Latin American version of the World Health Organization Composite International Diagnostic Interview (WHO-CIDI) (v 3.0) for use in Spain. *Gac Sanit*. 2012;27:325-33.

72. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12:189-198.

73. Lobo A, Saz P, Marcos G et al. [Revalidation and standardization of the cognition mini-exam (first Spanish version of the Mini-Mental Status Examination) in the general geriatric population]. *Med Clin (Barc)*. 1999;112:767-774.

74. Broadbent DE, Cooper PF, FitzGerald P et al. The Cognitive Failures Questionnaire (CFQ) and its correlates. *Br J Clin Psychol*. 1982;21 (Pt 1):1-16.

75. García Martínez J, Sánchez Cánovas J. Adaptación al cuestionario de fallos cognitivos de Broadbent, Cooper, Fitzgerald y Parkes (CFQ, Cognitive Failure Questionnaire). *Análisis y Modificación de Conducta*. 1994;20:727-752.

76. Fusté-Escolano A, Ruíz J. Estructura factorial de la versión reducida del "Eysenck Personality Profiler". *Psicothema*. 2000;12:406-411.

77. Sandín B, Valiente RM, Olmedo M et al. Versión española del cuestionario EPQR-Abreviado (EPQR-A) (II): Replicación factorial, fiabilidad y validez. *Revista de Psicopatología y Psicología Clínica*. 2002;7:207-216.

78. Sandín B, Valiente RM, Chorot P et al. Versión española del cuestionario EPQR-Abreviado (EPQE-A) (I): Análisis exploratorio de la estructura factorial. *Revista de Psicopatología y Psicología Clínica*. 2002;7:195-205.

79. Wagnild G, Young H. Development and psychometric evaluation of the resilience scale. *J Nurs Meas*. 1993;1:165-177.

80. Heilemann M, Lee K, Kury F. Psychometric properties of the spanish version of the Resilience Scale. *J Nurs Meas*. 2003;11:61-72.

81. Stefanis NC, Hanssen M, Smirnis NK et al. Evidence that three dimensions of psychosis have a distribution in the general population. *Psychol Med*. 2002;32:347-358.

82. Brugha T, Bebbington P, Tennant C et al. The List of Threatening Experiences: a subset of 12 life event categories with considerable long-term contextual threat. *Psychol Med*. 1985;15:189-194.

83. Brugha TS, Cragg D. The List of Threatening Experiences: the reliability and validity of a brief life events questionnaire. *Acta Psychiatr Scand*. 1990;82:77-81.

84. van Os J, Park S, Jones P. Neuroticism, life events and mental health: evidence for person--environment correlation. *The British Journal of Psychiatry*. 2001;178:s72-s77.

85. Badia X, Roset M, Montserrat S et al. [The Spanish version of EuroQol: a description and its applications. European Quality of Life scale]. *Med Clin (Barc)*. 1999;112:79-85.

86. Vilagut G, Valderas JM, Ferrer M et al. [Interpretation of SF-36 and SF-12 questionnaires in Spain: physical and mental components]. *Med Clin (Barc)*. 2008;130:726-735.

87. Subdirección General de Salud Mental. *Guía práctica clínica. Actuación en salud mental con mujeres maltratadas por su pareja.* 2010; Murcia: Servicio Murciano de Salud.

88. Abitan H, Bohr H, Buchhave P. Correction to the Beer-Lambert-Bouguer law for optical absorption. *Appl Opt*. 2008;47:5354-5357.

89. Commoner B, Lipkin D. The Application of the Beer-Lambert Law to Optically Anisotropic Systems. *Science*. 1949;110:41-43.

90. Trumbo TA, Schultz E, Borland MG et al. Applied spectrophotometry: Analysis of a biochemical mixture. *Biochem Mol Biol Educ*. 2013.

91. Mulot C, Stucker I, Clavel J et al. Collection of human genomic DNA from buccal cells for genetics studies: comparison between cytobrush, mouthwash, and treated card. *J Biomed Biotechnol*. 2005;2005:291-296.

92. Rogers NL, Cole SA, Lan HC et al. New saliva DNA collection method compared to buccal cell collection techniques for epidemiological studies. *Am J Hum Biol*. 2007;19:319-326.

93. Viltrop T, Krjutskov K, Palta P et al. Comparison of DNA extraction methods for multiplex polymerase chain reaction. *Anal Biochem*. 2010;398:260-262.

94. Valverde JC, Tormo MJ, Navarro C et al. Prevalence of diabetes in Murcia (Spain): a Mediterranean area characterised by obesity. *Diabetes Res Clin Pract*. 2006;71:202-209.

95. Huerta JM, Tormo MJ, Egea-Caparros JM et al. Accuracy of self-reported diabetes, hypertension and hyperlipidemia in the adult Spanish population. DINO study findings. *Rev Esp Cardiol*. 2009;62:143-152.

96. Hyman SE. The genetics of mental illness: implications for practice. *Bull World Health Organ*. 2000;78:455-463.

97. Merikangas KR, Low NC, Hardy J. Commentary: understanding sources of complexity in chronic diseases--the importance of integration of genetics and epidemiology. *Int J Epidemiol*. 2006;35:590-592.

98. Weich S, Araya R. International and regional variation in the prevalence of common mental disorders: do we need more surveys? *Br J Psychiatry*. 2004;184:289-290.

99. Kessler RC. Psychiatric epidemiology: selected recent advances and future directions. *Bull World Health Organ*. 2000;78:464-474.

100. Mercer KB, Orcutt HK, Quinn JF et al. Acute and Posttraumatic Stress Symptoms in a Prospective Gene x Environment Study of a University Campus Shooting. *Arch Gen Psychiatry*. 2011;69:89-97.

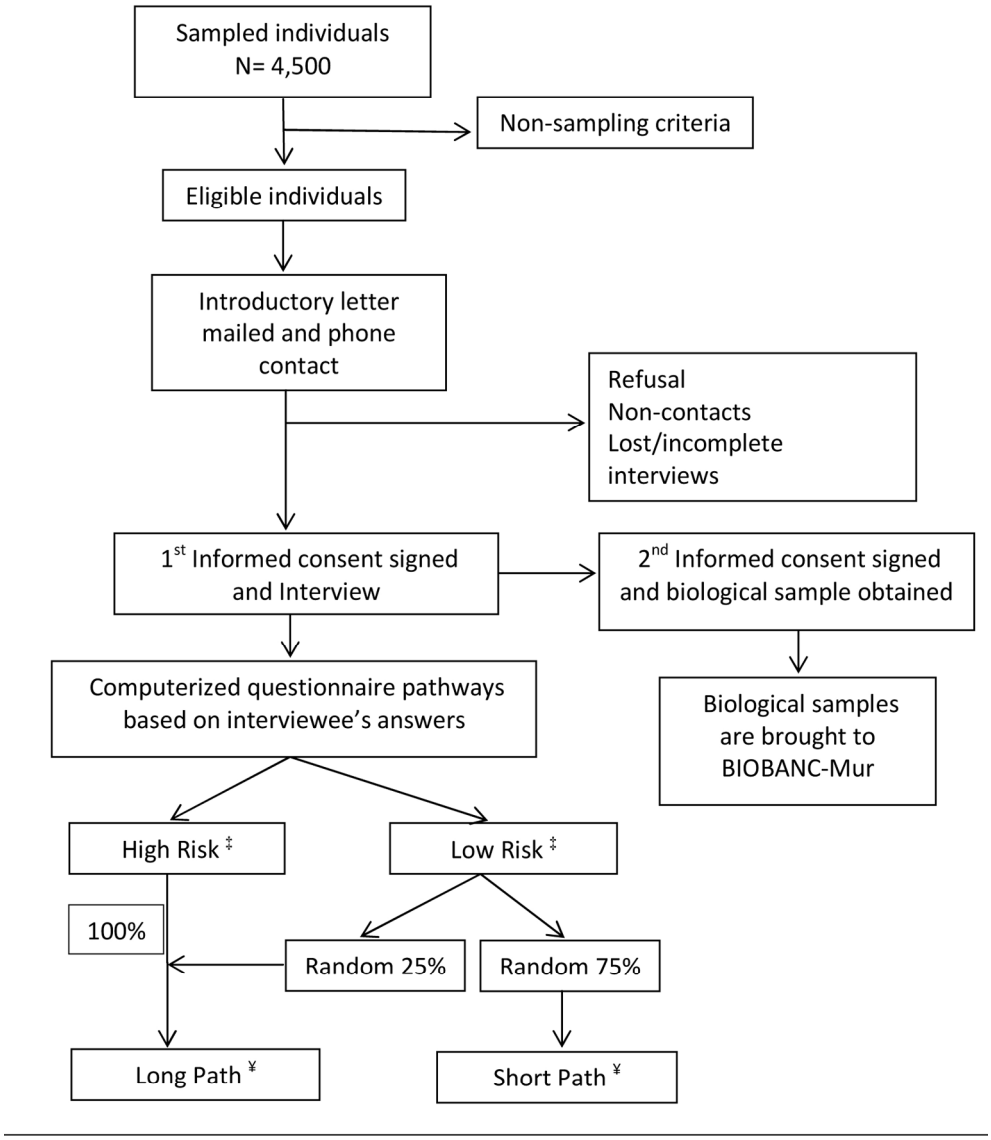
101. Tunbridge EM, Eastwood SL, Harrison PJ. Changed relative to what? Housekeeping genes and normalization strategies in human brain gene expression studies. *Biol Psychiatry*. 2011;69:173-179.

102. Abuhatzira L, Shemer R, Razin A. MeCP2 involvement in the regulation of neuronal alpha-tubulin production. *Hum Mol Genet*. 2009;18:1415-1423.

103. Bayou N, Belhadj A, Daoud H et al. Exploring the 7p22.1 chromosome as a candidate region for autism. *J Biomed Biotechnol*. 2010;2010:423894.

104. Fatjo-Vilas M, Papiol S, Estrada G et al. Dysbindin-1 gene contributes differentially to early- and adult-onset forms of functional psychosis. *Am J Med Genet B Neuropsychiatr Genet*. 2011;156B:322-333.

105. Lasken RS, Egholm M. Whole genome amplification: abundant supplies of DNA from precious samples or clinical specimens. *Trends Biotechnol*. 2003;21:531-535.



† The response rate is defined as: (completed interviews) / (total released respondent sample cases – respondent nonsample cases).

‡ High risk individuals: those who positively answer a number of specific questions related to mood and anxiety disorders in the screening section. Low risk individuals: those without symptoms related to mood and anxiety disorders in the screening section.

¥ Long Path inclusion criteria: a) all high risk individuals and b) a random subsample of 25% of the low risk individuals. The remaining 75% of respondents without screening symptoms not randomly selected for the long path will follow the Short Path of the questionnaire

273x333mm (240 x 240 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Pages
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1-2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2-3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-7
Objectives	3	State specific objectives, including any prespecified hypotheses	7-8
Methods			
Study design	4	Present key elements of study design early in the paper	8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	8-10
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	8-9
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	9-14
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*9-14*
Bias	9	Describe any efforts to address potential sources of bias	14-15
Study size	10	Explain how the study size was arrived at	8-9
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	9-14
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	16-18
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	
		(*e*) Describe any sensitivity analyses	

Continued on next page

Results			Pages
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	-
		(b) Give reasons for non-participation at each stage	-
		(c) Consider use of a flow diagram	-
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	-
		(b) Indicate number of participants with missing data for each variable of interest	-
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	-
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	-
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	-
		Cross-sectional study—Report numbers of outcome events or summary measures	-
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	-
		(*b*) Report category boundaries when continuous variables were categorized	-
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	-
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	-
Discussion			
Key results	18	Summarise key results with reference to study objectives	19-20
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	21
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	-
Generalisability	21	Discuss the generalisability (external validity) of the study results	-
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	22

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

Epidemiology and Genetics of Common Mental Disorders in the general population: the PEGASUS-Murcia project

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-004035.R1
Article Type:	Protocol
Date Submitted by the Author:	12-Nov-2013
Complete List of Authors:	Navarro-Mateu, Fernando; Unidad Docencia, Investigación y Formación en Salud Mental (UDIF-SM), Subdirección Gral. Salud Mental; CIBER de Epidemiología y Salud Pública (CIBERESP), Tormo, Mª José; Consejería de Sanidad, Servicio de Epidemiología; CIBER de Epidemiología y Salud Pública (CIBERESP), Vilagut, Gemma; IMIM-Institut Hospital del Mar d'Investigacions Médiques, ; CIBER de Epidemiología y Salud Pública (CIBERESP), Alonso, Jordi; IMIM-Institut Hospital del Mar d'Investigacions Médiques, ; CIBER de Epidemiología y Salud Pública (CIBERESP), Ruiz-Merino, Guadalupe; Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia, Escámez, Teresa; IMIB BIOBANC-MUR. Biobanco-HUVA-AECC-FFIS, ; Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia, Salmerón, Diego; Consejería de Sanidad, Servicio de Epidemiología; CIBER de Epidemiología y Salud Pública (CIBERESP), Júdez, Javier; Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia, ; IMIB BIOBANC-MUR. Biobanco-HUVA-AECC-FFIS, Martínez, Salvador; Instituto de Neurociencias. UMH-CSIC, Navarro, Carmen; Consejería de Sanidad, Servicio de Epidemiología; CIBER de Epidemiología y Salud Pública (CIBERESP),
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Epidemiology, Genetics and genomics, Public health
Keywords:	MENTAL HEALTH, EPIDEMIOLOGY, GENETICS

STUDY PROTOCOL

Title: **Epidemiology and Genetics of Common Mental Disorders in the general population: the PEGASUS-Murcia project**

Authors: F. Navarro-Mateu[1,2,*], M.J. Tormo[2,3,4], G. Vilagut[2,5], J. Alonso[2,5,6], G. Ruíz-Merino[7], T. Escámez[7,8], D. Salmerón[2,3,4], J. Júdez[7,8], S. Martínez[9], C. Navarro[2,3,4]

[1] Unidad de Docencia, Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud. Murcia. Spain.
[2] CIBER de Epidemiología y Salud Pública (CIBERESP). Spain.
[3] Servicio de Epidemiología. Consejería de Sanidad y Política Social. Murcia. Spain.
[4] Departamento de Ciencias Sociosanitarias. Universidad de Murcia. Spain.
[5] IMIM-Institut Hospital del Mar d´Investigacions Médiques. Barcelona. Spain.
[6] Departamento de Salud y Ciencias Experimentales. Universidad Pompeu Fabra. Barcelona. Spain.
[7] Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia. Spain.
[8] IMIB BIOBANC-MUR. Biobanco-HUVA-AECC-FFIS. Murcia. Spain.
[9] Instituto de Neurociencias. UMH-CSIC. Alicante. Spain.

* **Corresponding author:** Fernando Navarro-Mateu, MD, PhD. Unidad de Docencia, Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud. c/ Lorca, nº 58. 30120-El Palmar (Murcia). Spain. Email: fernando.navarro@carm.es

Emails of all coauthors

FNM: fernando.navarro@carm.es
MJT: mjose.tormo@carm.es
GV: gvilagut@imim.es
JA: jalonso@imim.es
GRM: guadalupe.ruiz@carm.es
TE: teresa.escamez@carm.es
DS: diego.salmeron@carm.es
JJ: javier.judez@ffis.es
SM: smartinez@umh.es
CN: carmen.navarro@carm.es

KEYWORDS: **Cross-sectional survey, mental disorders, prevalence, gene-environmental interactions, genome, epigenome, transcriptome.**

Word count (excluding title page, abstract, references, figures and tables)**: 5 689**

ABSTRACT (298 words)

Background: Multidisciplinary collaboration between clinicians, epidemiologists, neurogeneticists and statisticians on research projects has been encouraged to improve our knowledge of the complex mechanisms underlying the etiology and burden of mental disorders. The PEGASUS-Murcia project was designed to assess the prevalence of common mental disorders, to identify risk and protective factors and it also included the collection of biological samples to study gene-environmental interactions in the context of the World Mental Health Survey Initiative.

Methods and Analysis: The PEGASUS-Murcia project is a new cross-sectional face-to-face interview survey based on a representative sample of non-institutionalized adults in the Region of Murcia (Mediterranean Southeast, Spain). Trained lay interviewers used the latest version for use in Spain of the computer-assisted personal interview (CAPI) of the Composite International Diagnostic Interview (CIDI 3.0), specifically adapted for the project. Two biological samples of buccal mucosal epithelium will be collected from each interviewed participant, one for DNA extraction for genomic and epigenomic analyses and the other to obtain mRNA for gene expression quantification. Several quality control procedures will be implemented to assure the highest reliability and validity of the data. This paper describes the rationale, sampling methods and questionnaire content as well as the laboratory methodology.

Ethics and dissemination: Informed consent will be obtained from all participants and a Regional Ethics Research Committee has approved the protocol. Results will be disseminated in peer reviewed publications and presented at national and international conferences.

Discussion: Cross-sectional studies, which combine detailed personal information with biological data, offer new and exciting opportunities to study gene-environmental interactions in the etiology of common mental disorders in representative samples of the general population. A collaborative multidisciplinary research approach offers the potential to advance our knowledge of the underlying complex interactions and this opens the field for further innovative study designs in psychiatric epidemiology.

KEYWORDS: Cross-sectional survey, mental disorders, prevalence, gene-environmental interactions, genome, epigenome, transcriptome.

ARTICLE SUMMARY

Article focus

- Study protocol of the PEGASUS-Murcia project, a new cross-sectional face-to-face interview survey based on a representative sample of non-institutionalized adults in the Region of Murcia (Mediterranean Southeast, Spain).
- The first objective is to estimate the prevalence of the most common mental disorders in general population, analyzing the association with sociodemographic factors, quality of life, treatment, use of services, unmet need and quality of care received and comparing the results with those obtained from Spain, Europe and other non-European countries.
- The second objective it to study the genetic, epigenetic and transcriptomic influences associated with mental disorders.

Key messages

- Multidisciplinary research team better approaches the study of the complex interactions between environmental and genetic risk and protective factors involved in mental disorders.

Strengths and limitation of this study

- The major strength of this protocol is the assessment of environmental and genetic factors not only associated to mental disorder but also with positive mental health in a representative sample of the general population by a multidisciplinary research team.
- The limitation of this protocol is that its cross-sectional design which, while it allows association studies and the generation of new hypotheses, limits the possible causal interpretation of the findings.

BACKGROUND

The World Mental Health (WMH) Survey Initiative is a WHO (World Health Organization) initiative specifically designed to carry out epidemiological surveys in a representative number of countries in all major regions of the world.[1-3] All previous WMH surveys have used or are currently using the same diagnostic interview, the WHO Composite International Diagnostic Interview (WMH-CIDI, hereafter referred to as CIDI), a fully-structured research diagnostic interview questionnaire designed to be used by trained lay interviewers without clinical experience. This initiative has generated an enormous body of comparative cross-national data on the epidemiology of mental disorders all over the world.[3-7] As part of it, the European Study of the Epidemiology of Mental Disorders (ESEMeD) project was designed to collect data from representative samples of the adult population in six European countries: Belgium, France, Germany, Italy, the Netherlands and Spain.[2,8,9] It has also generated a large number of scientific papers on the most prevalent mental health disorders (mood, anxiety, and alcohol abuse) in Europe.[10-17] There is a general consensus on the importance of the ESEMeD project in terms of improving scientific knowledge of the epidemiology of mental disorders in Europe.[1,2,9]

Genes and environment factors in the etiology of mental disorders.

Despite decades of intensive research, it remains difficult to identify specific genes and to characterize those environmental factors primarily responsible for mental disorders.[18-22] The concept of genes and environmental factors as independent causes of mental disorders has been replaced by one of complex interactions between them. These Gene-Environment (GxE) interactions imply a genetic predisposition of some subjects to be expressed differently depending on the environment to which they are exposed.[23,24] For example, the important role of environmental factors, especially stressful life events (SLEs), is now widely accepted. Exposure to various SLEs (work or physical problems, assault, natural disasters, etc.), separately or cumulatively over the life of an individual, increases the risk of depression although in only a proportion of those exposed.[25,26] These data suggest the existence of genetic differences which might explain individual variation in the sensitivity of people to the depressogenic effects of

SLEs. On the other hand, the serotonin transporter (*SERT* or *5HTT*) gene, a key regulator of serotonergic neurotransmission and one of the most studied genetic polymorphisms in relation to affective disorders,[27] has been associated with depression,[28,29] neuroticism[30] and posttraumatic stress disorder (PTSD).[31] However, these findings have not always been replicated.[32-34]

These inconsistencies may be explained by, at least, three different factors. Firstly, in adults, higher levels of neuroticism are associated with an increased risk of depression,[35] anxiety[36] and PTSD after exposure to a traumatic even[37] and are a powerful predictor of comorbidity between depression and anxiety.[38] Neuroticism includes those personality traits that represent how some people perceive the world around them as threatening or stressful. In addition, some personality traits also influence the individual tendency to be potentially exposed to stressful environments. Predisposed individuals may tend to choose environments prone to having a high risk of exposure to stressful events. Specifically, this scenario, known as GxE correlation, may mediate the relationship between neuroticism and specific SLEs.[39] Secondly, the genetic factors influencing the level of neuroticism, including the *5-HTTLPR* polymorphism, are shared by persons having anxious-depressive spectrum disorders.[38,40] Lastly, GxE interactions have been described involving *5-HTTLPR* and depression,[41] anxiety[42] and PTSD.[43] Despite all of the above evidence, genetic association and GxE interaction studies do not usually analyze or control for the level of neuroticism in the relationship between *5-HTTLPR*, SLEs and anxious-depressive spectrum disorders.

However, the question arising in this context is how environmental and genetic factors interact to produce a mental disorder.[21,44] In recent years, increasing interest in epigenetic factors described in other human diseases has focused on its role in mental disorders.[45] The study of the epigenome, changes in gene expression by modulating the accessibility of information that occurs without modifying the DNA sequence, suggests that, although inheritable, these changes are not necessarily stable over the life span of individuals and can be modified under some environmental stimuli that modulate the activity of the enzymes involved, opening new prospects for developing therapeutic approaches based on epigenetic mechanisms.[46] Epigenetic mechanisms have been associated with different mental disorders including depression,[47]

PTSD,[48] schizophrenia,[49,50] autism,[49] bipolar disorder[50] and alcohol dependence.[51] In fact, epigenetic regulation of the glucocorticoid receptor signaling in neurons has been recently shown to be the mechanism underlying GxE interactions to explain risk and resilience of PTSD after SLE in childhood.[52]

In order to integrate all these findings and create new opportunities and challenges offered by the GxE interaction scenarios in the field of mental disorders, a multidisciplinary collaboration between clinicians, epidemiologists, geneticists and statisticians offers greater opportunities.[20,23,53] One of the proposed mechanisms for this collaboration includes carrying out community psychiatric surveys and this has been facilitated by the possibility of obtaining DNA and/or mRNA from peripheral tissues. Specifically, saliva or buccal cells offers an easy, save, inexpensive, and non-invasive method with accumulating scientific rationale to be added in general population surveys.[54-59] Changes in gene expression can also be due to transcriptional alterations. In order to deepen the understanding of molecular mechanisms implicated in mental disorders, it is relevant to take into account transcriptional analyses with the RNA obtained at the same time as the DNA samples. The opportunity to get both biological samples at the same time from saliva offers the challenge of testing the suitability of this material for transcriptional analyses in general population surveys. Population-based surveys offer several advantages over other study designs to contribute to the clarification of the GxE interactions in mental disorders.[44,60,61] Firstly, current knowledge of genes as risk factors is based almost exclusively on clinical and non-representative population samples. Secondly, the distribution of the gene polymorphisms of interest in the general population has not been well investigated. Thirdly, this type of study can provide samples for future case-control studies and can be the bases for future longitudinal ones. Finally, hypotheses generated from epidemiologic surveys may contribute to test new basic studies and can be considered as a complementary strategy to translational research.

Pegasus-Murcia Project

Spain actively participated in the ESEMeD Project with a representative sample of the adult general Spanish population (n=5473) and the results have been published in national and

international journals.[62-67] However, the sample size within most of the Autonomous Communities in Spain was too small to be able to achieve accurate and precise estimates at the Regional level where Health Care policies are decided. Moreover, several differences between the Autonomous Communities in Spain in important aspects related to mental health such as socioeconomic[68] and territorial inequalities in health care supply and in long-term care, access to and use of health care facilities,[69], premature deaths due to alcohol consumption[70] and the prevalence of psychological distress[71] have recently been described.

Murcia is one of the 17 Autonomous Communities of Spain. It is a located in the southeast of the country on the Mediterranean coast, with a population of 1,424,063 inhabitants at the time of the survey (INE 2008, National Statistical Institute of Spain), almost a third of them (30.7%) living in the capital.

The PEGASUS-Murcia ("Psychiatric Enquiry to General Population in Southeast Spain-Murcia") project has been designed in order to obtain regional data of the prevalence, burden and care of a representative sample of the general adult population of Murcia to allow planning of new regional mental health policies and to compare the results with the national sample of Spain, Europe and all other countries participating in the WMH Survey Initiative. The project also constitutes a unique opportunity to initiate a biological bank of a well-studied representative sample of the general population.

Objectives

The PEGASUS-Murcia project is a multi-purpose, observational, cross-sectional, comparative study of the non-institutionalized general population of Murcia Region whose objective is to improve knowledge about common psychiatric disorders in two main areas. The first one is the epidemiology of mental disorders and protective and risk factors in the general population of Murcia. The specific objectives are: i) to estimate the one-month, 12-month and lifetime prevalence of the most common mental disorders, specifically, mood and anxiety disorders, in the general population of Murcia; ii) to assess the independent association of mood and anxiety disorders with sociodemographic factors (gender, age, education and urban/rural location) and selected risk factors (family history, childhood experiences, religion, partnership status and

sexual problems, among others); iii) to assess the quality of life of persons with the most common psychiatric disorders and to analyze how other variables (physical medical conditions and sociodemographic factors) may influence this outcome; iv) to assess the treatment for these disorders and to evaluate the unmet need and the quality of care received; and v) to compare our results with those obtained from Spain, Europe and other non-European countries, including the United States. The second area is the genetic, epigenetic and transcriptomic influences associated with mental disorders. Its specific aims include: i) the estimation of the distribution of different candidate genes in the general population and their association with different psychiatric disorders; ii) the identification of sensitive alleles underlying potential GxE interactions and the study of epigenetic mechanisms involved, specifically, DNA methylation; and iii) the analysis of gene expression alterations through transcriptomic assays.

METHODS AND ANALYSIS

Study Design

The project is a cross-sectional face-to-face interview survey based on a representative sample of the adult and non-institutionalized general population of the Murcia Region. Those who complete the interview will be invited to provide two biological samples from their oral mucous membranes. The target population is defined as persons aged 18 or older residing in Murcia, not living in institutions and with an active health card (defined as persons included in PERSAN, a regional registry that contains all residents with a Health Card which is periodically up-dated. Exclusion criteria are: i) Confirmed irretrievable contact errors (e.g. telephone number and/or address); ii) Institutionalized individuals (e.g. in prison, in a hospital or in another institution) or those living outside the Autonomous Community during the survey field work; and iii) individuals not able to understand the Spanish language or not able to conduct the questionnaire due to his/her physical or mental condition.

Sampling plan

The geographical area of the survey is the Murcia Region and a two-stage, stratified sampling design has been used. The primary sampling unit is the Primary Health Centre and the second is

the individual. The sampling frame has been PERSAN, the regional health care population database in Murcia. Primary Health Centres have been grouped into nine strata, the current Health Care Areas in Murcia Region. The initial sample size was 4,500 adult individuals divided into the nine Health Care Areas with proportionate allocation. A representative sample of two centres has been chosen in each health area, without individual participant replacement. Selection probability for each centre was known a priori and it was proportional to the size of the centre (% of adult individuals registered in the centre) and the proportion of adult individuals in the centre whose place of residence was rural, semi-urban or urban. Within each of the two selected Health Centers, a stratified random sample procedure, performed for each combination of gender, age group (18-24, 25-34, 35-49, 50-64 and 65+) and type of residence (rural, semi-urban and urban), constitutes a stratum and individuals have been selected using simple random sampling.

For each Health Care Area, the sample size of each stratum has been selected such that individuals in with the same demographic characteristics had equal probability of being selected independently of the selected centre. If a high number of those fulfilling the exclusion criteria in one area is reached, a fixed number of additional individuals will be released (subsequent releases), according to the number of interviews completed in the area and following the same selection procedure within each of the centres as the ones used to select the initial release (no new centre will be selected for these releases). Any replacement of those persons who do not want to collaborate or who do not meet the non-eligibility criteria is not allowed.

Survey procedures and data control

Those selected will receive no financial incentive to participate and there will be no individual replacement procedure. Trained lay interviewers carry out the survey using the computer-assisted personal interview (CAPI) that was programmed centrally using the Blaise software system. This is an interviewing application developed by Statistics Netherlands (Herleen, the Netherlands) and designed to ease the handling of elaborate skip and complex randomization patterns and to facilitate data entry, allow the elaboration of some questions and direct the interviewer through the questioning sequence.

Periodically, the completed interviews will be submitted to the central project Data Center (Regional Mental Health Service, Murcia-Spain) for checking and storage following a predetermined security procedure. All raw data will be transferred to the Hospital del Mar Medical Research Institute (IMIM) and the Department of Health Care Policy at Harvard University, coordinating centers of the ESEMeD and WMH Survey Initiative projects respectively, via secure websites. The database has been declared to the Spanish Data Protection Agency.

A survey firm has been contracted to undertake the fieldwork and, in order to ensure the quality of the survey, several strategies are being implemented: i) a one week training course for all interviewers by WHO certified trainers on the original protocol and use of the CAPI version of the CIDI; ii) development of a written manual to standardize the interviewing procedure and all scientific and administrative elements that could affect comparability of data; iii) regular meetings with the survey firm to ensure adherence to the protocol and to deal with any difficulty that may have arisen; and iv) data quality analysis to detect any inconsistencies and/or incomplete data.

The survey firm has been provided with sufficient data to allow contact with each of the individuals of the selected sample and only after 10 unsuccessful attempts the person will be considered as not-contactable or after confirmation that the selected person does not live at that address and new contact information is unavailable. Several methods will be used to improve the participation of those selected: i) an informative flyer providing general information related to the project and giving notice of future contact will be sent by conventional post together with an invitation letter signed by a person from the Health Care Authority; ii) a phone call to invite them to participate in the interview process and to offer them the possibility to do the interview either at home or in their Primary Care Center; iii) Several informative sessions for the healthcare personnel of the Primary Care Centers will be organized to facilitate their collaboration should the participants ask them about the project; iv) During the period when the interviews will take place, some official posters will be put in public centres to inform people

about the project; v) All interviewers will be provided with an official identification and have been trained on how to explain the institutional nature of the research project.

The survey questionnaire

The questionnaire used in the PEGASUS-Murcia project is a revised version of the CIDI which, together with diagnostic information on the most common mental disorders, also includes specific information on the severity of the disorders, symptoms, disability, quality of life, use of services and medication and several risk factors.

The Composite International Diagnostic Interview (CIDI)

The CIDI is a comprehensive, highly structured interview specifically designed by the World Health Organization (WHO) for the purpose of ascertaining diagnoses of mental illnesses based on the WHO International Classification of Disease (ICD-10) and not exclusively on DSM definitions and criteria. This objective is particularly important for cross-national comparative research of the epidemiology of mental illnesses throughout the entire world [72]. It comprises nearly 5000 questions divided into 42 sections (Table 1) and these, in turn, are grouped into two main parts: diagnostic and other. The first includes the clinical part of the interview with an introductory screening section and 22 diagnostic sections that assess different psychiatric conditions. The second includes various non-clinical sections that assess utilization of services, use of psychotropic drugs, degree of functioning in several aspects, chronic physical conditions, risk factors, social networks, caregiver burden and socio-demographic variables.

Table 1: Description of the adapted version of the World Health Organization -Composite International Diagnostic Interview (WHO-CIDI) used in the PEGASUS-Murcia project

Sections	Module	Number of Items	Rules for administration *
Household Listing	Methodological	5	All respondents
Screening (SCR)	Screening	51	All respondents
Minimental State Examination	Risk Factors		If older than 60 years old
Quality/Lie subscale	Functioning and physical Disorder	24	Random assignment to the beginning of the questionnaire or at the end
Depression	Mood disorder	189	Screening questions (SCR)
Mania	Mood disorder	95	Screening questions (SCR)
Panic Disorder	Anxiety	106	Screening questions (SCR)
Specific Phobia	Anxiety	143	Screening questions (SCR)
Social Phobia	Anxiety	85	Screening questions (SCR)
Agoraphobia	Anxiety	84	Screening questions (SCR)

General Anxiety Disorder	Anxiety	116	Screening questions (SCR)
Suicidality	Other Diagnostic	46	All respondents
Use of Services	Treatment	243	All respondents
Group of Questions (Tobacco and physical exercise)	Risk/Protective Factors	22 to 32	All respondents
Pharmacoepidemiology	Treatment	241	All respondents
Substances	Substance abuse	182	Long path
Post-Traumatic Stress Disorder	Anxiety	464 to 491	Long path
Chronic Conditions	Functioning and physical Disorder	201	Long path
30 Days Functions	Functioning and physical Disorder	75	Long path
30 Days Symptoms	Functioning and physical Disorder	75	Long path
Eating Disorders	Other Diagnostic	80	50% of Long path
Obsessive-Compulsive Disorder	Anxiety	124	33% of Long path
CAPE[‡]	Psychosis	42 to 84	All respondents
CFQ[β]	Risk Factors	25	All respondents
SLE[€]	Risk Factors	13 to 39	All respondents
Neuroticism and Extroversion subscales[¥]	Risk/Protective Factors	12	All respondents
Resilience Scale	Protective Factors	25	All respondents
Employment	Socio-demographics	121	Long path
Finances	Socio-demographics	21	Long path
Marriage	Socio-demographics	91	All respondents
Partner violence	Risk Factors	2 to 15	All respondents
Children	Socio-demographics	44	Long path
Social Networks	Risk/Protective Factors	16	All respondents
Adult Demographics	Socio-demographics	68	Long path
Child Demographics	Socio-demographics	34	Long path
Demographic Short	Socio-demographics	25-36	Long path
Childhood	Risk/Protective Factors	110	Long path
Attention Hyperactivity	Childhood	90	Long path and Screening
Oppositional Defiant	Childhood	46	Long path and Screening
Conduct Disorder	Childhood	54	Long path
Separation Anxiety Disorder	Childhood	86	Screening questions (SCR)
Family Burden	Risk Factors	40	Long path
Quality/ Lie subscale	Functioning and physical Disorder	26	Random assignment to the beginning of the questionnaire or at the end
Respondent Contacts	Methodological	19	All respondents
Interviewer Observation	Methodological	14	All respondents

[$] EQ-5D: European Quality of Life Scale; [&] SF-12 v2: Short Form 12 Health Questionnaire; [†] Lie subscale of the abbreviated version of the Eysenck Personality Questionnaire (EPQR-A); [‡] CAPE: Community Assessment of Psychic Experiences; [β] CFQ: Cognitive Failure Questionnaire; [€] SLE: Stressful Life Events; [¥] Neuroticism and Extroversion subscales of the abbreviate version of the Eysenck Personality Questionnaire (EPQR-A)

* Long Path inclusion criteria: a) all individuals that could be considered as "high risk individuals", because they had positively answered a number of specific questions related to mood and anxiety disorders, and b) a random subsample (25%) of the respondents without symptoms ("low risk individuals"). The remaining 75% of respondents without screening symptoms not randomly selected for the long path followed the Short Path of the questionnaire

The most recent version of the CIDI (version 3.0) is the end result of a number of international studies and adaptations made since 2000 when it was first used in WMH surveys. It was first created in English and has been translated into more than 30 different languages using the

standard WHO protocol with a rigorous process of adaptation.[73,74] Several clinical reappraisal studies have been carried out and the concordance of the CIDI version 3.0 has been evaluated in different subgroups of WMH surveys using the Structured Clinical Interview for DSM-IV (SCID) as the clinical gold standard and a moderate to excellent concordance has been found for most mental disorders.[75,76] CIDI is available in two formats: the paper form or PAPI (Paper and Pencil Interviewing) and the computerized form or CAPI (Computer Assisted Personal Interviewing), designed to ease the handling of elaborate skip and complex randomization patterns and to facilitate data entry with a resulting reduction in interview time and errors in data collection and recording. The original Spanish CAPI version used in Spain had not been updated since it was used in the context of the ESEMeD project almost ten years ago. Since then, all improvements in the questionnaire have only been added to the CIDI Latin American (LA) v20.0 version. However, due to linguistic and cultural differences in Spanish-speaking populations, this CAPI version had to be culturally adapted for use in Spain by our research team and this process is fully described elsewhere.[77]

To further shorten the length of the questionnaire, some sections were not selected for the purposes of this project. These include Intermittent Explosive Disorder, Personality I and II, Neurasthenia and Pre-Menstrual and Gambling sections. Some others were substituted by other questions or questionnaires, e.g. the Tobacco Use section was simplified using some questions obtained from the Spanish National Health Survey and the Psychosis section with the CAPE instrument (Community Assessment of Psychic Experiences), both described below.

Other study instruments

Several other instruments were added to the original CIDI for the specific purposes of the PEGASUS-Murcia project. These include the Spanish version of different questionnaires: i) Mini-Mental State Examination for interviewees older than 60 years old;[78,79] ii) the Cognitive Failure Questionnaire (CFQ);[80,81] iii) the Neuroticism, Extroversion and Lie subscales of the abbreviated version of the Eysenck Personality Questionnaire (EPQR-A);[82-84] iv) the Resilience Scale;[85,86] v) the Community Assessment of Psychic Experiences (CAPE)[87] to measure attenuated psychotic symptoms in the general population instead of the Psychosis section of the

CIDI, as the latest is only used as a screening instrument in the detection of psychosis. Those who positively answer two items of the positive dimension with a score equal or superior to 3, have been hospitalized for psychiatric reasons and/or have received psychotropic medication during the last year will be evaluated by a clinic psychiatrist with the module C (Psychotic Disorders) of the SCID (Structured Clinical Interview for DSM Disorders) ; vi) a brief list of 12 stressful life events in the last 12 months was included by the combination of a List of Threatening Experiences (LTE)[88,89] and the emotional and life-changing impact of each event;[90] vii) the European Quality of Life Scale (EuroQol 5D)[91] and the Short Form 12 Health Questionnaire (SF-12 v2);[92] viii) an ad-hoc questionnaire of partner violence obtained from the Spanish National Health Survey and from the regional mental health clinical guidelines;[93] and, finally, ix) some questions related to tobacco use and physical exercises from the Spanish National Health Survey.

Questionnaire pathways

In order to optimize the duration of the interview, the WMH questionnaire was divided into two parts with questions in Part 1 administered to all respondents and those in Part 2 only to a subsample of individuals who followed the long path of the interview. Part 2 of the interview includes detailed information about a wide range of aspects related to the primary disorders and also to mental disorders of secondary interest (Table 1). The inclusion criteria for the long path are: a) all individuals that could be considered as “high risk individuals” because they positively answer a number of specific questions related to mood and anxiety disorders and b) a random subsample (25%) of the respondent without symptoms (“low risk individuals”). The remaining 75% of respondents without screening symptoms not randomly selected for the long path followed the short path. The computer, without any intervention of the interviewer, automatically makes all these pathways. In this shorter itinerary, a specific section, that included those questions needed to calculate some demographic indicators, substituted the sections omitted. Moreover, two sections were only used in a percentage of the long path itinerary, Eating Disorders (50 %) and Obsessive-Compulsive Disorder (33%).

Quality control procedures

Data quality will be controlled in a number of ways to ensure that the predetermined protocol has been followed achieving the greatest reliability and validity and these quality control procedures will be organized and supervised by members of the coordinating centers. The principal investigator will reviewed all responses to open-ended questions to check if narratives excludes a clinical diagnosis of mental disorders, i.e., whether symptoms were due to a physical illness. All these procedures will be verified by the coordinating centers and the final document included several aspects, for example, sample releases, the duration of the interviews and the proportion of positive responses to selected screening questions. Local members of the research team will be responsible for verifying the informed consent forms and the quality checking following computerized protocols. These procedures are similar to those implemented in the ESEMeD project and are fully described elsewhere.[8] Briefly, they consist of checks of individual pieces of information from the interviewees, for example, completion status, consistency across the questionnaire, questionnaire itinerary and length of the interview, and from the interviewers, number of disorders screened positively, verification of a random selection of almost 1% of interviews completed by a telephone contact to confirm the interview and some aspects related to it such as place, approximate duration and identification of the interviewer.

Laboratory Methods

On completion of the interview, interviewees will be asked to provide two biological samples of buccal mucosal epithelium, one for DNA extraction for genomic and epigenomic analysis and the other one to obtain mRNA for gene expression quantification (transcriptomic assays). These samples will be obtained only if the interviewee signs an informed consents specifically designed for this project based on international recommendations for population-based research involving genetics[94] and previously approved by the Regional Ethics Research Committee. Interviewers have been trained by one of the authors (TE) to adequately obtain the biological sample by scraping the oral mucosa using swabs compatible with molecular amplification techniques, as they do not interfere with the amplification process (FLOQSwabs Flocked Swabs, Copan Flock Technologies srl).

Samples for DNA extraction will be collected in sterile 1.5 ml tubes. Those to be used for RNA extraction will be harvested in dark sterile tubes containing RNA protect cell Reagent (QIAGEN, Hilden, Germany), which provides immediate stabilization of RNA. Cells will be thus stabilized at room temperature and can then be stored or transported at ambient temperature prior to RNA purification. Tubes will be labeled with tags (14C.B. 40X40 type) with a specific code for each sample and will be packaged and sent to BIOBANC-MUR (the biobank for biomedical research network of the Region of Murcia, RD09/0076/00065, as a partner of the Spanish National Biobanks Network; IMIB: Instituto Murciano de Investigación Biosanitaria) according to current Spanish legislation and following the regulations of the International Air Transport Association (IATA) on biological sample shipping.

Those sample accepted by BIOBANC-MUR will be registered using a specific biobanking software (bio-e-bank, VITROSOFT, SL), as part of a Laboratory Integrated Management System (LIMS). The nucleic acid extraction will be performed automatically (QIAcube system; QIAGEN, Hilden, Germany) to minimize variability due to manual handling using QIAamp DNA Blood Mini Kit and RNeasyPlus Mini Kit (QIAGEN, Hilden, Germany) for DNA and RNA extraction, respectively.

QIAamp DNA Blood Mini Kit provide fast and easy method for purification of total DNA for reliable PCR and Southern blotting from whole human blood, buffy coat, cultured cells, lymphocytes; plasma, serum, body fluids, and buccal swabs. The synthesis of complementary DNA (cDNA) from mRNA for expression studies will be developed for all samples by reverse transcription using the *High Capacity cDNA Reverse Transcription Kit* (Applied Biosystems). All processes will be performed according to the manufacturer's instructions.

Nucleic acids quantity and quality will be determined by the ratio A260/280 calculated based on 260 and 280 nm absorbance measured using a spectrophotometer.[95-97] The ratio A260/230 is commonly used as a secondary indicator of nucleic acid purity[98-100]. The integrity of DNA will be visualized by electrophoresis on 1% agarose gel (migration for 1 hour at 100 V) using 100 ng of total DNA and a 23 kb DNA ladder (Lambda DNA/HindIII Marker (Thermo Fisher Scientific) as DNA marker. All mRNA samples will be transformed into cDNA.

Specially trained technicians from the BIOBANC-MUR will be used to monitor the specimen collection by donors and to perform sample manipulations in order to minimize variability of results and to obtain the optimal quality of nucleic acids for this and future studies. The processed biospecimens (150 µl of DNA and 80 µl of cDNA) will be stored in 750 µl microtubes in an ultra-freezer at -80 °C located in BIOBANC-MUR.

Statistical methods

The expected response-rate (RR) has been set to a minimum of 65%, based on a previous regional community survey which included the donation of blood samples [101,102]. The response rate will be calculated based on the proportion of people interviewed and was defined as the number of completed interviews divided by the total number of cases minus the number of non-eligible cases.

Weighting procedures

Given that the interview is divided into two parts and only a portion of the sample will be selected for the second part, two types of weightings are considered to estimate population parameters. The first is to weight for the probability of selection for each Health Care Area, Health Centre and demographic stratum and the second is for the random skips included in the questionnaire. The method designed is described in Box 1.

BOX 1: Weighting procedures

First weighting procedure:

Step 1) For each Healthcare Area h, health centre c and demographic stratum (sex, age group and type of residence), all individuals have sampling weight $w_s=1/p_{hc}p^1_{hcsgr}$, where p_{hc} is the probability that the centre c was selected, $p^1_{hcsgr} = n_{hcsgr} / N_{hcsgr}$ and n_{hcsgr} is the sample size for the demographic stratum with N_{hcsgr} individuals registered in the sampling frame.

Step 2) Non-response weight (w_{nr}): if p^*_{hcsgr} is the proportion of eligible persons that is actually interviewed in the Healthcare Area h, centre c, sex s, age group g and type of residence r, the non-response weight of the persons in the Healthcare Area h, centre c, sex s, age group g and type of residence r is $w_{nr} = 1/ p^*_{hcsgr}$.

Step 3) Unadjusted weight (w_{unadj}): it was calculated as the product of sampling weight by non-response weight: $w_{unadj} = w_s w_{nr}$.

Step 4) Post-stratification weight (w_{ps}): data on population of the region of Murcia by sex, age and Healthcare Area were provided by the CREM (*Centro Regional de Estadística de Murcia; Padrón 2010* (http://www.carm.es/econet/sicrem/PU_padron/). The population for the age group 18-24 has been estimated as the population for the age group 18-19 plus the population for the age group 20-24. The population for the age group 18-19 has been estimated as the population for the age group 15-19 times the proportion of population aged 18-19 in the age group 15-19 in Murcia: 0.4116 for males and 0.4165 for females. A post-stratification weight was created to ensure that the joint distribution of the post-stratifying variables Healthcare Area, sex and age group matches the known population joint distribution of Murcia.

Step 5) Adjusted weight (w_{adj}): the adjusted weight of an individual in the Healthcare Area h, centre c,

sex *s*, age group *g* and type of residence *r* is $w_{adj} = w_{unadj} w_{psk}$..

Step 6) Normalized weight: $w_{norm} = w_{adj} \, n / \sum_{i=1}^{n} w_{adj_i}$.

Step 7) Trimmed weight (w_{trim}) : trim the normalized weight obtained from step 6. The upper and lower 5% were trimmed to the mean of each tail.

Step 8) Normalized trimmed weight: $w = w_{trim} \, n / \sum_{i=1}^{n} w_{trim_i}$.

Second weighting procedure:

To take into account the random skips in the CIDI questionnaire applied to define the long path we calculated the skip pattern weights. Only a portion of the sample completed the second part (Part 2) of the survey. The probability of inclusion into Part 2 is based on the presence or absence of disorder symptoms as defined in the interview schedule. Again, different steps will be followed:

Step 1) Part 2 selection weight (w_{p2s})**:** each individual *i* in the sample that accepted to respond the first part of the survey were selected into Part 2 with probability π_i where $\pi_i = 1$ for high risk individuals of having mental disorders and $\pi_i = 0.25$ for the rest. Then the Part 2 selection weight of individual *i* is $w_{p2s} = 1/ \pi_i$.

Step 2) Unadjusted part 2 weight ($w_{p2unadj}$): the product of w_{trim} (Part 1) and the Part 2 selection weights.

Step 3) Part 2 post-stratification weight (w_{p2psk}): similar to the previous post-stratification procedure, a post-stratification weight was created to ensure that the joint distribution of the variables Healthcare Area, sex and age group in Part 2 match the known population distribution of Murcia.

Step 4) Part 2 adjusted weight (w_{p2adj}): the adjusted weight of an individual *i* in the Healthcare Area *h*, centre *c*, sex *s*, age group *g* and type of residence *r* is $w_{p2adj} = w_{p2unadj} w_{p2psk}$.

Step 5) Part 2 Normalized weight: $w_{P2norm} = w_{P2adj} \, n / \sum_{i=1}^{n} w_{P2adj_i}$.

Analysis of the data and forthcoming research projects

There are three data analysis centres in the project: Harvard University (Boston, USA), IMIM (Barcelona, Spain) and the Regional Centers of Epidemiology and Mental Health (Murcia, Spain). Harvard will supervise all quality procedures and provides consultancy in many aspects of the analysis, including the sampling design, the weighting procedures and the verification of the CIDI diagnostic algorithms. All the analyses will be performed using SAS™ and SPSS programs.

Related to this research project, several other lines of research with different designs are being developed, for example, case-control studies and meta-analyses. An example of the former is a case-control study of the GxE interactions, involving *5-HTTLPR* polymorphisms, located in an area where a recent earthquake took place in Lorca (Murcia). It has been specifically designed to analyze its impact in the mental health of the general population exposed. Cases will be those people with a diagnostic of affective and/or anxiety disorder exposed to the earthquake attended in the Mental Health Care Centre and controls will be obtained from those exposed to the

earthquake that are going to be interviewed in the PEGASUS-Murcia project and without a diagnosis of any affective and/or anxiety disorder. Recently, our research team has published a meta-analysis of the relationship between *5-HTTLPR* polymorphism and PTSD.[34]

ETHICS AND DISSEMINATION

Eligible individuals will be asked to sign two independent informed consents to participate, the first one to be interviewed, including the possibility of future new contacts and the second to provide the biological samples but only those who had already completed the questionnaire. Name and contact information will be stored separately from any information provided as part of the study questionnaire. The Clinical Research Ethics Committee of the University Hospital Virgen de la Arrixaca of Murcia approved the protocol and the database of personal information has been registered with the National Data Protection Agency. Data from PEGASUS-Murcia project will be included in the WMH Cross National Sample for international comparisons. The study findings will be submitted to peer-reviewed journals for publication, and presented at national and international scientific meetings.

DISCUSSION

The epidemiology of mental illnesses is a fascinating but highly complex area of research. This complexity is primarily due to the wide range of factors, environmental and genetic, which combines to produce a recognized psychiatric disorder. Previous epidemiological research has resulted in the production of a great amount of data but it has been difficult to make cross-national comparisons due to methodological variability. The WMH Survey Initiative aimed to address this issue by using an international standardized protocol, allowing comparisons of the most common mental disorders and their associated factors throughout the world. Using this study design, it therefore offers the opportunity for new surveys to be performed in the context of an international collaborative initiative and the possibility to adapt the questionnaire according to the specific aims of the research being undertaken. The PEGASUS-Murcia project can be considered as an example of how the latter has been successfully achieved. It is a cross-sectional study designed to assess the prevalence of the most frequent mental disorders and their

correlates in a representative sample of the general population of Murcia. Its primary strengths are: i) the fact that it was specifically adapted to assess factors not only associated with mental disorders but also with positive mental health in a representative sample of the general population; ii) its context focused on regional needs where healthcare decisions are taken regarding resource allocation and mental health planning; iii) the collection of biological samples not only for DNA analysis but also for mRNA; iv) all the information collected in our study, including biological samples, can be correlated with past and future health events because all Spanish population had free access to the Healthcare System at the time of its inception and were thus registered and provided with a unique identification number and therefore; v) finally, the inclusion of a multidisciplinary research team is in accordance with the international consensus regarding the need for interdisciplinary collaboration between clinicians, epidemiologists and neuroscience researchers to increase their combined efforts to study the complex gene-gene and gene-environmental interactions underlying mental health disorders.[23,61,103,104]

Concerns have been expressed about the cost-effectiveness of psychiatric epidemiological surveys, such as World Mental Health 2000 (WMH-2000) projects,[105] an example being the rationale for starting a new psychiatric epidemiological survey in the Autonomous Community of Murcia if Spain had already participated in the ESEMeD project. However, there are several reasons to justify this regional initiative. Firstly, public health and healthcare agencies usually allocate mental health resources, including human, based on data from national epidemiologic surveys,[106] such as that provided by the Spanish participation in the ESEMeD Project. As previously mentioned, the involvement of the Region of Murcia in the Spanish ESEMeD survey did not allow evaluation of specific regional data. Nowadays, the main responsibility for planning and management of Healthcare resources in Spain lies with the Autonomous Communities and differences exist between them in terms of accessibility, amount of health-care resources and political decision-making.[68-71] Devolution of this responsibility to Murcia occurred in December 2001.

Secondly, the inclusion of biological data in a well-designed multidisciplinary epidemiological study offers great advantages in terms of a more global understanding of mental disorders. These are complex illnesses of the brain where social, familial, psychological and biological elements interact throughout the entire life of a person to influence his/her risk of developing a mental health disorder. To extend our understanding of the physiopathology and epidemiology of the more common ones (mood and anxiety), it is necessary to identify genetic loci and polymorphic alleles and their distribution in the healthy and affected population whose function in determining risk for, and protection against, these conditions probably depends on gene-gene and GxE interactions. The collection of genetic material from representative samples from the general population, well described using international diagnostic instruments such as CIDI, offers new and different possibilities to evaluate candidate genes in non-biased samples and to describe their distribution in the general population that may contribute to clarification of the complexity of mental disorders.

Thirdly, our project involving a multidisciplinary research team gives new opportunities to develop different study designs that can move from descriptive to analytical epidemiology. For example, this representative sample constitutes a good source of controls for future case-control studies, where cases will be provided from the public health care clinics, and can be the starting point for future cohort studies. Our project was designed to allow for all these possibilities.

Limitations of the study

Currently, the main limitations of the PEGASUS-Murcia project are related to: i) the cross-sectional design which, while it allows association studies, limits the possible causal interpretation of the findings. However, these findings may provide new hypotheses and enable the design of new studies; ii) not all interviewees will provide biological samples and this may affect the representativeness of some mental disorders in future analyses. To determine if this will result in selection bias, we will analyze whether there are distinguishing characteristics between donors and non-donors in the distribution of mental disorders and other characteristics of the participants; iii) the population stratification in our study which will be used for future genetic association analyses is performed by using the stated ancestral origin by participants[107]

instead of using genetic markers; and iv) biological samples will be obtained from oral mucosal scrapings and not from brain neurons. However, this is a general situation given the ethical issues and difficulties in obtaining neural tissues and, in any case, gene expression does not appear to be specific to neural tissue, at least in some genes that have ubiquitous expression, for example, 5-HTTLPR.[108-111]

Conclusions and Future Directions

The PEGASUS-Murcia project is a sound bases for multidisciplinary collaborative mental health research studies which will provide not only a huge amount of epidemiological information but will also offer exiting opportunities to clarify the complex interactions between genetic and environmental factors which result in a range of mental health disorders.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

FNM, MJT, GV, JA, TE, SM and CN conceived the design and supervised the whole process of the study. GV, JA and FNM have coordinated the project with the WMH Survey Initiative. MJT, JA and CN are coordinating the epidemiologic aspects. TE, JJ and SM are responsible for the genetic aspects. MJT, DS and GV were responsible for the sampling methods. GV, GRM and DS are responsible of the implementation of the qualitative procedures and the statistical analyses. All authors read and approved the final manuscript.

Acknowledgments: We thank Carlos Giribert Muñoz, Deputy Director of Mental Health and Psychiatric Services of Murcia for his support in developing the PEGASUS-Murcia project; Inés Morán-Sánchez, Mª Luisa Pujalte and Ascensión Garriga for their contribution to the initial phases of this project; Monica Ballesta Ruíz for her contribution to the sample selection procedure; David Martínez Martínez for his contribution to the management of the software; all the collaborators from the BIOBANC-MUR (R. Martínez Marín, B. Veas-Pérez de Tudela López, A. Parra Montoya, E. Sánchez Baeza); Pedro J. Bernal for his collaboration with the PERSAN database; and, finally, to Mike Tobin for his helpful discussions and contribution during the English translation of the document.

The PEGASUS-Murcia project is supported by the Regional Health Authorities of Murcia ("Servicio Murciano de Salud and Consejería de Sanidad y Política Social") (Decreto nº 455/2009), the "Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia" (Nº Expedientes: CM0829 I and FFIDS/EMER09/14) and the "Ayudas para proyectos

de Investigación en Salud –ISCIII- del Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica" (PI12/00809). None of the funders had any role in the design, analysis, interpretation of results, or preparation of this paper. The PEGASUS-Murcia project is carried out in conjunction with the WHOWMH Survey Initiative. We thank the WMH Coordinating Center staff at Harvard and Michigan Universities and, especially Professor Ron Kessler, for their assistance with the instrumentation, fieldwork and data analysis. These activities were supported by the United States National Institute of Mental Health (R01MH070884), the John D. and Catherine T. MacArthur Foundation, the Pfizer Foundation, the U.S. Public Health Service (R13-MH066849, R01-MH069864, and R01 DA016558), the Fogarty International Center (FIRCA R03- TW006481), the Pan American Health Organization, the Eli Lilly & Company Foundation, Ortho-McNeil Pharmaceutical, Inc., GlaxoSmithKline, Bristol-Myers Squibb and Shire. A complete list of WMH publications can be found at http://www.hcp.med.harvard.edu/wmh/.

Figure 1: Flow chart of the PEGASUS-Murcia project

† The response rate is defined as: ***(completed interviews) / (total released respondent sample cases –respondent nonsample cases).***

‡ **High risk individuals**: those who positively answer a number of specific questions related to mood and anxiety disorders in the screening section. **Low risk individuals**: those without symptoms related to mood and anxiety disorders in the screening section.

¥ **Long Path inclusion criteria**: a) all high risk individuals and b) a random subsample of 25% of the low risk individuals. The remaining 75% of respondents without screening symptoms not randomly selected for the long path will follow the **Short Path** of the questionnaire

612

References

1. Kessler RC. The global burden of anxiety and mood disorders: putting the European Study of the Epidemiology of Mental Disorders (ESEMeD) findings into perspective. *J Clin Psychiatry*. 2007;68 Suppl 2:10-9.:10-19.
2. Alonso J, Lepine JP. Overview of key data from the European Study of the Epidemiology of Mental Disorders (ESEMeD). *J Clin Psychiatry*. 2007;68 Suppl 2:3-9.
3. Kessler RC, Aguilar-Gaxiola S, Alonso J et al. The WHO World Mental Health (WMH) Surveys. *Psychiatrie (Stuttg)*. 2009;6:5-9.
4. Kessler RC. The World Health Organization International Consortium in Psychiatric Epidemiology (ICPE): initial work and future directions -- the NAPE Lecture 1998. Nordic Association for Psychiatric Epidemiology. *Acta Psychiatr Scand*. 1999;99:2-9.
5. Kessler RC, Angermeyer M, ANTHONY JC et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*. 2007;6:168-176.
6. de Graaf R, Kessler RC, Fayyad J et al. The prevalence and effects of adult attention-deficit/hyperactivity disorder (ADHD) on the performance of workers: results from the WHO World Mental Health Survey Initiative. *Occup Environ Med*. 2008;65:835-842.
7. Scott KM, von Korff M, Alonso J et al. Mental-physical co-morbidity and its relationship with disability: results from the World Mental Health Surveys. *Psychol Med*. 2009;39:33-43.
8. Alonso J, Angermeyer MC, Bernert S et al. Sampling and methods of the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;8-20.
9. Nutt DJ, Kessler RC, Alonso J et al. Consensus statement on the benefit to the community of ESEMeD (European Study of the Epidemiology of Mental Disorders) survey data on depression and anxiety. *J Clin Psychiatry*. 2007;68:42-48.

10. Alonso J, Angermeyer MC, Bernert S et al. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;21-27.

11. Alonso J, Angermeyer MC, Bernert S et al. 12-Month comorbidity patterns and associated factors in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;28-37.

12. Alonso J, Angermeyer MC, Bernert S et al. Disability and quality of life impact of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand*. 2004;420:38-46.

13. Alonso J, Angermeyer MC, Bernert S et al. Use of mental health services in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;47-54.

14. Alonso J, Buron A, Bruffaerts R et al. Association of perceived stigma and mood and anxiety disorders: results from the World Mental Health Surveys. *Acta Psychiatr Scand*. 2008;118:305-314.

15. Alonso J, Codony M, Kovess V et al. Population level of unmet need for mental healthcare in Europe. *Br J Psychiatry*. 2007;190:299-306.

16. Codony M, Alonso J, Almansa J et al. Perceived need for mental health care and service use among adults in Western Europe: results of the ESEMeD project. *Psychiatr Serv*. 2009;60:1051-1058.

17. Sevilla-Dedieu C, Kovess-Masfety V, Gilbert F et al. Mental health care and out-of-pocket expenditures in Europe: results from the ESEMeD project. *J Ment Health Policy Econ*. 2011;14:95-105.

18. Hyman SE. The millennium of mind, brain, and behavior. *Arch Gen Psychiatry*. 2000;57:88-89.

19. Kubota T, Miyake K, Hirasawa T. Epigenetic understanding of gene-environment interactions in psychiatric disorders: a new concept of clinical genetics. *Clin Epigenetics*. 2012;4:1.

20. Merikangas KR, Zhang H, Avenevoli S et al. Longitudinal trajectories of depression and anxiety in a prospective community study: the Zurich Cohort Study. *Arch Gen Psychiatry*. 2003;60:993-1000.

21. Toyokawa S, Uddin M, Koenen KC et al. How does the social environment 'get into the mind'? Epigenetics at the intersection of social and psychiatric epidemiology. *Soc Sci Med*. 2012;74:67-74.

22. Tsuang M, Bar J, Stone W et al. Gene-environment interactions in mental disorders. *World Psychiatry*. 2004;3:73-83.

23. Caspi A, Moffitt TE. Gene-environment interactions in psychiatry: joining forces with neuroscience. *Nat Rev Neurosci*. 2006;7:583-590.

24. Caspi A, Hariri AR, Holmes A et al. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *Am J Psychiatry*. 2010;167:509-527.

25. Koenen KC, Fu QJ, Ertel K et al. Common genetic liability to major depression and posttraumatic stress disorder in men. *J Affect Disord*. 2008;105:109-115.

26. Norris FH, Friedman MJ, Watson PJ et al. 60,000 disaster victims speak: Part I. An empirical review of the empirical literature, 1981-2001. *Psychiatry*. 2002;65:207-239.

27. Caspi A, Hariri AR, Holmes A et al. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *Am J Psychiatry*. 2010;167:509-527.

28. Clarke H, Flint J, Attwood AS et al. Association of the 5- HTTLPR genotype and unipolar depression: a meta-analysis. *Psychol Med*. 2010;40:1767-1778.

29. Karg K, Burmeister M, Shedden K et al. The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. *Arch Gen Psychiatry*. 2011;68:444-454.

30. Takano A, Arakawa R, Hayashi M et al. Relationship between neuroticism personality trait and serotonin transporter binding. *Biol Psychiatry*. 2007;62:588-592.

31. Koenen K, Kilpatrick D, Acierno R et al. Serotonin transporter gene and social support moderate post-traumatic stress disorder in hurricane exposed adults. *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics*. 2006;141B:742-743.

32. Risch N, Herrell R, Lehner T et al. Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis. *JAMA*. 2009;301:2462-2471.

33. Munafo MR, Durrant C, Lewis G et al. Gene X environment interactions at the serotonin transporter locus. *Biol Psychiatry*. 2009;65:211-219.

34. Navarro-Mateu F, Escámez T, Koenen KC et al. Meta-Analyses of the *5-HTTLPR* Polymorphisms and Post-Traumatic Stress Disorder. *PLOS One*. 2013;8:e66227.

35. Kendler KS, Neale MC, Kessler RC et al. A longitudinal twin study of personality and major depression in women. *Arch Gen Psychiatry*. 1993;50:853-862.

36. Kotov R, Gamez W, Schmidt F et al. Linking "big" personality traits to anxiety, depressive, and substance use disorders: a meta-analysis. *Psychol Bull*. 2010;136:768-821.

37. Wu D, Yin H, Xu S et al. Risk factors for posttraumatic stress reactions among Chinese students following exposure to a snowstorm disaster. *BMC Public Health*. 2011;11:96.

38. Hettema JM, Neale MC, Myers JM et al. A population-based twin study of the relationship between neuroticism and internalizing disorders. *Am J Psychiatry*. 2006;163:857-864.

39. Ploubidis GB, Frangou S. Neuroticism and psychological distress: to what extent is their association due to person-environment correlation? *European Psychiatry*. 2011;26:1-5.

40. Skelton K, Ressler KJ, Norrholm SD et al. PTSD and gene variants: new pathways and new thinking. *Neuropharmacology*. 2012;62:628-637.

41. Surtees PG, Wainwright NW, Willis-Owen SA et al. Social adversity, the serotonin transporter (5-HTTLPR) polymorphism and major depressive disorder. *Biol Psychiatry*. 2006;59:224-229.

42. Stein MB, Schork NJ, Gelernter J. Gene-by-environment (serotonin transporter and childhood maltreatment) interaction for anxiety sensitivity, an intermediate phenotype for anxiety disorders. *Neuropsychopharmacology*. 2008;33:312-319.

43. Koenen KC, Aiello AE, Bakshis E et al. Modification of the association between serotonin transporter genotype and risk of posttraumatic stress disorder in adults by county-level social environment. *Am J Epidemiol*. 2009;169:704-711.

44. Galea S, Uddin M, Koenen K. The urban environment and mental disorders: Epigenetic links. *Epigenetics*. 2011;6:400-404.

45. Esteller M. The necessity of a human epigenome project. *Carcinogenesis*. 2006;27:1121-1125.

46. Narayan P, Dragunow M. Pharmacology of epigenetics in brain disorders. *Br J Pharmacol*. 2010;159:285-303.

47. Olsson CA, Foley DL, Parkinson-Bates M et al. Prospects for epigenetic research within cohort studies of psychological disorder: a pilot investigation of a peripheral cell marker of epigenetic risk for depression. *Biol Psychol*. 2010;83:159-165.

48. Koenen KC, Uddin M, Chang SC et al. SLC6A4 methylation modifies the effect of the number of traumatic events on risk for posttraumatic stress disorder. *Depress Anxiety*. 2011;28:639-647.

49. Smith CL, Bolton A, Nguyen G. Genomic and Epigenomic Instability, Fragile Sites, Schizophrenia and Autism. *Current Genomics*. 2010;11:447-469.

50. Dempster EL, Pidsley R, Schalkwyk LC et al. Disease-associated epigenetic changes in monozygotic twins discordant for schizophrenia and bipolar disorder. *Hum Mol Genet*. 2011;20:4786-4796.

51. Park BY, Lee BC, Jung KH et al. Epigenetic changes of serotonin transporter in the patients with alcohol dependence: methylation of an serotonin transporter promoter CpG island. *Psychiatry Investig*. 2011;8:130-133.

52. Klengel T, Mehta D, Anacker C et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nat Neurosci*. 2013;16:33-41.

53. van Os J, Kenis G, Rutten BP. The environment and schizophrenia. *Nature*. 2010;468:203-212.

54. Lee YH, Wong DT. Saliva: an emerging biofluid for early detection of diseases. *Am J Dent*. 2009;22:241-248.

55. Yan W, Apweiler R, Balgley BM et al. Systematic comparison of the human saliva and plasma proteomes. *Proteomics Clin Appl*. 2009;3:116-134.

56. Ballantyne J. Validity of messenger RNA expression analyses of human saliva. *Clin Cancer Res*. 2007;13:1350.

57. Park NJ, Zhou X, Yu T et al. Characterization of salivary RNA by cDNA library analysis. *Arch Oral Biol*. 2007;52:30-35.

58. Lee YH, Zhou H, Reiss JK et al. Direct saliva transcriptome analysis. *Clin Chem*. 2011;57:1295-1302.

59. Palanisamy V, Wong DT. Transcriptomic analyses of saliva. *Methods Mol Biol*. 2010;666:43-51. doi: 10.1007/978-1-60761-820-1_4.:43-51.

60. Merikangas KR, Risch N. Will the Genomics Revolution Revolutionize Psychiatry? *American Journal of Psychiatry*. 2003;160:625-635.

61. Weissman MM, Brown AS, Talati A. Translational epidemiology in psychiatry: linking population to clinical and basic sciences. *Arch Gen Psychiatry*. 2011;68:600-608.

62. Autonell J, Vila F, Pinto-Meza A et al. [One year prevalence of mental disorders comorbidity and associated socio-demographic risk factors in the general population of Spain. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:4-11.

63. Haro JM, Palacin C, Vilagut G et al. [Prevalence of mental disorders and associated factors: results from the ESEMeD-Spain study]. *Med Clin (Barc)*. 2006;126:445-451.

64. Codony M, Alonso J, Almansa J et al. [Mental health care use in the Spanish general populations. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:21-28.

65. Codony M, Alonso J, Almansa J et al. [Psychotropic medications use in Spain. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:29-36.

66. Gabilondo A, Rojas-Farreras S, Rodriguez A et al. Use of primary and specialized mental health care for a major depressive episode in Spain by ESEMeD respondents. *Psychiatr Serv*. 2011;62:152-161.

67. Perales J, Olaya B, Fernandez A et al. Association of childhood adversities with the first onset of mental disorders in Spain: results from the ESEMeD project. *Soc Psychiatry Psychiatr Epidemiol*. 2013;48:371-384.

68. Borrell C, Peiro R, Ramon N et al. [Socioeconomic inequalities and health plans in the Autonomous Communities of Spain]. *Gac Sanit*. 2005;19:277-285.

69. Gonzalez B, Urbanos RM, Ortega P. [Public and private supply of health services by autonomous communities in Spain]. *Gac Sanit*. 2004;18 Suppl 1:82-89.

70. Fierro I, Yanez JL, Alvarez FJ. [Premature death and potential years of life lost due to alcohol consumption in Spain and the different autonomous communities in 2004]. *Aten Primaria*. 2010;42:95-101.

71. Ricci-Cabello I, Ruiz-Perez I, Plazaola-Castano J et al. [Mental disease, existence of diagnostic, use of psychotropic medication: differences by autonomous communities under the national health survey 2006]. *Rev Esp Salud Publica*. 2010;84:29-41.

72. Kessler RC, Ustun TB. The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res*. 2004;13:93-121.

73. Alonso J, Ferrer M, Romera B et al. The European Study of the Epidemiology of Mental Disorders (ESEMeD/MHEDEA 2000) project: rationale and methods. *Int J Methods Psychiatr Res*. 2002;11:55-67.

74. Harkness J, Pennell B, Villar A et al. Translation Procedures and Translation Assessment in the World Mental Health Survey Initiative. In: Kessler RC, Bedirhan Üstün T, eds. *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders*.2008; New York: Cambridge University Press. World Health Organization.

75. Haro JM, Arbabzadeh-Bouchez S, Bugha T et al. Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health Surveys. *Int J Methods Psychiatr Res*. 2006;15:167-180.

76. Kessler RC, Abelson J, Demler O et al. Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMHCIDI). *Int J Methods Psychiatr Res*. 2004;13:122-139.

77. Navarro-Mateu F, Moran-Sanchez I, Alonso J et al. Cultural adaptation of the Latin American version of the World Health Organization Composite International Diagnostic Interview (WHO-CIDI) (v 3.0) for use in Spain. *Gac Sanit*. 2012;27:325-331.

78. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12:189-198.

79. Lobo A, Saz P, Marcos G et al. [Revalidation and standardization of the cognition mini-exam (first Spanish version of the Mini-Mental Status Examination) in the general geriatric population]. *Med Clin (Barc)*. 1999;112:767-774.

80. Broadbent DE, Cooper PF, FitzGerald P et al. The Cognitive Failures Questionnaire (CFQ) and its correlates. *Br J Clin Psychol*. 1982;21 (Pt 1):1-16.

81. García Martínez J, Sánchez Cánovas J. Adaptación al cuestionario de fallos cognitivos de Broadbent, Cooper, Fitzgerald y Parkes (CFQ, Cognitive Failure Questionnaire). *Análisis y Modificación de Conducta*. 1994;20:727-752.

82. Fusté-Escolano A, Ruíz J. Estructura factorial de la versión reducida del "Eysenck Personality Profiler". *Psicothema*. 2000;12:406-411.

83. Sandín B, Valiente RM, Olmedo M et al. Versión española del cuestionario EPQR-Abreviado (EPQR-A) (II): Replicación factorial, fiabilidad y validez. *Revista de Psicopatología y Psicología Clínica*. 2002;7:207-216.

84. Sandín B, Valiente RM, Chorot P et al. Versión española del cuestionario EPQR-Abreviado (EPQE-A) (I): Análisis exploratorio de la estructura factorial. *Revista de Psicopatología y Psicología Clínica*. 2002;7:195-205.

85. Wagnild G, Young H. Development and psychometric evaluation of the resilience scale. *J Nurs Meas*. 1993;1:165-177.

86. Heilemann M, Lee K, Kury F. Psychometric properties of the spanish version of the Resilience Scale. *J Nurs Meas*. 2003;11:61-72.

87. Stefanis NC, Hanssen M, Smirnis NK et al. Evidence that three dimensions of psychosis have a distribution in the general population. *Psychol Med*. 2002;32:347-358.

88. Brugha T, Bebbington P, Tennant C et al. The List of Threatening Experiences: a subset of 12 life event categories with considerable long-term contextual threat. *Psychol Med*. 1985;15:189-194.

89. Brugha TS, Cragg D. The List of Threatening Experiences: the reliability and validity of a brief life events questionnaire. *Acta Psychiatr Scand*. 1990;82:77-81.

90. van Os J, Park S, Jones P. Neuroticism, life events and mental health: evidence for person--environment correlation. *The British Journal of Psychiatry*. 2001;178:s72-s77.

91. Badia X, Roset M, Montserrat S et al. [The Spanish version of EuroQol: a description and its applications. European Quality of Life scale]. *Med Clin (Barc)*. 1999;112:79-85.

92. Vilagut G, Valderas JM, Ferrer M et al. [Interpretation of SF-36 and SF-12 questionnaires in Spain: physical and mental components]. *Med Clin (Barc)*. 2008;130:726-735.

93. Subdirección General de Salud Mental. *Guía práctica clínica. Actuación en salud mental con mujeres maltratadas por su pareja.* 2010; Murcia: Servicio Murciano de Salud.

94. Beskow LM, Burke W, Merz JF et al. Informed consent for population-based research involving genetics. *JAMA*. 2001;286:2315-2321.

95. Abitan H, Bohr H, Buchhave P. Correction to the Beer-Lambert-Bouguer law for optical absorption. *Appl Opt*. 2008;47:5354-5357.

96. Commoner B, Lipkin D. The Application of the Beer-Lambert Law to Optically Anisotropic Systems. *Science*. 1949;110:41-43.

97. Trumbo TA, Schultz E, Borland MG et al. Applied spectrophotometry: Analysis of a biochemical mixture. *Biochem Mol Biol Educ*. 2013.

98. Mulot C, Stucker I, Clavel J et al. Collection of human genomic DNA from buccal cells for genetics studies: comparison between cytobrush, mouthwash, and treated card. *J Biomed Biotechnol*. 2005;2005:291-296.

99. Rogers NL, Cole SA, Lan HC et al. New saliva DNA collection method compared to buccal cell collection techniques for epidemiological studies. *Am J Hum Biol*. 2007;19:319-326.

100. Viltrop T, Krjutskov K, Palta P et al. Comparison of DNA extraction methods for multiplex polymerase chain reaction. *Anal Biochem*. 2010;398:260-262.

101. Valverde JC, Tormo MJ, Navarro C et al. Prevalence of diabetes in Murcia (Spain): a Mediterranean area characterised by obesity. *Diabetes Res Clin Pract*. 2006;71:202-209.

102. Huerta JM, Tormo MJ, Egea-Caparros JM et al. Accuracy of self-reported diabetes, hypertension and hyperlipidemia in the adult Spanish population. DINO study findings. *Rev Esp Cardiol*. 2009;62:143-152.

103. Hyman SE. The genetics of mental illness: implications for practice. *Bull World Health Organ*. 2000;78:455-463.

104. Merikangas KR, Low NC, Hardy J. Commentary: understanding sources of complexity in chronic diseases--the importance of integration of genetics and epidemiology. *Int J Epidemiol*. 2006;35:590-592.

105. Weich S, Araya R. International and regional variation in the prevalence of common mental disorders: do we need more surveys? *Br J Psychiatry*. 2004;184:289-290.

106. Kessler RC. Psychiatric epidemiology: selected recent advances and future directions. *Bull World Health Organ*. 2000;78:464-474.

107. Mercer KB, Orcutt HK, Quinn JF et al. Acute and Posttraumatic Stress Symptoms in a Prospective Gene x Environment Study of a University Campus Shooting. *Arch Gen Psychiatry*. 2011;69:89-97.

108. Tunbridge EM, Eastwood SL, Harrison PJ. Changed relative to what? Housekeeping genes and normalization strategies in human brain gene expression studies. *Biol Psychiatry*. 2011;69:173-179.

109. Abuhatzira L, Shemer R, Razin A. MeCP2 involvement in the regulation of neuronal alpha-tubulin production. *Hum Mol Genet*. 2009;18:1415-1423.

110. Bayou N, Belhadj A, Daoud H et al. Exploring the 7p22.1 chromosome as a candidate region for autism. *J Biomed Biotechnol*. 2010;2010:423894.

111. Fatjo-Vilas M, Papiol S, Estrada G et al. Dysbindin-1 gene contributes differentially to early- and adult-onset forms of functional psychosis. *Am J Med Genet B Neuropsychiatr Genet*. 2011;156B:322-333.

STUDY PROTOCOL

Title: **Epidemiology and Genetics of Common Mental Disorders in the general population: the PEGASUS-Murcia project**

Authors: F. Navarro-Mateu[1,2,*], M.J. Tormo[2,3,4], G. Vilagut[2,5], J. Alonso[2,5,6], G. Ruíz-Merino[7], T. Escámez[7,8], D. Salmerón[2,3,4], J. Júdez[7,8], S. Martínez[9], C. Navarro[2,3,4]

[1] Unidad de Docencia, Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud. Murcia. Spain.
[2] CIBER de Epidemiología y Salud Pública (CIBERESP). Spain.
[3] Servicio de Epidemiología. Consejería de Sanidad y Política Social. Murcia. Spain.
[4] Departamento de Ciencias Sociosanitarias. Universidad de Murcia. Spain.
[5] IMIM-Institut Hospital del Mar d´Investigacions Médiques. Barcelona. Spain.
[6] Departamento de Salud y Ciencias Experimentales. Universidad Pompeu Fabra. Barcelona. Spain.
[7] Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia. Spain.
[8] IMIB BIOBANC-MUR. Biobanco-HUVA-AECC-FFIS. Murcia. Spain.
[9] Instituto de Neurociencias. UMH-CSIC. Alicante. Spain.

* **Corresponding author:** Fernando Navarro-Mateu, MD, PhD. Unidad de Docencia, Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud. c/ Lorca, nº 58. 30120-El Palmar (Murcia). Spain. Email: fernando.navarro@carm.es

Emails of all coauthors

FNM: fernando.navarro@carm.es
MJT: mjose.tormo@carm.es
GV: gvilagut@imim.es
JA: jalonso@imim.es
GRM: guadalupe.ruiz@carm.es
TE: teresa.escamez@carm.es
DS: diego.salmeron@carm.es
JJ: javier.judez@ffis.es
SM: smartinez@umh.es
CN: carmen.navarro@carm.es

KEYWORDS: **Cross-sectional survey, mental disorders, prevalence, gene-environmental interactions, genome, epigenome, transcriptome.**

Word count (excluding title page, abstract, references, figures and tables)**: 5 689**

ABSTRACT (298 words)

Background: Multidisciplinary collaboration between clinicians, epidemiologists, neurogeneticists and statisticians on research projects has been encouraged to improve our knowledge of the complex mechanisms underlying the etiology and burden of mental disorders. The PEGASUS-Murcia project was designed to assess the prevalence of common mental disorders, to identify risk and protective factors and it also included the collection of biological samples to study gene-environmental interactions in the context of the World Mental Health Survey Initiative.

Methods and Analysis: The PEGASUS-Murcia project is a new cross-sectional face-to-face interview survey based on a representative sample of non-institutionalized adults in the Region of Murcia (Mediterranean Southeast, Spain). Trained lay interviewers used the latest version for use in Spain of the computer-assisted personal interview (CAPI) of the Composite International Diagnostic Interview (CIDI 3.0), specifically adapted for the project. Two biological samples of buccal mucosal epithelium ~~were~~ will be collected from each interviewed participant, one for DNA extraction for genomic and epigenomic analyses and the other to obtain mRNA for gene expression quantification. Several quality control procedures ~~were~~ will be implemented to assure the highest reliability and validity of the data. This paper describes the rationale, sampling methods and questionnaire content as well as the laboratory methodology.

Ethics and dissemination: Informed consent ~~was~~ will be obtained from all participants and a Regional Ethics Research Committee ~~the protocol was~~ has ~~been~~ approved the protocol~~by a Regional Ethics Research Committee~~. Results will be disseminated in peer reviewed publications and presented at national and international conferences.

Discussion: Cross-sectional ~~studies which combine detailed personal information with biological data~~studies, which combine detailed personal information with biological data, offer new and exciting opportunities to study gene-environmental interactions in the etiology of common mental disorders in representative samples of the general population. A collaborative multidisciplinary research approach offers the potential to advance our knowledge of the underlying complex interactions and this opens the field for further innovative study designs in psychiatric epidemiology.

KEYWORDS: Cross-sectional survey, mental disorders, prevalence, gene-environmental interactions, genome, epigenome, transcriptome.

ARTICLE SUMMARY

Article focus

- Study protocol of the PEGASUS-Murcia project, a new cross-sectional face-to-face interview survey based on a representative sample of non-institutionalized adults in the Region of Murcia (Mediterranean Southeast, Spain).
- The first objective is to estimate the prevalence of the most common mental disorders in general population, analyzing the association with sociodemographic factors, quality of life, treatment, use of services, unmet need and quality of care received and comparing the results with those obtained from Spain, Europe and other non-European countries.
- The second objective it to study the genetic, epigenetic and transcriptomic influences associated with mental disorders.

Key messages

- ~~The study of the complex interactions between environmental and genetic risk and protective factors involved in mental disorders is better approached by a multidisciplinary research team~~Multidisciplinary research team better approaches the study of the complex interactions between environmental and genetic risk and protective factors involved in mental disorders.

Strengths and limitation of this study

- The major strength of this protocol is the assessment of environmental and genetic factors not only associated to mental disorder but also with positive mental health in a representative sample of the general population by a multidisciplinary research team.
- The limitation of this protocol is that its cross-sectional design which, while it allows association studies and the generation of new hypotheses, limits the possible causal interpretation of the findings.

BACKGROUND

The World Mental Health (WMH) Survey Initiative is a WHO (World Health Organization) initiative specifically designed to carry out epidemiological surveys in a representative number of countries in all major regions of the world.[1-3] All previous WMH surveys have used or are currently using the same diagnostic interview, the WHO Composite International Diagnostic Interview (WMH-CIDI, hereafter referred to as CIDI), a fully-structured research diagnostic interview questionnaire designed to be used by trained lay interviewers without clinical experience. This initiative has generated an enormous body of comparative cross-national data on the epidemiology of mental disorders all over the world.[3-7] As part of it, the European Study of the Epidemiology of Mental Disorders (ESEMeD) project was designed to collect data from representative samples of the adult population in six European countries: Belgium, France, Germany, Italy, the Netherlands and Spain.[2,8,9] It has also generated a large number of scientific papers on the most prevalent mental health disorders (mood, anxiety, and alcohol abuse) in Europe.[10-17] There is a general consensus on the importance of the ESEMeD project in terms of improving scientific knowledge of the epidemiology of mental disorders in Europe.[1,2,9]

Genes and environment factors in the etiology of mental disorders.

Despite decades of intensive research, it remains difficult to identify specific genes and to characterize those environmental factors primarily responsible for mental disorders.[18-22] The concept of genes and environmental factors as independent causes of mental disorders has been replaced by one of complex interactions between them. These Gene-Environment (GxE) interactions imply a genetic predisposition of some subjects to be expressed differently depending on the environment to which they are exposed.[23,24] For example, the important role of environmental factors, especially stressful life events (SLEs), is now widely accepted. Exposure to various SLEs (work or physical problems, assault, natural disasters, etc.), separately or cumulatively over the life of an individual, increases the risk of depression although in only a proportion of those exposed.[25,26] These data suggest the existence of genetic differences which might explain individual variation in the sensitivity of people to the depressogenic effects of

SLEs. On the other hand, the serotonin transporter (*SERT* or *5HTT*) gene, a key regulator of serotonergic neurotransmission and one of the most studied genetic polymorphisms in relation to affective disorders,[27] has been associated with depression,[28,29] neuroticism[30] and posttraumatic stress disorder (PTSD).[31] However, these findings have not always been replicated.[32-34]

These inconsistencies may be explained by, at least, three different factors. Firstly, in adults, higher levels of neuroticism are associated with an increased risk of depression,[35] anxiety[36] and PTSD after exposure to a traumatic even[37] and are a powerful predictor of comorbidity between depression and anxiety.[38] Neuroticism includes those personality traits that represent how some people perceive the world around them as threatening or stressful. In addition, some personality traits also influence the individual tendency to be potentially exposed to stressful environments. Predisposed individuals may tend to choose environments prone to having a high risk of exposure to stressful events. Specifically, this scenario, known as GxE correlation, may mediate the relationship between neuroticism and specific SLEs.[39] Secondly, the genetic factors influencing the level of neuroticism, including the *5-HTTLPR* polymorphism, are shared by persons having anxious-depressive spectrum disorders.[38,40] Lastly, GxE interactions have been described involving *5-HTTLPR* and depression,[41] anxiety[42] and PTSD.[43] Despite all of the above evidence, genetic association and GxE interaction studies do not usually analyze or control for the level of neuroticism in the relationship between *5-HTTLPR*, SLEs and anxious-depressive spectrum disorders.

However, the question arising in this context is how environmental and genetic factors interact to produce a mental disorder.[21,44] In recent years, increasing interest in epigenetic factors described in other human diseases has focused on its role in mental disorders.[45] The study of the epigenome, changes in gene expression by modulating the accessibility of information that occurs without modifying the DNA sequence, suggests that, although inheritable, these changes are not necessarily stable over the life span of individuals and can be modified under some environmental stimuli that modulate the activity of the enzymes involved, opening new prospects for developing therapeutic approaches based on epigenetic mechanisms.[46] Epigenetic mechanisms have been associated with different mental disorders including depression,[47]

PTSD,[48] schizophrenia,[49,50] autism,[49] bipolar disorder[50] and alcohol dependence.[51] In fact, epigenetic regulation of the glucocorticoid receptor signaling in neurons has been recently shown to be the mechanism underlying GxE interactions to explain risk and resilience of PTSD after SLE in childhood.[52]

In order to integrate all these findings and create new opportunities and challenges offered by the GxE interaction scenarios in the field of mental disorders, a multidisciplinary collaboration between clinicians, epidemiologists, geneticists and statisticians offers greater opportunities.[20,23,53] One of the proposed mechanisms for this collaboration includes carrying out community psychiatric surveys and this has been facilitated by the possibility of obtaining DNA and/or mRNA from peripheral tissues ~~(blood, saliva or buccal cells)~~. Specifically, saliva or buccal cells offers an easy, save, inexpensive, and non-invasive method with accumulating scientific rationale to be added in general population surveys.[54-59] Changes in gene expression can also be due to transcriptional alterations. In order to deepen the understanding of molecular mechanisms implicated in mental disorders, it is relevant to take into account transcriptional analyses with the RNA obtained at the same time as the DNA samples. The opportunity to get both biological samples at the same time from saliva offers the challenge of testing the suitability of this material for transcriptional analyses in general population surveys. Population-based surveys offer several advantages over other study designs to contribute to the clarification of the GxE interactions in mental disorders.[44,60,61] Firstly, current knowledge of genes as risk factors is based almost exclusively on clinical and non-representative population samples. Secondly, the distribution of the gene polymorphisms of interest in the general population has not been well investigated. Thirdly, this type of study can provide samples for future case-control studies and can be the bases for future longitudinal ones. Finally, hypotheses generated from epidemiologic surveys may contribute to test new basic studies and can be considered as a complementary strategy to translational research.

Pegasus-Murcia Project

Spain actively participated in the ESEMeD Project with a representative sample of the adult general Spanish population (n=5473) and the results have been published in national and

international journals.[62-67] However, the sample size within most of the Autonomous Communities in Spain was too small to be able to achieve accurate and precise estimates at the Regional level where Health Care policies are decided. Moreover, several differences between the Autonomous Communities in Spain in important aspects related to mental health such as socioeconomic[68] and territorial inequalities in health care supply and in long-term care, access to and use of health care facilities,[69], premature deaths due to alcohol consumption[70] and the prevalence of psychological distress[71] have recently been described.

Murcia is one of the 17 Autonomous Communities of Spain. It is a located in the southeast of the country on the Mediterranean coast, with a population of 1,424,063 inhabitants at the time of the survey (INE 2008, National Statistical Institute of Spain), almost a third of them (30.7%) living in the capital.

The PEGASUS-Murcia ("Psychiatric Enquiry to General Population in Southeast Spain-Murcia") project has been designed in order to obtain regional data of the prevalence, burden and care of a representative sample of the general adult population of Murcia to allow planning of new regional mental health policies and to compare the results with the national sample of Spain, Europe and all other countries participating in the WMH Survey Initiative. The project also constitutes a unique opportunity to initiate a biological bank of a well-studied representative sample of the general population.

Objectives

The PEGASUS-Murcia project is a multi-purpose, observational, cross-sectional, comparative study of the non-institutionalized general population of Murcia Region whose objective is to improve knowledge about common psychiatric disorders in two main areas. The first one is the epidemiology of mental disorders and protective and risk factors in the general population of Murcia. The specific objectives are: i) to estimate the one-month, 12-month and lifetime prevalence of the most common mental disorders, specifically, mood and anxiety disorders, in the general population of Murcia; ii) to assess the independent association of mood and anxiety disorders with sociodemographic factors (gender, age, education and urban/rural location) and selected risk factors (family history, childhood experiences, religion, partnership status and

sexual problems, among others); iii) to assess the quality of life of persons with the most common psychiatric disorders and to analyze how other variables (physical medical conditions and sociodemographic factors) may influence this outcome; iv) to assess the treatment for these disorders and to evaluate the unmet need and the quality of care received; and v) to compare our results with those obtained from Spain, Europe and other non-European countries, including the United States.

The second ~~objective~~ area is the genetic, epigenetic and transcriptomic influences associated with mental disorders. Its specific aims include: i) the estimation of the distribution of different candidate genes in the general population and their association with different psychiatric disorders; ii) the identification of sensitive alleles underlying potential GxE interactions and the study of epigenetic mechanisms involved, specifically, DNA methylation; and iii) the analysis of gene expression alterations through transcriptomic assays.

METHODS AND ANALYSIS

Study Design

The project is a cross-sectional face-to-face interview survey based on a representative sample of the adult and non-institutionalized general population of the Murcia Region. Those who complete the interview will be invited to provide two biological samples from their oral mucous membranes. The target population is defined as persons aged 18 or older residing in Murcia, not living in institutions and with an active health card (defined as persons included in PERSAN, a regional registry that contains all residents with a Health Card which is periodically up-dated. Exclusion criteria are: i) Confirmed irretrievable contact errors (e.g. telephone number and/or address); ii) Institutionalized individuals (e.g. in prison, in a hospital or in another institution) or those living outside the Autonomous Community during the survey field work; and iii) individuals not able to understand the Spanish language or not able to conduct the questionnaire due to his/her physical or mental condition.

Sampling plan

The geographical area of the survey is the Murcia Region and a two-stage, stratified sampling design has been used. The primary sampling unit is the Primary Health Centre and the second is the individual. The sampling frame has been PERSAN, the regional health care population database in Murcia. Primary Health Centres have been grouped into nine strata, the current Health Care Areas in Murcia Region. The initial sample size was 4,500 adult individuals divided into the nine Health Care Areas with proportionate allocation. A representative sample of two centres has been chosen in each health area, without individual participant replacement. Selection probability for each centre was known a priori and it was proportional to the size of the centre (% of adult individuals registered in the centre) and the proportion of adult individuals in the centre whose place of residence was rural, semi-urban or urban. Within each of the two selected Health Centers, a stratified random sample procedure, performed for each combination of gender, age group (18-24, 25-34, 35-49, 50-64 and 65+) and type of residence (rural, semi-urban and urban), constitutes a stratum and individuals have been selected using simple random sampling.

For each Health Care Area, the sample size of each stratum has been selected such that individuals in with the same demographic characteristics had equal probability of being selected independently of the selected ~~center~~centre. If a high number of those fulfilling the exclusion criteria in one area is reached, a fixed number of additional individuals will be released (subsequent releases), according to the number of interviews completed in the area and following the same selection procedure within each of the ~~centers~~centres as the ones used to select the initial release (no new ~~center~~centre will be selected for these releases). Any replacement of those persons who do not want to collaborate or who do not meet the non-eligibility criteria is not allowed.

Survey procedures and data control

Those selected will receive no financial incentive to participate and there will be no individual replacement procedure. Trained lay interviewers carry out the survey ~~Questions are asked by trained lay interviewers~~ using the computer-assisted personal interview (CAPI) that was programmed centrally using the Blaise software system. This is an interviewing application

developed by Statistics Netherlands (Herleen, the Netherlands) and designed to ease the handling of elaborate skip and complex randomization patterns and to facilitate data entry, allow the elaboration of some questions and direct the interviewer through the questioning sequence. Periodically, the completed interviews will be submitted to the central project Data Center (Regional Mental Health Service, Murcia-Spain) for checking and storage following a predetermined security procedure. All raw data will be transferred to the Hospital del Mar Medical Research Institute (IMIM) and the Department of Health Care Policy at Harvard University, coordinating centers of the ESEMeD and WMH Survey Initiative projects respectively, via secure websites. The database has been declared to the Spanish Data Protection Agency.

A survey firm has been contracted to undertake the fieldwork and, in order to ensure the quality of the survey, several strategies are being implemented: i) a one week training course for all interviewers by WHO certified trainers on the original protocol and use of the CAPI version of the CIDI; ii) development of a written manual to standardize the interviewing procedure and all scientific and administrative elements that could affect comparability of data; iii) regular meetings with the survey firm to ensure adherence to the protocol and to deal with any difficulty that may have arisen; and iv) data quality analysis to detect any inconsistencies and/or incomplete data.

The survey firm has been provided with sufficient data to allow contact with each of the individuals of the selected sample and only after 10 unsuccessful attempts the person will be considered ~~to be~~as not~~ un~~contactable or after confirmation that the selected person does not live at that address and new contact information is unavailable. Several methods will be used to improve the participation of those selected: i) an informative flyer providing general information related to the project and giving notice of future contact will be sent by conventional post together with an invitation letter signed by a person from the Health Care Authority; ii) a phone call to invite them to participate in the interview process and to offer them the possibility to do the interview either at home or in their Primary Care Center; iii) Several informative sessions for the healthcare personnel of the Primary Care Centers will be

organized to facilitate their collaboration should the participants ask them about the project; iv) During the period when the interviews will take place, some official posters will be put in public centrers to inform people about the project; v) All interviewers will be provided with an official identification and have been trained on how to explain the institutional nature of the research project.

The survey questionnaire

The questionnaire used in the PEGASUS-Murcia project is a revised version of the CIDI which, together with diagnostic information on the most common mental disorders, also includes specific information on the severity of the disorders, symptoms, disability, quality of life, use of services and medication and several risk factors.

The Composite International Diagnostic Interview (CIDI)

The CIDI is a comprehensive, highly structuredhighly structured interview specifically designed by the World Health Organization (WHO) for the purpose of ascertaining diagnoses of mental illnesses based on the WHO International Classification of Disease (ICD-10) and not exclusively on DSM definitions and criteria. This objective is particularly important for cross-national comparative research of the epidemiology of mental illnesses throughout the entire world [72]. It comprises nearly 5000 questions divided into 42 sections (Table 1) and these, in turn, are grouped into two main parts: diagnostic and other. The first includes the clinical part of the interview with an introductory screening section and 22 diagnostic sections that assess different psychiatric conditions. The second includes various non-clinical sections whichsections that assess utilization of services, use of psychotropic drugs, degree of functioning in several aspects, chronic physical conditions, risk factors, social networks, caregiver burden and socio-demographic variables.

Table 1: Description of the adapted version of the World Health Organization -Composite International Diagnostic Interview (WHO-CIDI) used in the PEGASUS-Murcia project

Sections	Module	Number of Items	Rules for administration *
Household Listing	Methodological	5	All respondents
Screening (SCR)	Screening	51	All respondents
Minimental State Examination	Risk Factors		If older than 60 years old

Quality/Lie subscale	Functioning and physical Disorder	24	Random assignment to the beginning of the questionnaire or at the end
Depression	Mood disorder	189	Screening questions (SCR)
Mania	Mood disorder	95	Screening questions (SCR)
Panic Disorder	Anxiety	106	Screening questions (SCR)
Specific Phobia	Anxiety	143	Screening questions (SCR)
Social Phobia	Anxiety	85	Screening questions (SCR)
Agoraphobia	Anxiety	84	Screening questions (SCR)
General Anxiety Disorder	Anxiety	116	Screening questions (SCR)
Suicidality	Other Diagnostic	46	All respondents
Use of Services	Treatment	243	All respondents
Group of Questions (Tobacco and physical exercise)	Risk/Protective Factors	22 to 32	All respondents
Pharmacoepidemiology	Treatment	241	All respondents
Substances	Substance abuse	182	Long path
Post-Traumatic Stress Disorder	Anxiety	464 to 491	Long path
Chronic Conditions	Functioning and physical Disorder	201	Long path
30 Days Functions	Functioning and physical Disorder	75	Long path
30 Days Symptoms	Functioning and physical Disorder	75	Long path
Eating Disorders	Other Diagnostic	80	50% of Long path
Obsessive-Compulsive Disorder	Anxiety	124	33% of Long path
CAPE[‡]	Psychosis	42 to 84	All respondents
CFQ[ß]	Risk Factors	25	All respondents
SLE[€]	Risk Factors	13 to 39	All respondents
Neuroticism and Extroversion subscales[¥]	Risk/Protective Factors	12	All respondents
Resilience Scale	Protective Factors	25	All respondents
Employment	Socio-demographics	121	Long path
Finances	Socio-demographics	21	Long path
Marriage	Socio-demographics	91	All respondents
Partner violence	Risk Factors	2 to 15	All respondents
Children	Socio-demographics	44	Long path
Social Networks	Risk/Protective Factors	16	All respondents
Adult Demographics	Socio-demographics	68	Long path
Child Demographics	Socio-demographics	34	Long path
Demographic Short	Socio-demographics	25-36	Long path
Childhood	Risk/Protective Factors	110	Long path
Attention Hyperactivity	Childhood	90	Long path and Screening
Oppositional Defiant	Childhood	46	Long path and Screening
Conduct Disorder	Childhood	54	Long path
Separation Anxiety Disorder	Childhood	86	Screening questions (SCR)
Family Burden	Risk Factors	40	Long path
Quality/ Lie subscale	Functioning and physical Disorder	26	Random assignment to the beginning of the questionnaire or at the end
Respondent Contacts	Methodological	19	All respondents
Interviewer Observation	Methodological	14	All respondents

[$] EQ-5D: European Quality of Life Scale; [&] SF-12 v2: Short Form 12 Health Questionnaire; [†] Lie subscale of the abbreviated version of the Eysenck Personality Questionnaire (EPQR-A); [‡] CAPE: Community Assessment of Psychic Experiences; [ß] CFQ: Cognitive Failure Questionnaire; [€] SLE: Stressful Life Events; [¥] Neuroticism and Extroversion subscales of the abbreviate version of the Eysenck Personality Questionnaire (EPQR-A)

* Long Path inclusion criteria: a) all individuals that could be considered as "high risk individuals", because they had positively answered a number of specific questions related to mood and anxiety disorders, and b) a random subsample (25%) of the respondents without symptoms ("low risk individuals"). The remaining 75% of respondents without screening symptoms not randomly selected for the long path followed the Short Path of the questionnaire

The most recent version of the CIDI (version 3.0) is the end result of a number of international studies and adaptations made since 2000 when it was first used in WMH surveys. It was first created in English and has been translated into more than 30 different languages using the standard WHO protocol with a rigorous process of adaptation.[73,74] Several clinical reappraisal studies have been carried out and the concordance of the CIDI version 3.0 has been evaluated in different subgroups of WMH surveys using the Structured Clinical Interview for DSM-IV (SCID) as the clinical gold standard and a moderate to excellent concordance has been found for most mental disorders.[75,76] CIDI is available in two formats: the paper form or PAPI (Paper and Pencil Interviewing) and the computerized form or CAPI (Computer Assisted Personal Interviewing), designed to ease the handling of elaborate skip and complex randomization patterns and to facilitate data entry with a resulting reduction in interview time and errors in data collection and recording. The original Spanish CAPI version used in Spain had not been updated since it was used in the context of the ESEMeD project almost ten years ago. Since then, all improvements in the questionnaire have only been added to the CIDI Latin American (LA) v20.0 version. However, due to linguistic and cultural differences in Spanish-speaking populations, this CAPI version had to be culturally adapted for use in Spain by our research team and this process is fully described elsewhere.[77]

To further shorten the length of the questionnaire, some sections were not selected for the purposes of this project. These include Intermittent Explosive Disorder, Personality I and II, Neurasthenia and Pre-Menstrual and Gambling sections. Some others were substituted by other questions or questionnaires, e.g. the Tobacco Use section was simplified using some questions obtained from the Spanish National Health Survey and the Psychosis section with the CAPE instrument (Community Assessment of Psychic Experiences), both described below.

Other study instruments

Several other instruments were added to the original CIDI for the specific purposes of the PEGASUS-Murcia project. These include the Spanish version of different questionnaires: i) Mini-Mental State Examination for interviewees older than 60 years old;[78,79] ii) the Cognitive

Failure Questionnaire (CFQ);[80,81] iii) the Neuroticism, Extroversion and Lie subscales of the abbreviated version of the Eysenck Personality Questionnaire (EPQR-A);[82-84] iv) the Resilience Scale;[85,86] v) the Community Assessment of Psychic Experiences (CAPE)[87] to measure attenuated psychotic symptoms in the general population instead of the Psychosis section of the CIDI, as the latest is only used as a screening instrument in the detection of psychosis. Those who positively answer two items of the positive dimension with a score equal or superior to 3, have been hospitalized for psychiatric reasons and/or have received psychotropic medication during the last year will be evaluated by a clinic psychiatrist with the module C (Psychotic Disorders) of the SCID (Structured Clinical Interview for DSM Disorders) ; vi) a brief list of 12 stressful life events in the last 12 months was included by the combination of a List of Threatening Experiences (LTE)[88,89] and the emotional and life-changing impact of each event;[90] vii) the European Quality of Life Scale (EuroQol 5D)[91] and the Short Form 12 Health Questionnaire (SF-12 v2);[92] viii) an ad-hoc questionnaire of partner violence obtained from the Spanish National Health Survey and from the regional mental health clinical guidelines;[93] and, finally, ix) some questions related to tobacco use and physical exercises from the Spanish National Health Survey.

Questionnaire pathways

In order to optimize the duration of the interview, the WMH questionnaire was divided into two parts with questions in Part 1 administered to all respondents and those in Part 2 only to a subsample of individuals who followed the long path of the interview. Part 2 of the interview includes detailed information about a wide range of aspects related to the primary disorders and also to mental disorders of secondary interest (Table 1). The inclusion criteria for the long path are: a) all individuals that could be considered as "high risk individuals" because they positively answer a number of specific questions related to mood and anxiety disorders and b) a random subsample (25%) of the respondent without symptoms ("low risk individuals"). The remaining 75% of respondents without screening symptoms not randomly selected for the long path followed the short path. ~~All these pathways are automatically made by the computer without any intervention of the interviewer~~The computer, without any intervention of the interviewer,

automatically makes all these pathways. –In this shorter itinerary, ~~the sections omitted were substituted by a specific section that included those questions needed to calculate some demographic indicators~~a specific section, that included those questions needed to calculate some demographic indicators, substituted the sections omitted. Moreover, two sections were only used in a percentage of the long path itinerary, Eating Disorders (50 %) and Obsessive-Compulsive Disorder (33%).

Quality control procedures

Data quality will be controlled in a number of ways to ensure that the predetermined protocol has been followed achieving the greatest reliability and validity and these quality control procedures will be organized and supervised by members of the coordinating centers. The principal investigator will reviewed all responses to open-ended questions to check if narratives excludes a clinical diagnosis of mental disorders, i.e., whether symptoms were due to a physical illness. All these procedures will be verified by the coordinating centers and the final document included several aspects, for example, sample releases, the duration of the interviews and the proportion of positive responses to selected screening questions. Local members of the research team will be responsible for verifying the informed consent forms and the quality checking following computerized protocols. These procedures are similar to those implemented in the ESEMeD project and are fully described elsewhere.[8] Briefly, they consist of checks of individual pieces of information from the interviewees, for example, completion status, consistency across the questionnaire, questionnaire itinerary and length of the interview, and from the interviewers, number of disorders screened positively, verification of a random selection of almost 1% of interviews completed by a telephone contact to confirm the interview and some aspects related to it such as place, approximate duration and identification of the interviewer.

Laboratory Methods

On completion of the interview, interviewees will be asked to provide two biological samples of buccal mucosal epithelium, one for DNA extraction for genomic and epigenomic analysis and the other one to obtain mRNA for gene expression quantification (transcriptomic assays). These

samples will ~~–~~be obtained only if the interviewee signs an informed consents specifically designed for this project based on international recommendations for population-based research involving genetics[94] and previously approved by~~from each interviewee, one for DNA extraction for genomic and epigenomic analysis and the other one to obtain mRNA for gene expression quantification (transcriptomic assays~~ the Regional Ethics Research Committee. ~~These samples will be taken using swabs compatible with molecular amplification techniques, as they do not interfere with the amplification process (FLOQSwabs Flocked Swabs, Copan Flock Technologies srl).~~ Interviewers have been trained by one of the authors (TE) to adequately obtain the biological sample by scraping the oral mucosa using swabs compatible with molecular amplification techniques, as they do not interfere with the amplification process (FLOQSwabs Flocked Swabs, Copan Flock Technologies srl).

Samples for DNA extraction will be collected in sterile 1.5 ml tubes. Those to be used for RNA extraction will be harvested in dark sterile tubes containing RNA protect cell Reagent (QIAGEN, Hilden, Germany), which provides immediate stabilization of RNA. Cells will be thus stabilized at room temperature and can then be stored or transported at ambient temperature prior to RNA purification. Tubes will be labeled with tags (14C.B. 40X40 type) with a specific code for each sample and will be packaged and sent to BIOBANC-MUR ~~Mur~~ (the biobank for biomedical research network of the Region of Murcia, RD09/0076/00065, as a partner of the Spanish National Biobanks Network; IMIB: Instituto Murciano de Investigación Biosanitaria) according to current Spanish legislation and following the regulations of the International Air Transport Association (IATA) on biological sample shipping.

Those sample accepted by BIOBANC-MUR~~ur~~ will be registered using a specific biobanking software (bio-e-bank, VITROSOFT, SL), as part of a Laboratory Integrated Management System (LIMS). The nucleic acid extraction will be performed automatically (QIAcube system; QIAGEN, Hilden, Germany) to minimize variability due to manual handling using QIAamp DNA Blood Mini Kit and RNeasyPlus Mini Kit (QIAGEN, Hilden, Germany) for DNA and RNA extraction, respectively.

QIAamp DNA Blood Mini Kit provide fast and easy method for purification of total DNA for reliable PCR and Southern blotting from whole human blood, buffy coat, cultured cells, lymphocytes; plasma, serum, body fluids, and buccal swabs. The synthesis of complementary DNA (cDNA) from mRNA for expression studies will be developed for all samples by reverse transcription using the *High Capacity cDNA Reverse Transcription Kit* (Applied Biosystems). All processes will be performed according to the manufacturer's instructions.

Nucleic acids ~~DNA and RNA~~quantity and quality will be ~~be quantitated by measuring~~determined by the ratio A260/280 calculated based on 260 and 280 nm absorbance measured ~~absorbance at 260/280 nm~~ using a spectrophotometer.[95-97] The ratio ~~Between 260 nm and 230 nm (~~A260/230~~)~~ ~~absorbance~~ is commonly used as a secondary indicator of nucleic acid purity[98-100]. The integrity of DNA will be visualized by electrophoresis on 1% agarose gel (migration for 1 hour at 100 V) using 100 ng of total DNA and a 23 kb DNA ladder (Lambda DNA/HindIII Marker (Thermo Fisher Scientific) as DNA marker. All mRNA samples will be transformed into cDNA.

Specially trained technicians from the BIOBANC-M~~ur~~UR will be used to monitor the specimen collection by donors and to perform sample manipulations in order to minimize variability of results and to obtain the optimal quality of nucleic acids for this and future studies. The processed biospecimens (150 µl of DNA and 80 µl of cDNA) will be stored in 750 µl microtubes in an ultra-freezer at -80 °C located in BIOBANC-MUR~~ur~~.

Statistical methods

The expected response-rate (RR) has been set to a minimum of 65%, based on a previous regional community survey which included the donation of blood samples [101,102]. The response rate will be calculated based on the proportion of people interviewed and was defined as the number of completed interviews divided by the total number of cases minus the number of non-eligible cases.

Weighting procedures

Given that the interview is divided into two parts and only a portion of the sample will be selected for the second part, two types of weightings are considered to estimate population

parameters. The first is to weight for the probability of selection for each Health Care Area, Health ~~Center~~Centre and demographic stratum and the second is for the random skips included in the questionnaire. The method designed is described in Box 1.

BOX 1: Weighting procedures

First weighting procedure:

Step 1) For each Healthcare Area h, health centre c and demographic stratum (sex, age group and type of residence), all individuals have sampling weight $w_s=1/p_{hc}p^1_{hcsgr}$, where p_{hc} is the probability that the centre c was selected, $p^1_{hcsgr} = n_{hcsgr} / N_{hcsgr}$ and n_{hcsgr} is the sample size for the demographic stratum with N_{hcsgr} individuals registered in the sampling frame.

Step 2) Non-response weight (w_{nr}): if p^*_{hcsgr} is the proportion of eligible persons that is actually interviewed in the Healthcare Area h, centre c, sex s, age group g and type of residence r, the non-response weight of the persons in the Healthcare Area h, centre c, sex s, age group g and type of residence r is $w_{nr} = 1/ p^*_{hcsgr}$.

Step 3) Unadjusted weight (w_{unadj}): it was calculated as the product of sampling weight by non-response weight: $w_{unadj} = w_s w_{nr}$.

Step 4) Post-stratification weight (w_{ps}): data on population of the region of Murcia by sex, age and Healthcare Area were provided by the CREM (*Centro Regional de Estadística de Murcia; Padrón 2010* (http://www.carm.es/econet/sicrem/PU_padron/). The population for the age group 18-24 has been estimated as the population for the age group 18-19 plus the population for the age group 20-24. The population for the age group 18-19 has been estimated as the population for the age group 15-19 times the proportion of population aged 18-19 in the age group 15-19 in Murcia: 0.4116 for males and 0.4165 for females. A post-stratification weight was created to ensure that the joint distribution of the post-stratifying variables Healthcare Area, sex and age group matches the known population joint distribution of Murcia.

Step 5) Adjusted weight (w_{adj}): the adjusted weight of an individual in the Healthcare Area h, centre c, sex s, age group g and type of residence r is $w_{adj} = w_{unadj} w_{psk}$..

Step 6) Normalized weight: $w_{norm} = w_{adj}\, n \,/ \sum_{i=1}^{n} w_{adj_i}$.

Step 7) Trimmed weight (w_{trim}) : trim the normalized weight obtained from step 6. The upper and lower 5% were trimmed to the mean of each tail.

Step 8) Normalized trimmed weight: $w = w_{trim}\, n \,/ \sum_{i=1}^{n} w_{trim_i}$.

Second weighting procedure:

To take into account the random skips in the CIDI questionnaire applied to define the long path we calculated the skip pattern weights. Only a portion of the sample completed the second part (Part 2) of the survey. The probability of inclusion into Part 2 is based on the presence or absence of disorder symptoms as defined in the interview schedule. Again, different steps will be followed:

Step 1) Part 2 selection weight (w_{p2s})**:** each individual i in the sample that accepted to respond the first part of the survey were selected into Part 2 with probability π_i where $\pi_i = 1$ for high risk individuals of having mental disorders and $\pi_i = 0.25$ for the rest. Then the Part 2 selection weight of individual i is $w_{p2s} = 1/ \pi_i$.

Step 2) Unadjusted part 2 weight ($w_{p2unadj}$): the product of w_{trim} (Part 1) and the Part 2 selection weights.

Step 3) Part 2 post-stratification weight (w_{p2psk}): similar to the previous post-stratification procedure, a post-stratification weight was created to ensure that the joint distribution of the variables Healthcare Area, sex and age group in Part 2 match the known population distribution of Murcia.

Step 4) Part 2 adjusted weight (w_{p2adj}): the adjusted weight of an individual i in the Healthcare Area h, centre c, sex s, age group g and type of residence r is $w_{p2adj} = w_{p2unadj} w_{p2psk}$.

Step 5) Part 2 Normalized weight: $w_{P2norm} = w_{P2adj}\, n \,/ \sum_{i=1}^{n} w_{P2adj_i}$.

Analysis of the data and forthcoming research projects

There are three data analysis centres in the project: Harvard University (Boston, USA), IMIM (Barcelona, Spain) and the Regional Centers of Epidemiology and Mental Health (Murcia, Spain). Harvard will supervise all quality procedures and provides consultancy in many aspects of the analysis, including the sampling design, the weighting procedures and the verification of the CIDI diagnostic algorithms. All the analyses will be performed using SAS™ and SPSS programs.

Related to this research project, several other lines of research with different designs are being developed, for example, case-control studies and meta-analyses. An example of the former is a case-control study of the GxE interactions, involving *5-HTTLPR* polymorphisms, located in an area where a recent earthquake took place in Lorca (Murcia). It has been specifically designed to analyzse ~~the~~ its impact ~~of an earthquake~~ in the mental health of the general population exposed ~~have been recently been granted~~. Cases will be those people with a diagnostic of affective and/or anxiety disorder exposed to the earthquake attended in the Mental Health Care ~~center~~Centre and controls will be obtained from those exposed to the earthquake that are going to be interviewed in the PEGASUS-Murcia project and without a diagnosis of any affective and/or anxiety disorder. Recently, our research team has published a meta-analysis of the relationship between *5-HTTLPR* polymorphism and PTSD.[34]

ETHICS AND DISSEMINATION

Eligible individuals will be asked to sign two independent informed consents to participate, the first one to be interviewed, including the possibility of future new contacts and the second to provide the biological samples but only those who had already completed the questionnaire. Name and contact information will be stored separately from any information provided as part of the study questionnaire. The ~~protocol was approved by the Clinical Research Ethics Committee of the University Hospital *Virgen de la Arrixaca* of Murcia~~Clinical Research Ethics Committee of the University Hospital Virgen de la Arrixaca of Murcia approved the protocol and the database of personal information ~~was~~ has been registered with the National Data

Protection Agency. Data from PEGASUS-Murcia project will be included in the WMH Cross National Sample for international comparisons. The study findings will be submitted to peer-reviewed journals for publication, and presented at national and international scientific meetings.

DISCUSSION

The epidemiology of mental illnesses is a fascinating but highly complex area of research. This complexity is primarily due to the wide range of factors, environmental and genetic, which combines to produce a recognized psychiatric disorder. Previous epidemiological research has resulted in the production of a great amount of data but it has been difficult to make cross-national comparisons due to methodological variability. The WMH Survey Initiative aimed to address this issue by using an international standardized protocol, allowing comparisons of the most common mental disorders and their associated factors throughout the world. Using this study design, it therefore offers the opportunity for new surveys to be performed in the context of an international collaborative initiative and the possibility to adapt the questionnaire according to the specific aims of the research being undertaken. The PEGASUS-Murcia project can be considered as an example of how the latter has been successfully achieved. It is a cross-sectional study designed to assess the prevalence of the most frequent mental disorders and their correlates in a representative sample of the general population of Murcia. Its primary strengths are: i) the fact that it was specifically adapted to assess factors not only associated with mental disorders but also with positive mental health in a representative sample of the general population; ii) its context focused on regional needs where healthcare decisions are taken regarding resource allocation and mental health planning; iii) the collection of biological samples not only for DNA analysis but also for mRNA; iv) all the information collected in our study, including biological samples, can be correlated with past and future health events because all Spanish population had free access to the Healthcare System at the time of its inception and were thus registered and provided with a unique identification number and therefore; v) finally, the inclusion of a multidisciplinary research team is in accordance with the international

consensus regarding the need for interdisciplinary collaboration between clinicians, epidemiologists and neuroscience researchers to increase their combined efforts to study the complex gene-gene and gene-environmental interactions underlying mental health disorders.[23,61,103,104]

Concerns have been expressed about the cost-effectiveness of psychiatric epidemiological surveys, such as World Mental Health 2000 (WMH-2000) projects,[105] an example being the rationale for starting a new psychiatric epidemiological survey in the Autonomous Community of Murcia if Spain had already participated in the ESEMeD project. However, there are several reasons to justify this regional initiative. Firstly, public health and healthcare agencies usually allocate mental health resources, including human, based on data from national epidemiologic surveys,[106] such as that provided by the Spanish participation in the ESEMeD Project. As previously mentioned, the involvement of the Region of Murcia in the Spanish ESEMeD survey did not allow evaluation of specific regional data. Nowadays, the main responsibility for planning and management of Healthcare resources in Spain lies with the Autonomous Communities and differences exist between them in terms of accessibility, amount of health-care resources and political decision-making.[68-71] Devolution of this responsibility to Murcia occurred in December 2001.

Secondly, the inclusion of biological data in a well-designed multidisciplinary epidemiological study offers great advantages in terms of a more global understanding of mental disorders. These are complex illnesses of the brain where social, familial, psychological and biological elements interact throughout the entire life of a person to influence his/her risk of developing a mental health disorder. To extend our understanding of the physiopathology and epidemiology of the more common ones (mood and anxiety), it is necessary to identify genetic loci and polymorphic alleles and their distribution in the healthy and affected population whose function in determining risk for, and protection against, these conditions probably depends on gene-gene and GxE interactions. The collection of genetic material from representative samples from the general population, well described using international diagnostic instruments such as CIDI, offers new and different possibilities to evaluate candidate genes in non-biased samples and to

describe their distribution in the general population that may contribute to clarification of the complexity of mental disorders.

Thirdly, our project involving a multidisciplinary research team gives new opportunities to develop different study designs that can move from descriptive to analytical epidemiology. For example, this representative sample constitutes a good source of controls for future case-control studies, where cases will be provided from the public health care clinics, and can be the starting point for future cohort studies. Our project was designed to allow for all these possibilities.

Limitations of the study

Currently, the main limitations of the PEGASUS-Murcia project are related to: i) the cross-sectional design which, while it allows association studies, limits the possible causal interpretation of the findings. However, these findings may provide new hypotheses and enable the design of new studies; ii) not all interviewees will provide biological samples and this may affect the representativeness of some mental disorders in future analyses. To determine if this will result in selection bias, we will analyze whether there are distinguishing characteristics between donors and non-donors in the distribution of mental disorders and other characteristics of the participants; iii) the population stratification in our study which will be used for future genetic association analyses is performed by using the stated ancestral origin by participants[107] instead of using genetic markers; and iv) biological samples will be obtained from oral mucosal scrapings and not from brain neurons. However, this is a general situation given the ethical issues and difficulties in obtaining neural tissues and, in any case, gene expression does not appear to be specific to neural tissue, at least in some genes that have ubiquitous expression, for example, 5-HTTLPR.[108-111]

Conclusions and Future Directions

The PEGASUS-Murcia project is a sound bases for multidisciplinary collaborative mental health research studies which will provide not only a huge amount of epidemiological information but will also offer exiting opportunities to clarify the complex interactions between genetic and environmental factors which result in a range of mental health disorders.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

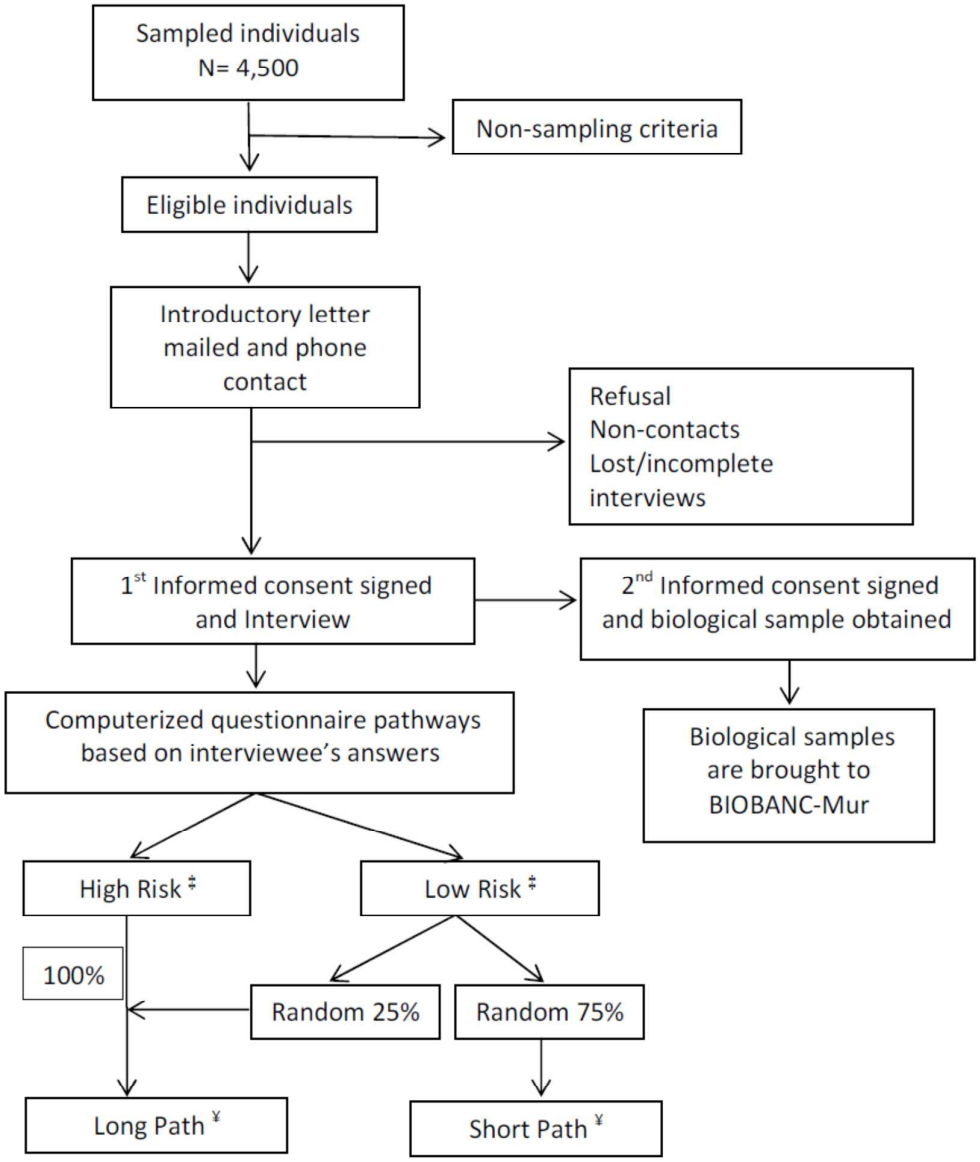
FNM, MJT, GV, JA, TE, SM and CN conceived the design and supervised the whole process of the study. GV, JA and FNM have coordinated the project with the WMH Survey Initiative~~International Consortium of Psychiatric Epidemiology (ICPE)~~. MJT, JA and CN are coordinating the epidemiologic aspects. TE, JJ and SM are responsible for the genetic aspects. MJT, DS and GV were responsible for the sampling methods. GV, GRM and DS are responsible of the implementation of the qualitative procedures and the statistical analyses. All authors read and approved the final manuscript.

Acknowledgments: We thank Carlos Giribert Muñoz, Deputy Director of Mental Health and Psychiatric Services of Murcia for his support in developing the PEGASUS-Murcia project; Inés Morán-Sánchez, Mª Luisa Pujalte and Ascensión Garriga for their contribution to the initial phases of this project; Monica Ballesta Ruíz for her contribution to the sample selection procedure; David Martínez Martínez for his contribution to the management of the software; all the collaborators from the BIOBANC-MUR~~ur~~ (R. Martínez Marín, B. Veas-Pérez de Tudela López, A. Parra Montoya, E. Sánchez Baeza); Pedro J. Bernal for his collaboration with the PERSAN database; and, finally, to Mike Tobin for his helpful discussions and contribution during the English translation of the document.

The PEGASUS-Murcia project ~~was~~ is supported by the Regional Health Authorities of Murcia ("Servicio Murciano de Salud and Consejería de Sanidad y Política Social") (Decreto nº 455/2009), the "Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia" (Nº Expedientes: CM0829 I and FFIDS/EMER09/14) and the "Ayudas para proyectos de Investigación en Salud –ISCIII- del Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica" (PI12/00809). None of the funders had any role in the design, analysis, interpretation of results, or preparation of this paper. The PEGASUS-Murcia project is carried out in conjunction with the WHOWMH Survey Initiative. We thank the WMH Coordinating Center staff at Harvard and Michigan Universities and, especially Professor Ron Kessler, for their assistance with the instrumentation, fieldwork and data analysis. These activities were supported by the United States National Institute of Mental Health (R01MH070884), the John D. and Catherine T. MacArthur Foundation, the Pfizer Foundation, the U.S. Public Health Service (R13-MH066849, R01-MH069864, and R01 DA016558), the Fogarty International Center (FIRCA R03- TW006481), the Pan American Health Organization, the Eli Lilly & Company Foundation, Ortho-McNeil Pharmaceutical, Inc., GlaxoSmithKline,

Bristol-Myers Squibb and Shire. A complete list of WMH publications can be found at http://www.hcp.med.harvard.edu/wmh/.

Figure 1: Flow chart of the PEGASUS-Murcia project



† The response rate is defined as: ***(completed interviews) / (total released respondent sample cases –respondent nonsample cases).***

‡ **High risk individuals**: those who positively answer a number of specific questions related to mood and anxiety disorders in the screening section. **Low risk individuals**: those without symptoms related to mood and anxiety disorders in the screening section.

¥ **Long Path inclusion criteria**: a) all high risk individuals and b) a random subsample of 25% of the low risk individuals. The remaining 75% of respondents without screening symptoms not randomly selected for the long path will follow the **Short Path** of the questionnaire

References

1. Kessler RC. The global burden of anxiety and mood disorders: putting the European Study of the Epidemiology of Mental Disorders (ESEMeD) findings into perspective. *J Clin Psychiatry*. 2007;68 Suppl 2:10-9.:10-19.
2. Alonso J, Lepine JP. Overview of key data from the European Study of the Epidemiology of Mental Disorders (ESEMeD). *J Clin Psychiatry*. 2007;68 Suppl 2:3-9.
3. Kessler RC, Aguilar-Gaxiola S, Alonso J et al. The WHO World Mental Health (WMH) Surveys. *Psychiatrie (Stuttg)*. 2009;6:5-9.
4. Kessler RC. The World Health Organization International Consortium in Psychiatric Epidemiology (ICPE): initial work and future directions -- the NAPE Lecture 1998. Nordic Association for Psychiatric Epidemiology. *Acta Psychiatr Scand*. 1999;99:2-9.
5. Kessler RC, Angermeyer M, ANTHONY JC et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*. 2007;6:168-176.
6. de Graaf R, Kessler RC, Fayyad J et al. The prevalence and effects of adult attention-deficit/hyperactivity disorder (ADHD) on the performance of workers: results from the WHO World Mental Health Survey Initiative. *Occup Environ Med*. 2008;65:835-842.
7. Scott KM, von Korff M, Alonso J et al. Mental-physical co-morbidity and its relationship with disability: results from the World Mental Health Surveys. *Psychol Med*. 2009;39:33-43.
8. Alonso J, Angermeyer MC, Bernert S et al. Sampling and methods of the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;8-20.
9. Nutt DJ, Kessler RC, Alonso J et al. Consensus statement on the benefit to the community of ESEMeD (European Study of the Epidemiology of Mental Disorders) survey data on depression and anxiety. *J Clin Psychiatry*. 2007;68:42-48.

10. Alonso J, Angermeyer MC, Bernert S et al. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;21-27.

11. Alonso J, Angermeyer MC, Bernert S et al. 12-Month comorbidity patterns and associated factors in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;28-37.

12. Alonso J, Angermeyer MC, Bernert S et al. Disability and quality of life impact of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand*. 2004;420:38-46.

13. Alonso J, Angermeyer MC, Bernert S et al. Use of mental health services in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;47-54.

14. Alonso J, Buron A, Bruffaerts R et al. Association of perceived stigma and mood and anxiety disorders: results from the World Mental Health Surveys. *Acta Psychiatr Scand*. 2008;118:305-314.

15. Alonso J, Codony M, Kovess V et al. Population level of unmet need for mental healthcare in Europe. *Br J Psychiatry*. 2007;190:299-306.

16. Codony M, Alonso J, Almansa J et al. Perceived need for mental health care and service use among adults in Western Europe: results of the ESEMeD project. *Psychiatr Serv*. 2009;60:1051-1058.

17. Sevilla-Dedieu C, Kovess-Masfety V, Gilbert F et al. Mental health care and out-of-pocket expenditures in Europe: results from the ESEMeD project. *J Ment Health Policy Econ*. 2011;14:95-105.

18. Hyman SE. The millennium of mind, brain, and behavior. *Arch Gen Psychiatry*. 2000;57:88-89.

19. Kubota T, Miyake K, Hirasawa T. Epigenetic understanding of gene-environment interactions in psychiatric disorders: a new concept of clinical genetics. *Clin Epigenetics*. 2012;4:1.

20. Merikangas KR, Zhang H, Avenevoli S et al. Longitudinal trajectories of depression and anxiety in a prospective community study: the Zurich Cohort Study. *Arch Gen Psychiatry*. 2003;60:993-1000.

21. Toyokawa S, Uddin M, Koenen KC et al. How does the social environment 'get into the mind'? Epigenetics at the intersection of social and psychiatric epidemiology. *Soc Sci Med*. 2012;74:67-74.

22. Tsuang M, Bar J, Stone W et al. Gene-environment interactions in mental disorders. *World Psychiatry*. 2004;3:73-83.

23. Caspi A, Moffitt TE. Gene-environment interactions in psychiatry: joining forces with neuroscience. *Nat Rev Neurosci*. 2006;7:583-590.

24. Caspi A, Hariri AR, Holmes A et al. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *Am J Psychiatry*. 2010;167:509-527.

25. Koenen KC, Fu QJ, Ertel K et al. Common genetic liability to major depression and posttraumatic stress disorder in men. *J Affect Disord*. 2008;105:109-115.

26. Norris FH, Friedman MJ, Watson PJ et al. 60,000 disaster victims speak: Part I. An empirical review of the empirical literature, 1981-2001. *Psychiatry*. 2002;65:207-239.

27. Caspi A, Hariri AR, Holmes A et al. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *Am J Psychiatry*. 2010;167:509-527.

28. Clarke H, Flint J, Attwood AS et al. Association of the 5- HTTLPR genotype and unipolar depression: a meta-analysis. *Psychol Med*. 2010;40:1767-1778.

29. Karg K, Burmeister M, Shedden K et al. The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. *Arch Gen Psychiatry*. 2011;68:444-454.

30. Takano A, Arakawa R, Hayashi M et al. Relationship between neuroticism personality trait and serotonin transporter binding. *Biol Psychiatry*. 2007;62:588-592.

31. Koenen K, Kilpatrick D, Acierno R et al. Serotonin transporter gene and social support moderate post-traumatic stress disorder in hurricane exposed adults. *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics*. 2006;141B:742-743.

32. Risch N, Herrell R, Lehner T et al. Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis. *JAMA*. 2009;301:2462-2471.

33. Munafo MR, Durrant C, Lewis G et al. Gene X environment interactions at the serotonin transporter locus. *Biol Psychiatry*. 2009;65:211-219.

34. Navarro-Mateu F, Escámez T, Koenen KC et al. Meta-Analyses of the *5-HTTLPR* Polymorphisms and Post-Traumatic Stress Disorder. *PLOS One*. 2013;8:e66227.

35. Kendler KS, Neale MC, Kessler RC et al. A longitudinal twin study of personality and major depression in women. *Arch Gen Psychiatry*. 1993;50:853-862.

36. Kotov R, Gamez W, Schmidt F et al. Linking "big" personality traits to anxiety, depressive, and substance use disorders: a meta-analysis. *Psychol Bull*. 2010;136:768-821.

37. Wu D, Yin H, Xu S et al. Risk factors for posttraumatic stress reactions among Chinese students following exposure to a snowstorm disaster. *BMC Public Health*. 2011;11:96.

38. Hettema JM, Neale MC, Myers JM et al. A population-based twin study of the relationship between neuroticism and internalizing disorders. *Am J Psychiatry*. 2006;163:857-864.

39. Ploubidis GB, Frangou S. Neuroticism and psychological distress: to what extent is their association due to person-environment correlation? *European Psychiatry*. 2011;26:1-5.

40. Skelton K, Ressler KJ, Norrholm SD et al. PTSD and gene variants: new pathways and new thinking. *Neuropharmacology*. 2012;62:628-637.

41. Surtees PG, Wainwright NW, Willis-Owen SA et al. Social adversity, the serotonin transporter (5-HTTLPR) polymorphism and major depressive disorder. *Biol Psychiatry*. 2006;59:224-229.

42. Stein MB, Schork NJ, Gelernter J. Gene-by-environment (serotonin transporter and childhood maltreatment) interaction for anxiety sensitivity, an intermediate phenotype for anxiety disorders. *Neuropsychopharmacology*. 2008;33:312-319.

43. Koenen KC, Aiello AE, Bakshis E et al. Modification of the association between serotonin transporter genotype and risk of posttraumatic stress disorder in adults by county-level social environment. *Am J Epidemiol*. 2009;169:704-711.

44. Galea S, Uddin M, Koenen K. The urban environment and mental disorders: Epigenetic links. *Epigenetics*. 2011;6:400-404.

45. Esteller M. The necessity of a human epigenome project. *Carcinogenesis*. 2006;27:1121-1125.

46. Narayan P, Dragunow M. Pharmacology of epigenetics in brain disorders. *Br J Pharmacol*. 2010;159:285-303.

47. Olsson CA, Foley DL, Parkinson-Bates M et al. Prospects for epigenetic research within cohort studies of psychological disorder: a pilot investigation of a peripheral cell marker of epigenetic risk for depression. *Biol Psychol*. 2010;83:159-165.

48. Koenen KC, Uddin M, Chang SC et al. SLC6A4 methylation modifies the effect of the number of traumatic events on risk for posttraumatic stress disorder. *Depress Anxiety*. 2011;28:639-647.

49. Smith CL, Bolton A, Nguyen G. Genomic and Epigenomic Instability, Fragile Sites, Schizophrenia and Autism. *Current Genomics*. 2010;11:447-469.

50. Dempster EL, Pidsley R, Schalkwyk LC et al. Disease-associated epigenetic changes in monozygotic twins discordant for schizophrenia and bipolar disorder. *Hum Mol Genet*. 2011;20:4786-4796.

51. Park BY, Lee BC, Jung KH et al. Epigenetic changes of serotonin transporter in the patients with alcohol dependence: methylation of an serotonin transporter promoter CpG island. *Psychiatry Investig*. 2011;8:130-133.

52. Klengel T, Mehta D, Anacker C et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nat Neurosci*. 2013;16:33-41.

53. van Os J, Kenis G, Rutten BP. The environment and schizophrenia. *Nature*. 2010;468:203-212.

54. Lee YH, Wong DT. Saliva: an emerging biofluid for early detection of diseases. *Am J Dent*. 2009;22:241-248.

55. Yan W, Apweiler R, Balgley BM et al. Systematic comparison of the human saliva and plasma proteomes. *Proteomics Clin Appl*. 2009;3:116-134.

56. Ballantyne J. Validity of messenger RNA expression analyses of human saliva. *Clin Cancer Res*. 2007;13:1350.

57. Park NJ, Zhou X, Yu T et al. Characterization of salivary RNA by cDNA library analysis. *Arch Oral Biol*. 2007;52:30-35.

58. Lee YH, Zhou H, Reiss JK et al. Direct saliva transcriptome analysis. *Clin Chem*. 2011;57:1295-1302.

59. Palanisamy V, Wong DT. Transcriptomic analyses of saliva. *Methods Mol Biol*. 2010;666:43-51. doi: 10.1007/978-1-60761-820-1_4.:43-51.

60. Merikangas KR, Risch N. Will the Genomics Revolution Revolutionize Psychiatry? *American Journal of Psychiatry*. 2003;160:625-635.

61. Weissman MM, Brown AS, Talati A. Translational epidemiology in psychiatry: linking population to clinical and basic sciences. *Arch Gen Psychiatry*. 2011;68:600-608.

62. Autonell J, Vila F, Pinto-Meza A et al. [One year prevalence of mental disorders comorbidity and associated socio-demographic risk factors in the general population of Spain. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:4-11.

63. Haro JM, Palacin C, Vilagut G et al. [Prevalence of mental disorders and associated factors: results from the ESEMeD-Spain study]. *Med Clin (Barc)*. 2006;126:445-451.

64. Codony M, Alonso J, Almansa J et al. [Mental health care use in the Spanish general populations. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:21-28.

65. Codony M, Alonso J, Almansa J et al. [Psychotropic medications use in Spain. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:29-36.

66. Gabilondo A, Rojas-Farreras S, Rodriguez A et al. Use of primary and specialized mental health care for a major depressive episode in Spain by ESEMeD respondents. *Psychiatr Serv*. 2011;62:152-161.

67. Perales J, Olaya B, Fernandez A et al. Association of childhood adversities with the first onset of mental disorders in Spain: results from the ESEMeD project. *Soc Psychiatry Psychiatr Epidemiol*. 2013;48:371-384.

68. Borrell C, Peiro R, Ramon N et al. [Socioeconomic inequalities and health plans in the Autonomous Communities of Spain]. *Gac Sanit*. 2005;19:277-285.

69. Gonzalez B, Urbanos RM, Ortega P. [Public and private supply of health services by autonomous communities in Spain]. *Gac Sanit*. 2004;18 Suppl 1:82-89.

70. Fierro I, Yanez JL, Alvarez FJ. [Premature death and potential years of life lost due to alcohol consumption in Spain and the different autonomous communities in 2004]. *Aten Primaria*. 2010;42:95-101.

71. Ricci-Cabello I, Ruiz-Perez I, Plazaola-Castano J et al. [Mental disease, existence of diagnostic, use of psychotropic medication: differences by autonomous communities under the national health survey 2006]. *Rev Esp Salud Publica*. 2010;84:29-41.

72. Kessler RC, Ustun TB. The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res*. 2004;13:93-121.

73. Alonso J, Ferrer M, Romera B et al. The European Study of the Epidemiology of Mental Disorders (ESEMeD/MHEDEA 2000) project: rationale and methods. *Int J Methods Psychiatr Res*. 2002;11:55-67.

74. Harkness J, Pennell B, Villar A et al. Translation Procedures and Translation Assessment in the World Mental Health Survey Initiative. In: Kessler RC, Bedirhan Üstün T, eds. *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders*.2008; New York: Cambridge University Press. World Health Organization.

75. Haro JM, Arbabzadeh-Bouchez S, Bugha T et al. Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health Surveys. *Int J Methods Psychiatr Res*. 2006;15:167-180.

76. Kessler RC, Abelson J, Demler O et al. Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMHCIDI). *Int J Methods Psychiatr Res*. 2004;13:122-139.

77. Navarro-Mateu F, Moran-Sanchez I, Alonso J et al. Cultural adaptation of the Latin American version of the World Health Organization Composite International Diagnostic Interview (WHO-CIDI) (v 3.0) for use in Spain. *Gac Sanit*. 2012;27:325-331.

78. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12:189-198.

79. Lobo A, Saz P, Marcos G et al. [Revalidation and standardization of the cognition mini-exam (first Spanish version of the Mini-Mental Status Examination) in the general geriatric population]. *Med Clin (Barc)*. 1999;112:767-774.

80. Broadbent DE, Cooper PF, FitzGerald P et al. The Cognitive Failures Questionnaire (CFQ) and its correlates. *Br J Clin Psychol*. 1982;21 (Pt 1):1-16.

81. García Martínez J, Sánchez Cánovas J. Adaptación al cuestionario de fallos cognitivos de Broadbent, Cooper, Fitzgerald y Parkes (CFQ, Cognitive Failure Questionnaire). *Análisis y Modificación de Conducta*. 1994;20:727-752.

82. Fusté-Escolano A, Ruíz J. Estructura factorial de la versión reducida del "Eysenck Personality Profiler". *Psicothema*. 2000;12:406-411.

83. Sandín B, Valiente RM, Olmedo M et al. Versión española del cuestionario EPQR-Abreviado (EPQR-A) (II): Replicación factorial, fiabilidad y validez. *Revista de Psicopatología y Psicología Clínica*. 2002;7:207-216.

84. Sandín B, Valiente RM, Chorot P et al. Versión española del cuestionario EPQR-Abreviado (EPQE-A) (I): Análisis exploratorio de la estructura factorial. *Revista de Psicopatología y Psicología Clínica*. 2002;7:195-205.

85. Wagnild G, Young H. Development and psychometric evaluation of the resilience scale. *J Nurs Meas*. 1993;1:165-177.

86. Heilemann M, Lee K, Kury F. Psychometric properties of the spanish version of the Resilience Scale. *J Nurs Meas*. 2003;11:61-72.

87. Stefanis NC, Hanssen M, Smirnis NK et al. Evidence that three dimensions of psychosis have a distribution in the general population. *Psychol Med*. 2002;32:347-358.

88. Brugha T, Bebbington P, Tennant C et al. The List of Threatening Experiences: a subset of 12 life event categories with considerable long-term contextual threat. *Psychol Med*. 1985;15:189-194.

89. Brugha TS, Cragg D. The List of Threatening Experiences: the reliability and validity of a brief life events questionnaire. *Acta Psychiatr Scand*. 1990;82:77-81.

90. van Os J, Park S, Jones P. Neuroticism, life events and mental health: evidence for person--environment correlation. *The British Journal of Psychiatry*. 2001;178:s72-s77.

91. Badia X, Roset M, Montserrat S et al. [The Spanish version of EuroQol: a description and its applications. European Quality of Life scale]. *Med Clin (Barc)*. 1999;112:79-85.

92. Vilagut G, Valderas JM, Ferrer M et al. [Interpretation of SF-36 and SF-12 questionnaires in Spain: physical and mental components]. *Med Clin (Barc)*. 2008;130:726-735.

93. Subdirección General de Salud Mental. *Guía práctica clínica. Actuación en salud mental con mujeres maltratadas por su pareja.* 2010; Murcia: Servicio Murciano de Salud.

94. Beskow LM, Burke W, Merz JF et al. Informed consent for population-based research involving genetics. *JAMA*. 2001;286:2315-2321.

95. Abitan H, Bohr H, Buchhave P. Correction to the Beer-Lambert-Bouguer law for optical absorption. *Appl Opt*. 2008;47:5354-5357.

96. Commoner B, Lipkin D. The Application of the Beer-Lambert Law to Optically Anisotropic Systems. *Science*. 1949;110:41-43.

97. Trumbo TA, Schultz E, Borland MG et al. Applied spectrophotometry: Analysis of a biochemical mixture. *Biochem Mol Biol Educ*. 2013.

98. Mulot C, Stucker I, Clavel J et al. Collection of human genomic DNA from buccal cells for genetics studies: comparison between cytobrush, mouthwash, and treated card. *J Biomed Biotechnol*. 2005;2005:291-296.

99. Rogers NL, Cole SA, Lan HC et al. New saliva DNA collection method compared to buccal cell collection techniques for epidemiological studies. *Am J Hum Biol*. 2007;19:319-326.

100. Viltrop T, Krjutskov K, Palta P et al. Comparison of DNA extraction methods for multiplex polymerase chain reaction. *Anal Biochem*. 2010;398:260-262.

101. Valverde JC, Tormo MJ, Navarro C et al. Prevalence of diabetes in Murcia (Spain): a Mediterranean area characterised by obesity. *Diabetes Res Clin Pract*. 2006;71:202-209.

102. Huerta JM, Tormo MJ, Egea-Caparros JM et al. Accuracy of self-reported diabetes, hypertension and hyperlipidemia in the adult Spanish population. DINO study findings. *Rev Esp Cardiol*. 2009;62:143-152.

103. Hyman SE. The genetics of mental illness: implications for practice. *Bull World Health Organ*. 2000;78:455-463.

104. Merikangas KR, Low NC, Hardy J. Commentary: understanding sources of complexity in chronic diseases--the importance of integration of genetics and epidemiology. *Int J Epidemiol*. 2006;35:590-592.

105. Weich S, Araya R. International and regional variation in the prevalence of common mental disorders: do we need more surveys? *Br J Psychiatry*. 2004;184:289-290.

106. Kessler RC. Psychiatric epidemiology: selected recent advances and future directions. *Bull World Health Organ*. 2000;78:464-474.

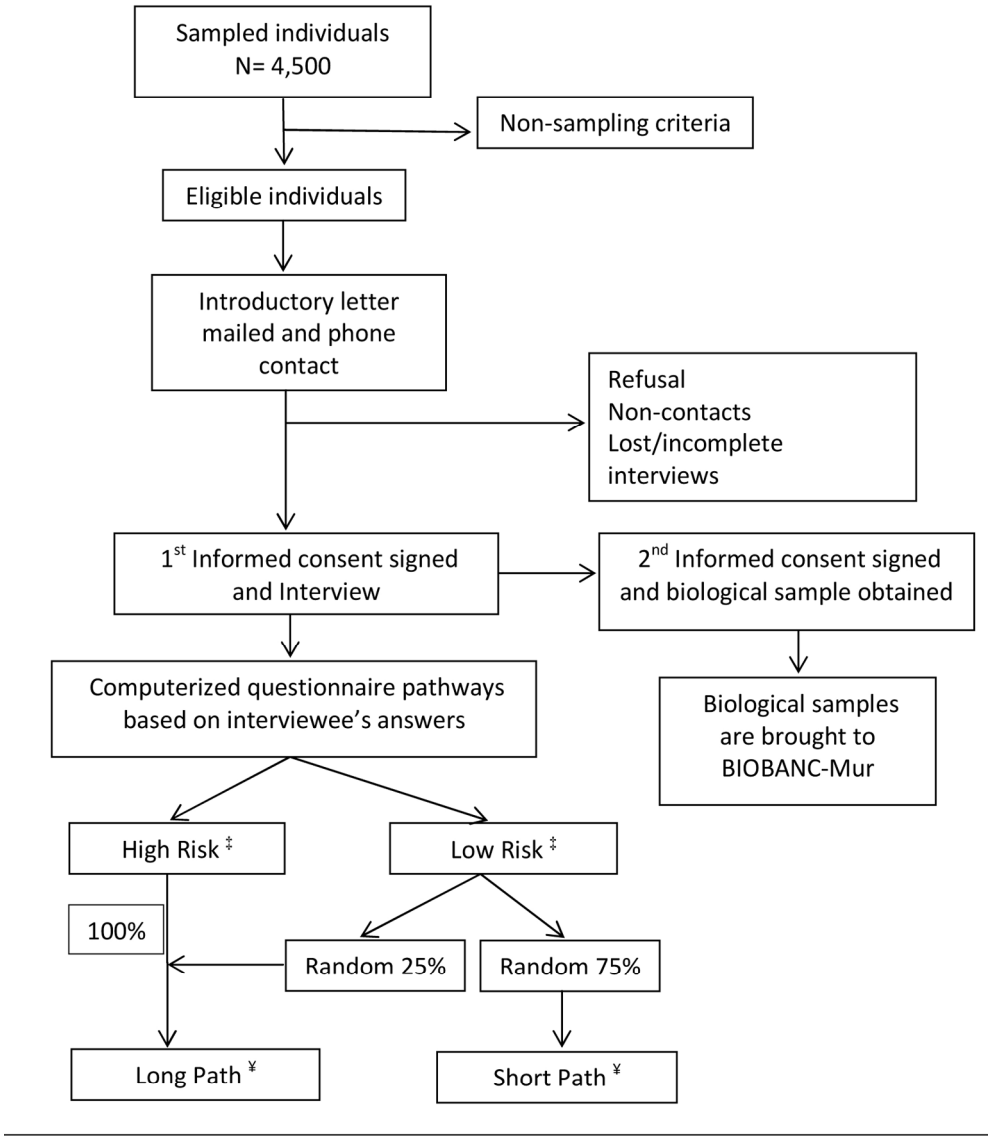
107. Mercer KB, Orcutt HK, Quinn JF et al. Acute and Posttraumatic Stress Symptoms in a Prospective Gene x Environment Study of a University Campus Shooting. *Arch Gen Psychiatry*. 2011;69:89-97.

108. Tunbridge EM, Eastwood SL, Harrison PJ. Changed relative to what? Housekeeping genes and normalization strategies in human brain gene expression studies. *Biol Psychiatry*. 2011;69:173-179.

109. Abuhatzira L, Shemer R, Razin A. MeCP2 involvement in the regulation of neuronal alpha-tubulin production. *Hum Mol Genet*. 2009;18:1415-1423.

110. Bayou N, Belhadj A, Daoud H et al. Exploring the 7p22.1 chromosome as a candidate region for autism. *J Biomed Biotechnol*. 2010;2010:423894.

111. Fatjo-Vilas M, Papiol S, Estrada G et al. Dysbindin-1 gene contributes differentially to early- and adult-onset forms of functional psychosis. *Am J Med Genet B Neuropsychiatr Genet*. 2011;156B:322-333.



† The response rate is defined as: (completed interviews) / (total released respondent sample cases – respondent nonsample cases).

‡ High risk individuals: those who positively answer a number of specific questions related to mood and anxiety disorders in the screening section. Low risk individuals: those without symptoms related to mood and anxiety disorders in the screening section.

¥ Long Path inclusion criteria: a) all high risk individuals and b) a random subsample of 25% of the low risk individuals. The remaining 75% of respondents without screening symptoms not randomly selected for the long path will follow the Short Path of the questionnaire

273x333mm (240 x 240 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Pages
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1-2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2-3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-7
Objectives	3	State specific objectives, including any prespecified hypotheses	7-8
Methods			
Study design	4	Present key elements of study design early in the paper	8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	8-10
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	8-9
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	9-14
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*9-14*
Bias	9	Describe any efforts to address potential sources of bias	14-15
Study size	10	Explain how the study size was arrived at	8-9
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	9-14
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	16-18
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	
		(*e*) Describe any sensitivity analyses	

Continued on next page

Results			Pages
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	-
		(b) Give reasons for non-participation at each stage	-
		(c) Consider use of a flow diagram	-
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	-
		(b) Indicate number of participants with missing data for each variable of interest	-
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	-
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	-
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	-
		Cross-sectional study—Report numbers of outcome events or summary measures	-
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	-
		(*b*) Report category boundaries when continuous variables were categorized	-
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	-
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	-
Discussion			
Key results	18	Summarise key results with reference to study objectives	19-20
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	21
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	-
Generalisability	21	Discuss the generalisability (external validity) of the study results	-
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	22

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.