



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

Journal:	<i>BMJ Open</i>
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, M <sup>a</sup> José; Consejería de Sanidad, Servicio de Epidemiología; CIBER de Epidemiología y Salud Pública (CIBERESP), Vilagut, Gemma; IMIM-Institut Hospital del Mar d'Investigacions Mèdiques, ; CIBER de Epidemiología y Salud Pública (CIBERESP), Alonso, Jordi; IMIM-Institut Hospital del Mar d'Investigacions Mèdiques, ; 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),
<b>Primary Subject Heading</b>:	Mental health
Secondary Subject Heading:	Epidemiology, Genetics and genomics, Public health
Keywords:	MENTAL HEALTH, EPIDEMIOLOGY, GENETICS

SCHOLARONE™  
Manuscripts

1  
2  
3 1 STUDY PROTOCOL4  
5 2 Title: **Epidemiology and Genetics of Common Mental Disorders in the general population:**  
6  
7 3 **the PEGASUS-Murcia project**8  
9  
10 4 **Authors:** F. Navarro-Mateu<sup>1,2,\*</sup>, M.J. Tormo<sup>2,3,4</sup>, G. Vilagut<sup>2,5</sup>, J. Alonso<sup>2,5,6</sup>, G. Ruíz-Merino<sup>7</sup>,  
11  
12 5 T. Escámez<sup>7,8</sup>, D. Salmerón<sup>2,3,4</sup>, J. Júdez<sup>7,8</sup>, S. Martínez<sup>9</sup>, C. Navarro<sup>2,3,4</sup>13  
14  
15  
16 6  
17 7 <sup>1</sup> Unidad de Docencia, Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de  
18 8 Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud. Murcia. Spain.19 9 <sup>2</sup> CIBER de Epidemiología y Salud Pública (CIBERESP). Spain.20 10 <sup>3</sup> Servicio de Epidemiología. Consejería de Sanidad y Política Social. Murcia. Spain.21 11 <sup>4</sup> Departamento de Ciencias Sociosanitarias. Universidad de Murcia. Spain.22 12 <sup>5</sup> IMIM-Institut Hospital del Mar d'Investigacions Mèdiques. Barcelona. Spain.23 13 <sup>6</sup> Departamento de Salud y Ciencias Experimentales. Universidad Pompeu Fabra. Barcelona. Spain.24 14 <sup>7</sup> Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia. Spain.25 15 <sup>8</sup> IMIB BIOBANC-MUR. Biobanco-HUVA-AECC-FFIS. Murcia. Spain.26 16 <sup>9</sup> Instituto de Neurociencias. UMH-CSIC. Alicante. Spain.27  
28  
29 \* **Corresponding author:** Fernando Navarro-Mateu, MD, PhD. Unidad de Docencia,  
30 20 Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de Salud Mental  
31 21 y Asistencia Psiquiátrica. Servicio Murciano de Salud. c/ Lorca, nº 58. 30120-El Palmar  
32 22 (Murcia). Spain. Email: [fernando.navarro@carm.es](mailto:fernando.navarro@carm.es)33  
34 23 **Emails of all coauthors**35  
36 24  
37 25 FNM: [fernando.navarro@carm.es](mailto:fernando.navarro@carm.es)  
38 26 MJT: [mjose.tormo@carm.es](mailto:mjose.tormo@carm.es)  
39 27 GV: [gvilagut@imim.es](mailto:gvilagut@imim.es)  
40 28 JA: [jalonso@imim.es](mailto:jalonso@imim.es)  
41 29 GRM: [guadalupe.ruiz@carm.es](mailto:guadalupe.ruiz@carm.es)  
42 30 TE: [teresa.escamez@carm.es](mailto:teresa.escamez@carm.es)  
43 31 DS: [diego.salmeron@carm.es](mailto:diego.salmeron@carm.es)  
44 32 JJ: [javier.judez@ffis.es](mailto:javier.judez@ffis.es)  
45 33 SM: [smartinez@umh.es](mailto:smartinez@umh.es)  
46 34 CN: [carmen.navarro@carm.es](mailto:carmen.navarro@carm.es)47  
48 37 **KEYWORDS:** Cross-sectional survey, mental disorders, prevalence, gene-environmental  
49 38 interactions, genome, epigenome, transcriptome.50  
51 40 **Word count** (excluding title page, abstract, references, figures and tables): **5 498**

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80

## 41 **ABSTRACT (296 words)**

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

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

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

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

68 **KEYWORDS:** Cross-sectional survey, mental disorders, prevalence, gene-environmental  
69 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 is 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.

71

## 72 **BACKGROUND**

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

### 87 **Genes and environment factors in the etiology of mental disorders.**

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

1  
2  
3 99 SLEs. On the other hand, the serotonin transporter (*SERT* or *5HTT*) gene, a key regulator of  
4 serotonergic neurotransmission and one of the most studied genetic polymorphisms in relation  
5 100 serotonergic neurotransmission and one of the most studied genetic polymorphisms in relation  
6 to affective disorders,<sup>27</sup> has been associated with depression,<sup>28,29</sup> neuroticism<sup>30</sup> and posttraumatic  
7 101 to affective disorders,<sup>27</sup> has been associated with depression,<sup>28,29</sup> neuroticism<sup>30</sup> and posttraumatic  
8 stress disorder (PTSD).<sup>31</sup> However, these findings have not always been replicated.<sup>32-34</sup>  
9 102 stress disorder (PTSD).<sup>31</sup> However, these findings have not always been replicated.<sup>32-34</sup>  
10  
11 103 These inconsistencies may be explained by, at least, three different factors. Firstly, in adults,  
12 higher levels of neuroticism are associated with an increased risk of depression,<sup>35</sup> anxiety<sup>36</sup> and  
13 104 higher levels of neuroticism are associated with an increased risk of depression,<sup>35</sup> anxiety<sup>36</sup> and  
14 PTSD after exposure to a traumatic event<sup>37</sup> and are a powerful predictor of comorbidity between  
15 105 PTSD after exposure to a traumatic event<sup>37</sup> and are a powerful predictor of comorbidity between  
16 depression and anxiety.<sup>38</sup> Neuroticism includes those personality traits that represent how some  
17 106 depression and anxiety.<sup>38</sup> Neuroticism includes those personality traits that represent how some  
18 people perceive the world around them as threatening or stressful. In addition, some personality  
19 107 people perceive the world around them as threatening or stressful. In addition, some personality  
20 traits also influence the individual tendency to be potentially exposed to stressful environments.  
21 108 traits also influence the individual tendency to be potentially exposed to stressful environments.  
22 Predisposed individuals may tend to choose environments prone to having a high risk of  
23 109 Predisposed individuals may tend to choose environments prone to having a high risk of  
24 exposure to stressful events. Specifically, this scenario, known as GxE correlation, may mediate  
25 110 exposure to stressful events. Specifically, this scenario, known as GxE correlation, may mediate  
26 the relationship between neuroticism and specific SLEs.<sup>39</sup> Secondly, the genetic factors  
27 111 the relationship between neuroticism and specific SLEs.<sup>39</sup> Secondly, the genetic factors  
28 influencing the level of neuroticism, including the *5-HTTLPR* polymorphism, are shared by  
29 112 influencing the level of neuroticism, including the *5-HTTLPR* polymorphism, are shared by  
30 persons having anxious-depressive spectrum disorders.<sup>38,40</sup> Lastly, GxE interactions have been  
31 113 persons having anxious-depressive spectrum disorders.<sup>38,40</sup> Lastly, GxE interactions have been  
32 described involving *5-HTTLPR* and depression,<sup>41</sup> anxiety<sup>42</sup> and PTSD.<sup>43</sup> Despite all of the above  
33 114 described involving *5-HTTLPR* and depression,<sup>41</sup> anxiety<sup>42</sup> and PTSD.<sup>43</sup> Despite all of the above  
34 evidence, genetic association and GxE interaction studies do not usually analyze or control for  
35 115 evidence, genetic association and GxE interaction studies do not usually analyze or control for  
36 the level of neuroticism in the relationship between *5-HTTLPR*, SLEs and anxious-depressive  
37 116 the level of neuroticism in the relationship between *5-HTTLPR*, SLEs and anxious-depressive  
38 spectrum disorders.  
39 117 spectrum disorders.  
40  
41 118 However, the question arising in this context is how environmental and genetic factors interact  
42 to produce a mental disorder.<sup>21,44</sup> In recent years, increasing interest in epigenetic factors  
43 119 to produce a mental disorder.<sup>21,44</sup> In recent years, increasing interest in epigenetic factors  
44 described in other human diseases has focused on its role in mental disorders.<sup>45</sup> The study of the  
45 120 described in other human diseases has focused on its role in mental disorders.<sup>45</sup> The study of the  
46 epigenome, changes in gene expression by modulating the accessibility of information that  
47 121 epigenome, changes in gene expression by modulating the accessibility of information that  
48 occurs without modifying the DNA sequence, suggests that, although inheritable, these changes  
49 122 occurs without modifying the DNA sequence, suggests that, although inheritable, these changes  
50 are not necessarily stable over the life span of individuals and can be modified under some  
51 123 are not necessarily stable over the life span of individuals and can be modified under some  
52 environmental stimuli that modulate the activity of the enzymes involved, opening new  
53 124 environmental stimuli that modulate the activity of the enzymes involved, opening new  
54 prospects for developing therapeutic approaches based on epigenetic mechanisms.<sup>46</sup> Epigenetic  
55 125 prospects for developing therapeutic approaches based on epigenetic mechanisms.<sup>46</sup> Epigenetic  
56 mechanisms have been associated with different mental disorders including depression,<sup>47</sup>  
57 126 mechanisms have been associated with different mental disorders including depression,<sup>47</sup>  
58  
59  
60

1  
2  
3 127 PTSD,<sup>48</sup> schizophrenia,<sup>49,50</sup> autism,<sup>49</sup> bipolar disorder<sup>50</sup> and alcohol dependence.<sup>51</sup> In fact,  
4  
5 128 epigenetic regulation of the glucocorticoid receptor signaling in neurons has been recently  
6  
7 129 shown to be the mechanism underlying GxE interactions to explain risk and resilience of PTSD  
8  
9 130 after SLE in childhood.<sup>52</sup>

10  
11 131 In order to integrate all these findings and create new opportunities and challenges offered by  
12  
13 132 the GxE interaction scenarios in the field of mental disorders, a multidisciplinary collaboration  
14  
15 133 between clinicians, epidemiologists, geneticists and statisticians offers greater  
16  
17 134 opportunities.<sup>20,23,53</sup> One of the proposed mechanisms for this collaboration includes carrying  
18  
19 135 out community psychiatric surveys and this has been facilitated by the possibility of obtaining  
20  
21 136 DNA and/or mRNA from peripheral tissues (blood, saliva or buccal cells). Population-based  
22  
23 137 surveys offer several advantages over other study designs to contribute to the clarification of the  
24  
25 138 GxE interactions in mental disorders.<sup>44,54,55</sup> Firstly, current knowledge of genes as risk factors is  
26  
27 139 based almost exclusively on clinical and non-representative population samples. Secondly, the  
28  
29 140 distribution of the gene polymorphisms of interest in the general population has not been well  
30  
31 141 investigated. Thirdly, this type of study can provide samples for future case-control studies and  
32  
33 142 can be the bases for future longitudinal ones. Finally, hypotheses generated from epidemiologic  
34  
35 143 surveys may contribute to test new basic studies and can be considered as a complementary  
36  
37 144 strategy to translational research.

#### 39 145 **Pegasus-Murcia Project**

40  
41 146 Spain actively participated in the ESEMeD Project with a representative sample of the adult  
42  
43 147 general Spanish population (n=5473) and the results have been published in national and  
44  
45 148 international journals.<sup>56-61</sup> However, the sample size within most of the Autonomous  
46  
47 149 Communities in Spain was too small to be able to achieve accurate and precise estimates at the  
48  
49 150 Regional level where Health Care policies are decided. Moreover, several differences between  
50  
51 151 the Autonomous Communities in Spain in important aspects related to mental health such as  
52  
53 152 socioeconomic<sup>62</sup> and territorial inequalities in health care supply and in long-term care, access  
54  
55 153 to and use of health care facilities,<sup>63</sup> premature deaths due to alcohol consumption<sup>64</sup> and the  
56  
57 154 prevalence of psychological distress<sup>65</sup> have recently been described.

1  
2  
3 155 Murcia is one of the 17 Autonomous Communities of Spain. It is located in the southeast of  
4  
5 156 the country on the Mediterranean coast, with a population of 1,424,063 inhabitants at the time  
6  
7 157 of the survey (INE 2008, National Statistical Institute of Spain), almost a third of them (30.7%)  
8  
9 158 living in the capital.

10  
11 159 The PEGASUS-Murcia (“Psychiatric Enquiry to General Population in Southeast Spain-  
12  
13 160 Murcia”) project has been designed in order to obtain regional data of the prevalence, burden  
14  
15 161 and care of a representative sample of the general adult population of Murcia to allow planning  
16  
17 162 of new regional mental health policies and to compare the results with the national sample of  
18  
19 163 Spain, Europe and all other countries participating in the WMH Survey Initiative. The project  
20  
21 164 also constitutes a unique opportunity to initiate a biological bank of a well-studied  
22  
23 165 representative sample of the general population.

## 24 25 166 **Objectives**

26  
27 167 The PEGASUS-Murcia project is a multi-purpose, observational, cross-sectional, comparative  
28  
29 168 study of the non-institutionalized general population of Murcia Region whose objective is to  
30  
31 169 improve knowledge about common psychiatric disorders in two main areas. The first is the  
32  
33 170 epidemiology of mental disorders and protective and risk factors in the general population of  
34  
35 171 Murcia. The specific objectives are: i) to estimate the one-month, 12-month and lifetime  
36  
37 172 prevalence of the most common mental disorders, specifically, mood and anxiety disorders, in  
38  
39 173 the general population of Murcia; ii) to assess the independent association of mood and anxiety  
40  
41 174 disorders with sociodemographic factors (gender, age, education and urban/rural location) and  
42  
43 175 selected risk factors (family history, childhood experiences, religion, partnership status and  
44  
45 176 sexual problems, among others); iii) to assess the quality of life of persons with the most  
46  
47 177 common psychiatric disorders and to analyze how other variables (physical medical conditions  
48  
49 178 and sociodemographic factors) may influence this outcome; iv) to assess the treatment for these  
50  
51 179 disorders and to evaluate the unmet need and the quality of care received; and v) to compare our  
52  
53 180 results with those obtained from Spain, Europe and other non-European countries, including the  
54  
55 181 United States.



1  
2  
3 182 The second objective is the genetic, epigenetic and transcriptomic influences associated with  
4  
5 183 mental disorders. Its specific aims include: i) the estimation of the distribution of different  
6  
7 184 candidate genes in the general population and their association with different psychiatric  
8  
9 185 disorders; ii) the identification of sensitive alleles underlying potential GxE interactions and the  
10  
11 186 study of epigenetic mechanisms involved, specifically, DNA methylation; and iii) the analysis  
12  
13 187 of gene expression alterations through transcriptomic assays.

## 14 15 188 **METHODS AND ANALYSIS**

### 16 17 18 189 **Study Design**

19  
20 190 The project is a cross-sectional face-to-face interview survey based on a representative sample  
21  
22 191 of the adult and non-institutionalized general population of the Murcia Region. Those who  
23  
24 192 complete the interview will be invited to provide two biological samples from their oral mucous  
25  
26 193 membranes. The target population is defined as persons aged 18 or older residing in Murcia, not  
27  
28 194 living in institutions and with an active health card (defined as persons included in PERSAN, a  
29  
30 195 regional registry that contains all residents with a Health Card which is periodically up-dated.  
31  
32 196 Exclusion criteria are: i) Confirmed irretrievable contact errors (e.g. telephone number and/or  
33  
34 197 address); ii) Institutionalized individuals (e.g. in prison, in a hospital or in another institution) or  
35  
36 198 those living outside the Autonomous Community during the survey field work; and iii)  
37  
38 199 individuals not able to understand the Spanish language or not able to conduct the questionnaire  
39  
40 200 due to his/her physical or mental condition.

### 41 42 201 **Sampling plan**

43  
44 202 The geographical area of the survey is the Murcia Region and a two-stage, stratified sampling  
45  
46 203 design has been used. The primary sampling unit is the Primary Health Centre and the second is  
47  
48 204 the individual. The sampling frame has been PERSAN, the regional health care population  
49  
50 205 database in Murcia. Primary Health Centres have been grouped into nine strata, the current  
51  
52 206 Health Care Areas in Murcia Region. The initial sample size was 4,500 adult individuals  
53  
54 207 divided into the nine Health Care Areas with proportionate allocation. A representative sample  
55  
56 208 of two centres has been chosen in each health area, without individual participant replacement.

1  
2  
3 209 Selection probability for each centre was known a priori and it was proportional to the size of  
4  
5 210 the centre (% of adult individuals registered in the centre) and the proportion of adult  
6  
7 211 individuals in the centre whose place of residence was rural, semi-urban or urban. Within each  
8  
9 212 of the two selected Health Centers, a stratified random sample procedure, performed for each  
10  
11 213 combination of gender, age group (18-24, 25-34, 35-49, 50-64 and 65+) and type of residence  
12  
13 214 (rural, semi-urban and urban), constitutes a stratum and individuals have been selected using  
14  
15 215 simple random sampling.

16  
17 216 For each Health Care Area, the sample size of each stratum has been selected such that  
18  
19 217 individuals in with the same demographic characteristics had equal probability of being selected  
20  
21 218 independently of the selected center. If a high number of those fulfilling the exclusion criteria in  
22  
23 219 one area is reached, a fixed number of additional individuals will be released (subsequent  
24  
25 220 releases), according to the number of interviews completed in the area and following the same  
26  
27 221 selection procedure within each of the centers as the ones used to select the initial release (no  
28  
29 222 new center will be selected for these releases). Any replacement of those persons who do not  
30  
31 223 want to collaborate or who do not meet the non-eligibility criteria is not allowed.

#### 32 33 224 **Survey procedures and data control**

34  
35 225 Those selected will receive no financial incentive to participate and there will be no individual  
36  
37 226 replacement procedure. Questions are asked by trained lay interviewers using the computer-  
38  
39 227 assisted personal interview (CAPI) that was programmed centrally using the Blaise software  
40  
41 228 system. This is an interviewing application developed by Statistics Netherlands (Herleen, the  
42  
43 229 Netherlands) and designed to ease the handling of elaborate skip and complex randomization  
44  
45 230 patterns and to facilitate data entry, allow the elaboration of some questions and direct the  
46  
47 231 interviewer through the questioning sequence.

48  
49 232 Periodically, the completed interviews will be submitted to the central project Data Center  
50  
51 233 (Regional Mental Health Service, Murcia-Spain) for checking and storage following a  
52  
53 234 predetermined security procedure. All raw data will be transferred to the Hospital del Mar  
54  
55 235 Medical Research Institute (IMIM) and the Department of Health Care Policy at Harvard  
56  
57 236 University, coordinating centers of the ESEMeD and WMH Survey Initiative projects  
58  
59  
60

1  
2  
3 237 respectively, via secure websites. The database has been declared to the Spanish Data Protection  
4  
5 238 Agency.

6  
7 239 A survey firm has been contracted to undertake the fieldwork and, in order to ensure the quality  
8  
9 240 of the survey, several strategies are being implemented: i) a one week training course for all  
10  
11 241 interviewers by WHO certified trainers on the original protocol and use of the CAPI version of  
12  
13 242 the CIDI; ii) development of a written manual to standardize the interviewing procedure and all  
14  
15 243 scientific and administrative elements that could affect comparability of data; iii) regular  
16  
17 244 meetings with the survey firm to ensure adherence to the protocol and to deal with any difficulty  
18  
19 245 that may have arisen; and iv) data quality analysis to detect any inconsistencies and/or  
20  
21 246 incomplete data.

22  
23 247 The survey firm has been provided with sufficient data to allow contact with each of the  
24  
25 248 individuals of the selected sample and only after 10 unsuccessful attempts the person will be  
26  
27 249 considered to be uncontactable or after confirmation that the selected person does not live at that  
28  
29 250 address and new contact information is unavailable. Several methods will be used to improve  
30  
31 251 the participation of those selected: i) an informative flyer providing general information related  
32  
33 252 to the project and giving notice of future contact will be sent by conventional post together with  
34  
35 253 an invitation letter signed by a person from the Health Care Authority; ii) a phone call to invite  
36  
37 254 them to participate in the interview process and to offer them the possibility to do the interview  
38  
39 255 either at home or in their Primary Care Center; iii) Several informative sessions for the  
40  
41 256 healthcare personnel of the Primary Care Centers will be organized to facilitate their  
42  
43 257 collaboration should the participants ask them about the project; iv) During the period when the  
44  
45 258 interviews will take place, some official posters will be put in public centers to inform people  
46  
47 259 about the project; v) All interviewers will be provided with an official identification and have  
48  
49 260 been trained on how to explain the institutional nature of the research project.

50  
51  
52 **261 The survey questionnaire**

53  
54 262 The questionnaire used in the PEGASUS-Murcia project is a revised version of the CIDI which,  
55  
56 263 together with diagnostic information on the most common mental disorders, also includes  
57  
58  
59  
60

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

266 *The Composite International Diagnostic Interview (CIDI)*

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

278 **Table 1: Description of the adapted version of the World Health Organization -Composite**  
 279 **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 <sup>‡</sup>	Psychosis	42 to 84	All respondents
CFQ <sup>§</sup>	Risk Factors	25	All respondents
SLE <sup>¶</sup>	Risk Factors	13 to 39	All respondents
Neuroticism and Extroversion subscales <sup>¥</sup>	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 Childhood	Socio-demographics	25-36	Long path
Attention Hyperactivity	Risk/Protective Factors	110	Long path
Oppositional Defiant	Childhood	90	Long path and Screening
Conduct Disorder	Childhood	46	Long path and Screening
Separation Anxiety Disorder	Childhood	54	Long path
Family Burden	Childhood	86	Screening questions (SCR)
Quality/ Lie subscale	Risk Factors	40	Long path
	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

280 <sup>§</sup>EQ-5D: European Quality of Life Scale; <sup>¶</sup>SF-12 v2: Short Form 12 Health Questionnaire; <sup>†</sup>Lie subscale of the abbreviated version  
 281 of the Eysenck Personality Questionnaire (EPQR-A); <sup>‡</sup>CAPE: Community Assessment of Psychic Experiences; <sup>§</sup>CFQ: Cognitive  
 282 Failure Questionnaire; <sup>¶</sup>SLE: Stressful Life Events; <sup>¥</sup>Neuroticism and Extroversion subscales of the abbreviated version of the  
 283 Eysenck Personality Questionnaire (EPQR-A)  
 284 \* Long Path inclusion criteria: a) all individuals that could be considered as “high risk individuals”, because they had positively  
 285 answered a number of specific questions related to mood and anxiety disorders, and b) a random subsample (25%) of the  
 286 respondents without symptoms (“low risk individuals”). The remaining 75% of respondents without screening symptoms not  
 287 randomly selected for the long path followed the Short Path of the questionnaire  
 288

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

1  
2  
3 297 Pencil Interviewing) and the computerized form or CAPI (Computer Assisted Personal  
4  
5 298 Interviewing), designed to ease the handling of elaborate skip and complex randomization  
6  
7 299 patterns and to facilitate data entry with a resulting reduction in interview time and errors in data  
8  
9 300 collection and recording. The original Spanish CAPI version used in Spain had not been  
10  
11 301 updated since it was used in the context of the ESEMeD project almost ten years ago. Since  
12  
13 302 then, all improvements in the questionnaire have only been added to the CIDI Latin American  
14  
15 303 (LA) v20.0 version. However, due to linguistic and cultural differences in Spanish-speaking  
16  
17 304 populations, this CAPI version had to be culturally adapted for use in Spain by our research  
18  
19 305 team and this process is fully described elsewhere.<sup>71</sup>

20  
21 306 To further shorten the length of the questionnaire, some sections were not selected for the  
22  
23 307 purposes of this project. These include Intermittent Explosive Disorder, Personality I and II,  
24  
25 308 Neurasthenia and Pre-Menstrual and Gambling sections. Some others were substituted by other  
26  
27 309 questions or questionnaires, e.g. the Tobacco Use section was simplified using some questions  
28  
29 310 obtained from the Spanish National Health Survey and the Psychosis section with the CAPE  
30  
31 311 instrument (Community Assessment of Psychic Experiences), both described below.

### 32 312 *Other study instruments*

33  
34  
35 313 Several other instruments were added to the original CIDI for the specific purposes of the  
36  
37 314 PEGASUS-Murcia project. These include the Spanish version of different questionnaires: i)  
38  
39 315 Mini-Mental State Examination for interviewees older than 60 years old;<sup>72,73</sup> ii) the Cognitive  
40  
41 316 Failure Questionnaire (CFQ);<sup>74,75</sup> iii) the Neuroticism, Extroversion and Lie subscales of the  
42  
43 317 abbreviated version of the Eysenck Personality Questionnaire (EPQR-A);<sup>76-78</sup> iv) the Resilience  
44  
45 318 Scale;<sup>79,80</sup> v) the Community Assessment of Psychic Experiences (CAPE)<sup>81</sup> to measure  
46  
47 319 attenuated psychotic symptoms in the general population instead of the Psychosis section of the  
48  
49 320 CIDI, as the latest is only used as a screening instrument in the detection of psychosis. Those  
50  
51 321 who positively answer two items of the positive dimension with a score equal or superior to 3,  
52  
53 322 have been hospitalized for psychiatric reasons and/or have received psychotropic medication  
54  
55 323 during the last year will be evaluated by a clinic psychiatrist with the module C (Psychotic  
56  
57 324 Disorders) of the SCID (Structured Clinical Interview for DSM Disorders) ; vi) a brief list of 12  
58  
59  
60

1  
2  
3 325 stressful life events in the last 12 months was included by the combination of a List of  
4  
5 326 Threatening Experiences (LTE)<sup>82,83</sup> and the emotional and life-changing impact of each event;<sup>84</sup>  
6  
7 327 vii) the European Quality of Life Scale (EuroQol 5D)<sup>85</sup> and the Short Form 12 Health  
8  
9 328 Questionnaire (SF-12 v2);<sup>86</sup> viii) an ad-hoc questionnaire of partner violence obtained from the  
10  
11 329 Spanish National Health Survey and from the regional mental health clinical guidelines;<sup>87</sup> and,  
12  
13 330 finally, ix) some questions related to tobacco use and physical exercises from the Spanish  
14  
15 331 National Health Survey.

### 332 **Questionnaire pathways**

333 In order to optimize the duration of the interview, the WMH questionnaire was divided into two  
334 parts with questions in Part 1 administered to all respondents and those in Part 2 only to a  
335 subsample of individuals who followed the long path of the interview. Part 2 of the interview  
336 includes detailed information about a wide range of aspects related to the primary disorders and  
337 also to mental disorders of secondary interest (Table 1). The inclusion criteria for the long path  
338 are: a) all individuals that could be considered as “high risk individuals” because they positively  
339 answer a number of specific questions related to mood and anxiety disorders and b) a random  
340 subsample (25%) of the respondent without symptoms (“low risk individuals”). The remaining  
341 75% of respondents without screening symptoms not randomly selected for the long path  
342 followed the short path. All these pathways are automatically made by the computer without  
343 any intervention of the interviewer. In this short itinerary, the sections omitted were substituted  
344 by a specific section that included those questions needed to calculate some demographic  
345 indicators. Moreover, two sections were only used in a percentage of the long path itinerary,  
346 Eating Disorders (50 %) and Obsessive-Compulsive Disorder (33%).

### 347 **Quality control procedures**

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

1  
2  
3 353 illness. All these procedures will be verified by the coordinating centers and the final document  
4  
5 354 included several aspects, for example, sample releases, the duration of the interviews and the  
6  
7 355 proportion of positive responses to selected screening questions. Local members of the research  
8  
9 356 team will be responsible for verifying the informed consent forms and the quality checking  
10  
11 357 following computerized protocols. These procedures are similar to those implemented in the  
12  
13 358 ESEMeD project and are fully described elsewhere.<sup>8</sup> Briefly, they consist of checks of  
14  
15 359 individual pieces of information from the interviewees, for example, completion status,  
16  
17 360 consistency across the questionnaire, questionnaire itinerary and length of the interview, and  
18  
19 361 from the interviewers, number of disorders screened positively, verification of a random  
20  
21 362 selection of almost 1% of interviews completed by a telephone contact to confirm the interview  
22  
23 363 and some aspects related to it such as place, approximate duration and identification of the  
24  
25 364 interviewer.

### 27 365 **Laboratory Methods**

28  
29 366 On completion of the interview, two biological samples will be obtained from each interviewee,  
30  
31 367 one for DNA extraction for genomic and epigenomic analysis and the other one to obtain  
32  
33 368 mRNA for gene expression quantification (transcriptomic assays). These samples will be taken  
34  
35 369 using swabs compatible with molecular amplification techniques, as they do not interfere with  
36  
37 370 the amplification process (FLOQSwabs Flocked Swabs, Copan Flock Technologies srl).

38  
39 371 Samples for DNA extraction will be collected in sterile 1.5 ml tubes. Those to be used for RNA  
40  
41 372 extraction will be harvested in dark sterile tubes containing RNA protect cell Reagent  
42  
43 373 (QIAGEN, Hilden, Germany), which provides immediate stabilization of RNA. Cells will be  
44  
45 374 thus stabilized at room temperature and can then be stored or transported at ambient temperature  
46  
47 375 prior to RNA purification. Tubes will be labeled with tags (14C.B. 40X40 type) with a specific  
48  
49 376 code for each sample and will be packaged and sent to BIOBANC-Mur (the biobank for  
50  
51 377 biomedical research network of the Region of Murcia, RD09/0076/00065, as a partner of the  
52  
53 378 Spanish National Biobanks Network; IMIB: Instituto Murciano de Investigación Biosanitaria)  
54  
55 379 according to current Spanish legislation and following the regulations of the International Air  
56  
57 380 Transport Association (IATA) on biological sample shipping.



1  
2  
3 381 Those sample accepted by BIOBANC-Mur will be registered using a specific biobanking  
4  
5 382 software (bio-e-bank, VITROSOFT, SL), as part of a Laboratory Integrated Management  
6  
7 383 System (LIMS). The nucleic acid extraction will be performed automatically (QIAcube system;  
8  
9 384 QIAGEN, Hilden, Germany) to minimize variability due to manual handling using QIAamp  
10  
11 385 DNA Blood Mini Kit and RNeasyPlus Mini Kit (QIAGEN, Hilden, Germany) for DNA and  
12  
13 386 RNA extraction, respectively. The synthesis of complementary DNA (cDNA) from mRNA for  
14  
15 387 expression studies will be developed for all samples by reverse transcription using the High  
16  
17 388 Capacity cDNA Reverse Transcription Kit (Applied Biosystems). All processes will be  
18  
19 389 performed according to the manufacturer's instructions.

20  
21 390 DNA and RNA will be quantitated by measuring absorbance at 260/280 nm using a  
22  
23 391 spectrophotometer.<sup>88-90</sup> Between 260 nm and 230 nm (A<sub>260/230</sub>) absorbance is commonly used  
24  
25 392 as a secondary indicator of nucleic acid purity<sup>91-93</sup>. The integrity of DNA will be visualized by  
26  
27 393 electrophoresis on 1% agarose gel (migration for 1 hour at 100 V) using 100 ng of total DNA  
28  
29 394 and a 23 kb DNA ladder (Lambda DNA/HindIII Marker (Thermo Fisher Scientific) as DNA  
30  
31 395 marker. All mRNA samples will be transformed into cDNA.

32  
33 396 Specially trained technicians will be used to monitor the specimen collection by donors and to  
34  
35 397 perform sample manipulations in order to minimize variability of results and to obtain the  
36  
37 398 optimal quality of nucleic acids for this and future studies. The processed biospecimens (150 µl  
38  
39 399 of DNA and 80 µl of cDNA) will be stored in 750 µl microtubes in an ultra-freezer at -80 °C  
40  
41 400 located in BIOBANC-Mur.

#### 42 43 401 **Statistical methods**

44  
45 402 The expected response-rate (RR) has been set to a minimum of 65%, based on a previous  
46  
47 403 regional community survey which included the donation of blood samples<sup>94,95</sup>. The response  
48  
49 404 rate will be calculated based on the proportion of people interviewed and was defined as the  
50  
51 405 number of completed interviews divided by the total number of cases minus the number of non-  
52  
53 406 eligible cases.

#### 54 55 407 **Weighting procedures**

1  
2  
3 408 Given that the interview is divided into two parts and only a portion of the sample will be  
4  
5 409 selected for the second part, two types of weightings are considered to estimate population  
6  
7 410 parameters. The first is to weight for the probability of selection for each Health Care Area,  
8  
9 411 Health Center and demographic stratum and the second is for the random skips included in the  
10  
11 412 questionnaire. The method designed is described in Box 1.  
12

13 **BOX 1: Weighting procedures**

14 **First weighting procedure:**

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

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

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

25 Step 4) Post-stratification weight ( $w_{ps}$ ): data on population of the region of Murcia by sex, age and  
26 Healthcare Area were provided by the CREM (Centro Regional de Estadística de Murcia;  
27 Padrón 2010 ([http://www.carm.es/econet/sicrem/PU\\_padron/](http://www.carm.es/econet/sicrem/PU_padron/)). The population for the age group  
28 18-24 has been estimated as the population for the age group 18-19 plus the population for the  
29 age group 20-24. The population for the age group 18-19 has been estimated as the population  
30 for the age group 15-19 times the proportion of population aged 18-19 in the age group 15-19 in  
31 Murcia: 0.4116 for males and 0.4165 for females. A post-stratification weight was created to  
32 ensure that the joint distribution of the post-stratifying variables Healthcare Area, sex and age  
33 group matches the known population joint distribution of Murcia.

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

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

37 Step 7) Trimmed weight ( $w_{trim}$ ): trim the normalized weight obtained from step 6. The upper and lower  
38 5% were trimmed to the mean of each tail.

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

40 **Second weighting procedure:**

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

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

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

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

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

56 Step 5) Part 2 Normalized weight:  $w_{p2norm} = w_{p2adj} n / \sum_{i=1}^n w_{p2adj_i}$ .

#### 414 **Analysis of the data and forthcoming research projects**

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

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

#### 431 **ETHICS AND DISSEMINATION**

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

1  
2  
3 440 international comparisons. The study findings will be submitted to peer-reviewed journals for  
4  
5 441 publication, and presented at national and international scientific meetings.  
6

## 7 442 **DISCUSSION**

8  
9  
10 443 The epidemiology of mental illnesses is a fascinating but highly complex area of research. This  
11  
12 444 complexity is primarily due to the wide range of factors, environmental and genetic, which  
13  
14 445 combines to produce a recognized psychiatric disorder. Previous epidemiological research has  
15  
16 446 resulted in the production of a great amount of data but it has been difficult to make cross-  
17  
18 447 national comparisons due to methodological variability. The WMH Survey Initiative aimed to  
19  
20 448 address this issue by using an international standardized protocol, allowing comparisons of the  
21  
22 449 most common mental disorders and their associated factors throughout the world. Using this  
23  
24 450 study design, it therefore offers the opportunity for new surveys to be performed in the context  
25  
26 451 of an international collaborative initiative and the possibility to adapt the questionnaire  
27  
28 452 according to the specific aims of the research being undertaken. The PEGASUS-Murcia project  
29  
30 453 can be considered as an example of how the latter has been successfully achieved. It is a cross-  
31  
32 454 sectional study designed to assess the prevalence of the most frequent mental disorders and their  
33  
34 455 correlates in a representative sample of the general population of Murcia. Its primary strengths  
35  
36 456 are: i) the fact that it was specifically adapted to assess factors not only associated with mental  
37  
38 457 disorders but also with positive mental health in a representative sample of the general  
39  
40 458 population; ii) its context focused on regional needs where healthcare decisions are taken  
41  
42 459 regarding resource allocation and mental health planning; iii) the collection of biological  
43  
44 460 samples not only for DNA analysis but also for mRNA; iv) all the information collected in our  
45  
46 461 study, including biological samples, can be correlated with past and future health events because  
47  
48 462 all Spanish population had free access to the Healthcare System at the time of its inception and  
49  
50 463 were thus registered and provided with a unique identification number and therefore; v) finally,  
51  
52 464 the inclusion of a multidisciplinary research team is in accordance with the international  
53  
54 465 consensus regarding the need for interdisciplinary collaboration between clinicians,  
55  
56 466 epidemiologists and neuroscience researchers to increase their combined efforts to study the  
57  
58  
59  
60

1  
2  
3 467 complex gene-gene and gene-environmental interactions underlying mental health  
4  
5 468 disorders.<sup>23,55,96,97</sup>

6  
7 469 Concerns have been expressed about the cost-effectiveness of psychiatric epidemiological  
8  
9 470 surveys, such as World Mental Health 2000 (WMH-2000) projects,<sup>98</sup> an example being the  
10  
11 471 rationale for starting a new psychiatric epidemiological survey in the Autonomous Community  
12  
13 472 of Murcia if Spain had already participated in the ESEMeD project. However, there are several  
14  
15 473 reasons to justify this regional initiative. Firstly, public health and healthcare agencies usually  
16  
17 474 allocate mental health resources, including human, based on data from national epidemiologic  
18  
19 475 surveys,<sup>99</sup> such as that provided by the Spanish participation in the ESEMeD Project. As  
20  
21 476 previously mentioned, the involvement of the Region of Murcia in the Spanish ESEMeD survey  
22  
23 477 did not allow evaluation of specific regional data. Nowadays, the main responsibility for  
24  
25 478 planning and management of Healthcare resources in Spain lies with the Autonomous  
26  
27 479 Communities and differences exist between them in terms of accessibility, amount of health-  
28  
29 480 care resources and political decision-making.<sup>62-65</sup> Devolution of this responsibility to Murcia  
30  
31 481 occurred in December 2001.

32  
33 482 Secondly, the inclusion of biological data in a well-designed multidisciplinary epidemiological  
34  
35 483 study offers great advantages in terms of a more global understanding of mental disorders.  
36  
37 484 These are complex illnesses of the brain where social, familial, psychological and biological  
38  
39 485 elements interact throughout the entire life of a person to influence his/her risk of developing a  
40  
41 486 mental health disorder. To extend our understanding of the physiopathology and epidemiology  
42  
43 487 of the more common ones (mood and anxiety), it is necessary to identify genetic loci and  
44  
45 488 polymorphic alleles and their distribution in the healthy and affected population whose function  
46  
47 489 in determining risk for, and protection against, these conditions probably depends on gene-gene  
48  
49 490 and GxE interactions. The collection of genetic material from representative samples from the  
50  
51 491 general population, well described using international diagnostic instruments such as CIDI,  
52  
53 492 offers new and different possibilities to evaluate candidate genes in non-biased samples and to  
54  
55 493 describe their distribution in the general population that may contribute to clarification of the  
56  
57 494 complexity of mental disorders.

1  
2  
3 495 Thirdly, our project involving a multidisciplinary research team gives new opportunities to  
4  
5 496 develop different study designs that can move from descriptive to analytical epidemiology. For  
6  
7 497 example, this representative sample constitutes a good source of controls for future case-control  
8  
9 498 studies, where cases will be provided from the public health care clinics, and can be the starting  
10  
11 499 point for future cohort studies. Our project was designed to allow for all these possibilities.

### 12 **500 Limitations of the study**

13  
14  
15 501 Currently, the main limitations of the PEGASUS-Murcia project are related to: i) the cross-  
16  
17 502 sectional design which, while it allows association studies, limits the possible causal  
18  
19 503 interpretation of the findings. However, these findings may provide new hypotheses and enable  
20  
21 504 the design of new studies; ii) not all interviewees will provide biological samples and this may  
22  
23 505 affect the representativeness of some mental disorders in future analyses. To determine if this  
24  
25 506 will result in selection bias, we will analyze whether there are distinguishing characteristics  
26  
27 507 between donors and non-donors in the distribution of mental disorders and other characteristics  
28  
29 508 of the participants; iii) the population stratification in our study which will be used for future  
30  
31 509 genetic association analyses is performed by using the stated ancestral origin by participants<sup>100</sup>  
32  
33 510 instead of using genetic markers; and iv) biological samples will be obtained from oral mucosal  
34  
35 511 scrapings and not from brain neurons. However, this is a general situation given the ethical  
36  
37 512 issues and difficulties in obtaining neural tissues and, in any case, gene expression does not  
38  
39 513 appear to be specific to neural tissue, at least in some genes that have ubiquitous expression, for  
40  
41 514 example, 5-HTTLPR.<sup>101-104</sup>

### 42 **515 Conclusions and Future Directions**

43  
44  
45 516 The PEGASUS-Murcia project is a sound bases for multidisciplinary collaborative mental  
46  
47 517 health research studies which will provide not only a huge amount of epidemiological  
48  
49 518 information but will also offer exiting opportunities to clarify the complex interactions between  
50  
51 519 genetic and environmental factors which result in a range of mental health disorders.

### 52 **520 Competing interests**

53  
54  
55 521 The authors declare that they have no competing interests.  
56  
57  
58  
59  
60

## 522 Author's contributions

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

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

539 The PEGASUS-Murcia project was supported by the Regional Health Authorities of Murcia  
540 (“Servicio Murciano de Salud and Consejería de Sanidad y Política Social”) (Decreto n<sup>o</sup>  
541 455/2009), the “Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de  
542 Murcia” (N<sup>o</sup> Expedientes: CM0829 I and FFIDS/EMER09/14) and the “Ayudas para proyectos  
543 de Investigación en Salud –ISCIH- del Plan Nacional de Investigación Científica, Desarrollo e  
544 Innovación Tecnológica” (PI12/00809). None of the funders had any role in the design,  
545 analysis, interpretation of results, or preparation of this paper. The PEGASUS-Murcia project is  
546 carried out in conjunction with the WHOWMH Survey Initiative. We thank the WMH  
547 Coordinating Center staff at Harvard and Michigan Universities and, especially Professor Ron  
548 Kessler, for their assistance with the instrumentation, fieldwork and data analysis. These  
549 activities were supported by the United States National Institute of Mental Health  
550 (R01MH070884), the John D. and Catherine T. MacArthur Foundation, the Pfizer Foundation,  
551 the U.S. Public Health Service (R13-MH066849, R01-MH069864, and R01 DA016558), the  
552 Fogarty International Center (FIRCA R03- TW006481), the Pan American Health Organization,  
553 the Eli Lilly & Company Foundation, Ortho-McNeil Pharmaceutical, Inc., GlaxoSmithKline,  
554 Bristol-Myers Squibb and Shire. A complete list of WMH publications can be found at  
555 <http://www.hcp.med.harvard.edu/wmh/>.

556  
557  
558  
559  
560

1  
2  
3 558 **Figure 1: Flow chart of the PEGASUS-Murcia project**  
4 559  
5 560  
6 561  
7 562  
8 563  
9 564  
10 565  
11 566  
12 567  
13 568  
14 569  
15 570  
16 571  
17 572  
18 573  
19 574  
20 575  
21 576  
22 577  
23 578  
24 579  
25 580  
26 581  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

† 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.

- 1  
2  
3 608 10. Alonso J, Angermeyer MC, Bernert S et al. Prevalence of mental disorders in Europe:  
4 609 results from the European Study of the Epidemiology of Mental Disorders (ESEMeD)  
5 610 project. *Acta Psychiatr Scand Suppl.* 2004;21-27.  
6  
7  
8  
9 611 11. Alonso J, Angermeyer MC, Bernert S et al. 12-Month comorbidity patterns and  
10 612 associated factors in Europe: results from the European Study of the Epidemiology of  
11 613 Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl.* 2004;28-37.  
12  
13 614 12. Alonso J, Angermeyer MC, Bernert S et al. Disability and quality of life impact of  
14 615 mental disorders in Europe: results from the European Study of the Epidemiology of  
15 616 Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand.* 2004;420:38-46.  
16  
17  
18  
19 617 13. Alonso J, Angermeyer MC, Bernert S et al. Use of mental health services in Europe:  
20 618 results from the European Study of the Epidemiology of Mental Disorders (ESEMeD)  
21 619 project. *Acta Psychiatr Scand Suppl.* 2004;47-54.  
22  
23  
24  
25 620 14. Alonso J, Buron A, Bruffaerts R et al. Association of perceived stigma and mood and  
26 621 anxiety disorders: results from the World Mental Health Surveys. *Acta Psychiatr Scand.*  
27 622 2008;118:305-314.  
28  
29  
30  
31 623 15. Alonso J, Codony M, Kovess V et al. Population level of unmet need for mental  
32 624 healthcare in Europe. *Br J Psychiatry.* 2007;190:299-306.  
33  
34  
35 625 16. Codony M, Alonso J, Almansa J et al. Perceived need for mental health care and service  
36 626 use among adults in Western Europe: results of the ESEMeD project. *Psychiatr Serv.*  
37 627 2009;60:1051-1058.  
38  
39  
40  
41 628 17. Sevilla-Dedieu C, Kovess-Masfety V, Gilbert F et al. Mental health care and out-of-  
42 629 pocket expenditures in Europe: results from the ESEMeD project. *J Ment Health Policy*  
43 630 *Econ.* 2011;14:95-105.  
44  
45  
46  
47  
48 631 18. Hyman SE. The millennium of mind, brain, and behavior. *Arch Gen Psychiatry.*  
49 632 2000;57:88-89.  
50  
51  
52  
53 633 19. Kubota T, Miyake K, Hirasawa T. Epigenetic understanding of gene-environment  
54 634 interactions in psychiatric disorders: a new concept of clinical genetics. *Clin*  
55 635 *Epigenetics.* 2012;4:1.  
56  
57  
58  
59  
60

- 1  
2  
3 636 20. Merikangas KR, Zhang H, Avenevoli S et al. Longitudinal trajectories of depression  
4  
5 637 and anxiety in a prospective community study: the Zurich Cohort Study. *Arch Gen*  
6  
7 638 *Psychiatry*. 2003;60:993-1000.
- 8  
9 639 21. Toyokawa S, Uddin M, Koenen KC et al. How does the social environment 'get into the  
10  
11 640 mind'? Epigenetics at the intersection of social and psychiatric epidemiology. *Soc Sci*  
12  
13 641 *Med*. 2012;74:67-74.
- 14  
15 642 22. Tsuang M, Bar J, Stone W et al. Gene-environment interactions in mental disorders.  
16  
17 643 *World Psychiatry*. 2004;3:73-83.
- 18  
19 644 23. Caspi A, Moffitt TE. Gene-environment interactions in psychiatry: joining forces with  
20  
21 645 neuroscience. *Nat Rev Neurosci*. 2006;7:583-590.
- 22  
23 646 24. Caspi A, Hariri AR, Holmes A et al. Genetic sensitivity to the environment: the case of  
24  
25 647 the serotonin transporter gene and its implications for studying complex diseases and  
26  
27 648 traits. *Am J Psychiatry*. 2010;167:509-527.
- 28  
29 649 25. Koenen KC, Fu QJ, Ertel K et al. Common genetic liability to major depression and  
30  
31 650 posttraumatic stress disorder in men. *J Affect Disord*. 2008;105:109-115.
- 32  
33 651 26. Norris FH, Friedman MJ, Watson PJ et al. 60,000 disaster victims speak: Part I. An  
34  
35 652 empirical review of the empirical literature, 1981-2001. *Psychiatry*. 2002;65:207-239.
- 36  
37 653 27. Caspi A, Hariri AR, Holmes A et al. Genetic sensitivity to the environment: the case of  
38  
39 654 the serotonin transporter gene and its implications for studying complex diseases and  
40  
41 655 traits. *Am J Psychiatry*. 2010;167:509-527.
- 42  
43 656 28. Clarke H, Flint J, Attwood AS et al. Association of the 5-HTTLPR genotype and  
44  
45 657 unipolar depression: a meta-analysis. *Psychol Med*. 2010;40:1767-1778.
- 46  
47 658 29. Karg K, Burmeister M, Shedden K et al. The serotonin transporter promoter variant (5-  
48  
49 659 HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic  
50  
51 660 moderation. *Arch Gen Psychiatry*. 2011;68:444-454.
- 52  
53 661 30. Takano A, Arakawa R, Hayashi M et al. Relationship between neuroticism personality  
54  
55 662 trait and serotonin transporter binding. *Biol Psychiatry*. 2007;62:588-592.

- 1  
2  
3 663 31. Koenen K, Kilpatrick D, Acierno R et al. Serotonin transporter gene and social support  
4 moderate post-traumatic stress disorder in hurricane exposed adults. *American Journal*  
5 664 *of Medical Genetics Part B-Neuropsychiatric Genetics*. 2006;141B:742-743.  
6 665  
7  
8  
9 666 32. Risch N, Herrell R, Lehner T et al. Interaction between the serotonin transporter gene  
10 (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis. *JAMA*.  
11 667 2009;301:2462-2471.  
12 668  
13  
14 669 33. Munafò MR, Durrant C, Lewis G et al. Gene X environment interactions at the  
15 serotonin transporter locus. *Biol Psychiatry*. 2009;65:211-219.  
16 670  
17  
18 671 34. Navarro-Mateu F, Escámez T, Koenen KC et al. Meta-Analyses of the 5-HTTLPR  
19 Polymorphisms and Post-Traumatic Stress Disorder. *PLOS One*. 2013;8:e66227.  
20 672  
21  
22 673 35. Kendler KS, Neale MC, Kessler RC et al. A longitudinal twin study of personality and  
23 major depression in women. *Arch Gen Psychiatry*. 1993;50:853-862.  
24 674  
25  
26 675 36. Kotov R, Gamez W, Schmidt F et al. Linking "big" personality traits to anxiety,  
27 depressive, and substance use disorders: a meta-analysis. *Psychol Bull*. 2010;136:768-  
28 676 821.  
29  
30 677  
31  
32 678 37. Wu D, Yin H, Xu S et al. Risk factors for posttraumatic stress reactions among Chinese  
33 students following exposure to a snowstorm disaster. *BMC Public Health*. 2011;11:96.  
34 679  
35  
36 680 38. Hettema JM, Neale MC, Myers JM et al. A population-based twin study of the  
37 relationship between neuroticism and internalizing disorders. *Am J Psychiatry*.  
38 681 2006;163:857-864.  
39 682  
40  
41 683 39. Ploubidis GB, Frangou S. Neuroticism and psychological distress: to what extent is  
42 their association due to person-environment correlation? *European Psychiatry*.  
43 684 2011;26:1-5.  
44 685  
45  
46 686 40. Skelton K, Ressler KJ, Norrholm SD et al. PTSD and gene variants: new pathways and  
47 new thinking. *Neuropharmacology*. 2012;62:628-637.  
48 687  
49  
50 688 41. Surtees PG, Wainwright NW, Willis-Owen SA et al. Social adversity, the serotonin  
51 transporter (5-HTTLPR) polymorphism and major depressive disorder. *Biol Psychiatry*.  
52 689 2006;59:224-229.  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 691 42. Stein MB, Schork NJ, Gelernter J. Gene-by-environment (serotonin transporter and  
4 childhood maltreatment) interaction for anxiety sensitivity, an intermediate phenotype  
5 692  
6 for anxiety disorders. *Neuropsychopharmacology*. 2008;33:312-319.  
7 693  
8  
9 694 43. Koenen KC, Aiello AE, Bakshis E et al. Modification of the association between  
10 serotonin transporter genotype and risk of posttraumatic stress disorder in adults by  
11 695  
12 county-level social environment. *Am J Epidemiol*. 2009;169:704-711.  
13 696  
14  
15 697 44. Galea S, Uddin M, Koenen K. The urban environment and mental disorders: Epigenetic  
16 links. *Epigenetics*. 2011;6:400-404.  
17 698  
18  
19 699 45. Esteller M. The necessity of a human epigenome project. *Carcinogenesis*.  
20 2006;27:1121-1125.  
21 700  
22  
23 701 46. Narayan P, Dragunow M. Pharmacology of epigenetics in brain disorders. *Br J*  
24 *Pharmacol*. 2010;159:285-303.  
25 702  
26  
27 703 47. Olsson CA, Foley DL, Parkinson-Bates M et al. Prospects for epigenetic research  
28 within cohort studies of psychological disorder: a pilot investigation of a peripheral cell  
29 704  
30 marker of epigenetic risk for depression. *Biol Psychol*. 2010;83:159-165.  
31 705  
32  
33 706 48. Koenen KC, Uddin M, Chang SC et al. SLC6A4 methylation modifies the effect of the  
34 number of traumatic events on risk for posttraumatic stress disorder. *Depress Anxiety*.  
35 707  
36 2011;28:639-647.  
37 708  
38  
39 709 49. Smith CL, Bolton A, Nguyen G. Genomic and Epigenomic Instability, Fragile Sites,  
40 Schizophrenia and Autism. *Current Genomics*. 2010;11:447-469.  
41 710  
42  
43 711 50. Dempster EL, Pidsley R, Schalkwyk LC et al. Disease-associated epigenetic changes in  
44 monozygotic twins discordant for schizophrenia and bipolar disorder. *Hum Mol Genet*.  
45 712  
46 2011;20:4786-4796.  
47 713  
48  
49 714 51. Park BY, Lee BC, Jung KH et al. Epigenetic changes of serotonin transporter in the  
50 patients with alcohol dependence: methylation of an serotonin transporter promoter  
51 715  
52 CpG island. *Psychiatry Investig*. 2011;8:130-133.  
53 716  
54  
55 717 52. Klengel T, Mehta D, Anacker C et al. Allele-specific FKBP5 DNA demethylation  
56 mediates gene-childhood trauma interactions. *Nat Neurosci*. 2013;16:33-41.  
57 718  
58  
59  
60

- 1  
2  
3 719 53. van Os J, Kenis G, Rutten BP. The environment and schizophrenia. *Nature*.  
4  
5 720 2010;468:203-212.  
6  
7 721 54. Merikangas KR, Risch N. Will the Genomics Revolution Revolutionize Psychiatry?  
8  
9 722 *American Journal of Psychiatry*. 2003;160:625-635.  
10  
11 723 55. Weissman MM, Brown AS, Talati A. Translational epidemiology in psychiatry: linking  
12  
13 724 population to clinical and basic sciences. *Arch Gen Psychiatry*. 2011;68:600-608.  
14  
15 725 56. Autonell J, Vila F, Pinto-Meza A et al. [One year prevalence of mental disorders  
16  
17 726 comorbidity and associated socio-demographic risk factors in the general population of  
18  
19 727 Spain. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:4-11.  
20  
21 728 57. Haro JM, Palacin C, Vilagut G et al. [Prevalence of mental disorders and associated  
22  
23 729 factors: results from the ESEMeD-Spain study]. *Med Clin (Barc)*. 2006;126:445-451.  
24  
25 730 58. Codony M, Alonso J, Almansa J et al. [Mental health care use in the Spanish general  
26  
27 731 populations. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:21-28.  
28  
29 732 59. Codony M, Alonso J, Almansa J et al. [Psychotropic medications use in Spain. Results  
30  
31 733 of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:29-36.  
32  
33 734 60. Gabilondo A, Rojas-Farreras S, Rodriguez A et al. Use of primary and specialized  
34  
35 735 mental health care for a major depressive episode in Spain by ESEMeD respondents.  
36  
37 736 *Psychiatr Serv*. 2011;62:152-161.  
38  
39 737 61. Perales J, Olaya B, Fernandez A et al. Association of childhood adversities with the first  
40  
41 738 onset of mental disorders in Spain: results from the ESEMeD project. *Soc Psychiatry*  
42  
43 739 *Psychiatr Epidemiol*. 2013;48:371-384.  
44  
45 740 62. Borrell C, Peiro R, Ramon N et al. [Socioeconomic inequalities and health plans in the  
46  
47 741 Autonomous Communities of Spain]. *Gac Sanit*. 2005;19:277-285.  
48  
49 742 63. Gonzalez B, Urbanos RM, Ortega P. [Public and private supply of health services by  
50  
51 743 autonomous communities in Spain]. *Gac Sanit*. 2004;18 Suppl 1:82-89.  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 744 64. Fierro I, Yanez JL, Alvarez FJ. [Premature death and potential years of life lost due to  
4 alcohol consumption in Spain and the different autonomous communities in 2004]. *Aten*  
5 745  
6 *Primaria*. 2010;42:95-101.  
7 746
- 8  
9 747 65. Ricci-Cabello I, Ruiz-Perez I, Plazaola-Castano J et al. [Mental disease, existence of  
10 748  
11 diagnostic, use of psychotropic medication: differences by autonomous communities  
12 749  
13 under the national health survey 2006]. *Rev Esp Salud Publica*. 2010;84:29-41.  
14 750
- 15 750 66. Kessler RC, Ustun TB. The World Mental Health (WMH) Survey Initiative Version of  
16 751  
17 the World Health Organization (WHO) Composite International Diagnostic Interview  
18 752  
19 (CIDI). *Int J Methods Psychiatr Res*. 2004;13:93-121.  
20 753
- 21 753 67. Alonso J, Ferrer M, Romera B et al. The European Study of the Epidemiology of  
22 754  
23 Mental Disorders (ESEMEd/MHEDEA 2000) project: rationale and methods. *Int J*  
24 755  
25 *Methods Psychiatr Res*. 2002;11:55-67.  
26 756
- 27 756 68. Harkness J, Pennell B, Villar A et al. Translation Procedures and Translation  
28 757  
29 Assessment in the World Mental Health Survey Initiative. In: Kessler RC, Bedirhan  
30 758  
31 Üstün T, eds. *The WHO World Mental Health Surveys: Global Perspectives on the*  
32 759  
33 *Epidemiology of Mental Disorders*. 2008; New York: Cambridge University Press.  
34 760  
35 World Health Organization.  
36 761
- 37 761 69. Haro JM, Arbabzadeh-Bouchez S, Bugha T et al. Concordance of the Composite  
38 762  
39 International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical  
40 763  
41 assessments in the WHO World Mental Health Surveys. *Int J Methods Psychiatr Res*.  
42 764  
43 2006;15:167-180.  
44 765
- 45 765 70. Kessler RC, Abelson J, Demler O et al. Clinical calibration of DSM-IV diagnoses in the  
46 766  
47 World Mental Health (WMH) version of the World Health Organization (WHO)  
48 767  
49 Composite International Diagnostic Interview (WMHCIDI). *Int J Methods Psychiatr*  
50 768  
51 *Res*. 2004;13:122-139.  
52 769
- 53 769 71. Navarro-Mateu F, Moran-Sanchez I, Alonso J et al. Cultural adaptation of the Latin  
54 770  
55 American version of the World Health Organization Composite International  
56 771  
57 Diagnostic Interview (WHO-CIDI) (v 3.0) for use in Spain. *Gac Sanit*. 2012;27:325-33.  
58  
59  
60

- 1  
2  
3 772 72. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for  
4  
5 773 grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12:189-  
6  
7 774 198.
- 8  
9 775 73. Lobo A, Saz P, Marcos G et al. [Revalidation and standardization of the cognition mini-  
10  
11 776 exam (first Spanish version of the Mini-Mental Status Examination) in the general  
12  
13 777 geriatric population]. *Med Clin (Barc).* 1999;112:767-774.
- 14  
15 778 74. Broadbent DE, Cooper PF, FitzGerald P et al. The Cognitive Failures Questionnaire  
16  
17 779 (CFQ) and its correlates. *Br J Clin Psychol.* 1982;21 (Pt 1):1-16.
- 18  
19 780 75. García Martínez J, Sánchez Cánovas J. Adaptación al cuestionario de fallos cognitivos  
20  
21 781 de Broadbent, Cooper, Fitzgerald y Parkes (CFQ, Cognitive Failure Questionnaire).  
22  
23 782 *Análisis y Modificación de Conducta.* 1994;20:727-752.
- 24  
25 783 76. Fusté-Escolano A, Ruíz J. Estructura factorial de la versión reducida del "Eysenck  
26  
27 784 Personality Profiler". *Psicothema.* 2000;12:406-411.
- 28  
29 785 77. Sandín B, Valiente RM, Olmedo M et al. Versión española del cuestionario EPQR-  
30  
31 786 Abreviado (EPQR-A) (II): Replicación factorial, fiabilidad y validez. *Revista de*  
32  
33 787 *Psicopatología y Psicología Clínica.* 2002;7:207-216.
- 34  
35 788 78. Sandín B, Valiente RM, Chorot P et al. Versión española del cuestionario EPQR-  
36  
37 789 Abreviado (EPQE-A) (I): Análisis exploratorio de la estructura factorial. *Revista de*  
38  
39 790 *Psicopatología y Psicología Clínica.* 2002;7:195-205.
- 40  
41 791 79. Wagnild G, Young H. Development and psychometric evaluation of the resilience scale.  
42  
43 792 *J Nurs Meas.* 1993;1:165-177.
- 44  
45 793 80. Heilemann M, Lee K, Kury F. Psychometric properties of the spanish version of the  
46  
47 794 Resilience Scale. *J Nurs Meas.* 2003;11:61-72.
- 48  
49 795 81. Stefanis NC, Hanssen M, Smirnis NK et al. Evidence that three dimensions of  
50  
51 796 psychosis have a distribution in the general population. *Psychol Med.* 2002;32:347-358.
- 52  
53 797 82. Brugha T, Bebbington P, Tennant C et al. The List of Threatening Experiences: a subset  
54  
55 798 of 12 life event categories with considerable long-term contextual threat. *Psychol Med.*  
56  
57 799 1985;15:189-194.

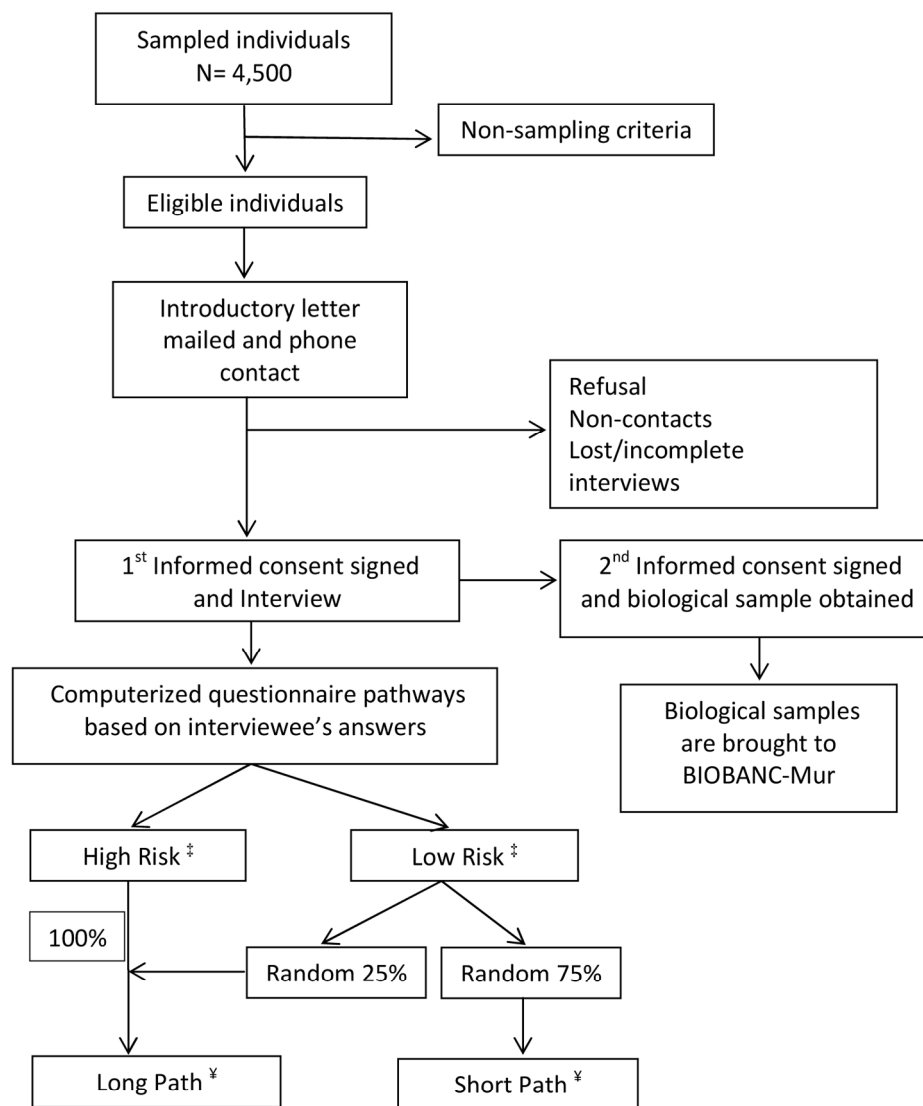


- 1  
2  
3 800 83. Brugha TS, Cragg D. The List of Threatening Experiences: the reliability and validity  
4  
5 801 of a brief life events questionnaire. *Acta Psychiatr Scand.* 1990;82:77-81.  
6  
7 802 84. van Os J, Park S, Jones P. Neuroticism, life events and mental health: evidence for  
8  
9 803 person--environment correlation. *The British Journal of Psychiatry.* 2001;178:s72-s77.  
10  
11 804 85. Badia X, Roset M, Montserrat S et al. [The Spanish version of EuroQol: a description  
12  
13 805 and its applications. European Quality of Life scale]. *Med Clin (Barc).* 1999;112:79-85.  
14  
15 806 86. Vilagut G, Valderas JM, Ferrer M et al. [Interpretation of SF-36 and SF-12  
16  
17 807 questionnaires in Spain: physical and mental components]. *Med Clin (Barc).*  
18  
19 808 2008;130:726-735.  
20  
21 809 87. Subdirección General de Salud Mental. *Guía práctica clínica. Actuación en salud*  
22  
23 810 *mental con mujeres maltratadas por su pareja.* 2010; Murcia: Servicio Murciano de  
24  
25 811 Salud.  
26  
27 812 88. Abitan H, Bohr H, Buchhave P. Correction to the Beer-Lambert-Bouguer law for  
28  
29 813 optical absorption. *Appl Opt.* 2008;47:5354-5357.  
30  
31 814 89. Commoner B, Lipkin D. The Application of the Beer-Lambert Law to Optically  
32  
33 815 Anisotropic Systems. *Science.* 1949;110:41-43.  
34  
35 816 90. Trumbo TA, Schultz E, Borland MG et al. Applied spectrophotometry: Analysis of a  
36  
37 817 biochemical mixture. *Biochem Mol Biol Educ.* 2013.  
38  
39 818 91. Mulot C, Stucker I, Clavel J et al. Collection of human genomic DNA from buccal cells  
40  
41 819 for genetics studies: comparison between cytobrush, mouthwash, and treated card. *J*  
42  
43 820 *Biomed Biotechnol.* 2005;2005:291-296.  
44  
45 821 92. Rogers NL, Cole SA, Lan HC et al. New saliva DNA collection method compared to  
46  
47 822 buccal cell collection techniques for epidemiological studies. *Am J Hum Biol.*  
48  
49 823 2007;19:319-326.  
50  
51 824 93. Viltrop T, Krjutskov K, Palta P et al. Comparison of DNA extraction methods for  
52  
53 825 multiplex polymerase chain reaction. *Anal Biochem.* 2010;398:260-262.  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 826 94. Valverde JC, Tormo MJ, Navarro C et al. Prevalence of diabetes in Murcia (Spain): a  
4  
5 827 Mediterranean area characterised by obesity. *Diabetes Res Clin Pract.* 2006;71:202-  
6  
7 828 209.  
8  
9 829 95. Huerta JM, Tormo MJ, Egea-Caparrós JM et al. Accuracy of self-reported diabetes,  
10  
11 830 hypertension and hyperlipidemia in the adult Spanish population. DINO study findings.  
12  
13 831 *Rev Esp Cardiol.* 2009;62:143-152.  
14  
15 832 96. Hyman SE. The genetics of mental illness: implications for practice. *Bull World Health*  
16  
17 833 *Organ.* 2000;78:455-463.  
18  
19 834 97. Merikangas KR, Low NC, Hardy J. Commentary: understanding sources of complexity  
20  
21 835 in chronic diseases--the importance of integration of genetics and epidemiology. *Int J*  
22  
23 836 *Epidemiol.* 2006;35:590-592.  
24  
25 837 98. Weich S, Araya R. International and regional variation in the prevalence of common  
26  
27 838 mental disorders: do we need more surveys? *Br J Psychiatry.* 2004;184:289-290.  
28  
29 839 99. Kessler RC. Psychiatric epidemiology: selected recent advances and future directions.  
30  
31 840 *Bull World Health Organ.* 2000;78:464-474.  
32  
33 841 100. Mercer KB, Orcutt HK, Quinn JF et al. Acute and Posttraumatic Stress Symptoms in a  
34  
35 842 Prospective Gene x Environment Study of a University Campus Shooting. *Arch Gen*  
36  
37 843 *Psychiatry.* 2011;69:89-97.  
38  
39 844 101. Tunbridge EM, Eastwood SL, Harrison PJ. Changed relative to what? Housekeeping  
40  
41 845 genes and normalization strategies in human brain gene expression studies. *Biol*  
42  
43 846 *Psychiatry.* 2011;69:173-179.  
44  
45 847 102. Abuhatzira L, Shemer R, Razin A. MeCP2 involvement in the regulation of neuronal  
46  
47 848 alpha-tubulin production. *Hum Mol Genet.* 2009;18:1415-1423.  
48  
49 849 103. Bayou N, Belhadj A, Daoud H et al. Exploring the 7p22.1 chromosome as a candidate  
50  
51 850 region for autism. *J Biomed Biotechnol.* 2010;2010:423894.  
52  
53 851 104. Fatjo-Vilas M, Papiol S, Estrada G et al. Dysbindin-1 gene contributes differentially to  
54  
55 852 early- and adult-onset forms of functional psychosis. *Am J Med Genet B Neuropsychiatr*  
56  
57 853 *Genet.* 2011;156B:322-333.  
58  
59  
60

- 1  
2  
3 854 105. Lasken RS, Egholm M. Whole genome amplification: abundant supplies of DNA from  
4  
5 855 precious samples or clinical specimens. *Trends Biotechnol.* 2003;21:531-535.  
6  
7 856  
8  
9 857  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For peer review only



† The response rate is defined as:  $(\text{completed interviews}) / (\text{total released respondent sample cases} - \text{respondent nonsample cases})$ .

‡ 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
<b>Title and abstract</b>	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
<b>Introduction</b>			
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
<b>Methods</b>			
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) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —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 <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	8-9
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —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) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	

Continued on next page

<b>Results</b>			<b>Pages</b>
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) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	-
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	-
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	-
		<i>Cross-sectional study</i> —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	-
<b>Discussion</b>			
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	-
<b>Other information</b>			
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](http://www.strobe-statement.org).



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

Journal:	<i>BMJ Open</i>
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, M <sup>a</sup> José; Consejería de Sanidad, Servicio de Epidemiología; CIBER de Epidemiología y Salud Pública (CIBERESP), Vilagut, Gemma; IMIM-Institut Hospital del Mar d'Investigacions Mèdiques, ; CIBER de Epidemiología y Salud Pública (CIBERESP), Alonso, Jordi; IMIM-Institut Hospital del Mar d'Investigacions Mèdiques, ; 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),
<b>Primary Subject Heading</b>:	Mental health
Secondary Subject Heading:	Epidemiology, Genetics and genomics, Public health
Keywords:	MENTAL HEALTH, EPIDEMIOLOGY, GENETICS

SCHOLARONE™  
Manuscripts

1  
2  
3 1 STUDY PROTOCOL4  
5 2 Title: **Epidemiology and Genetics of Common Mental Disorders in the general population:**  
6  
7 3 **the PEGASUS-Murcia project**8  
9  
10 4 **Authors:** F. Navarro-Mateu<sup>1,2,\*</sup>, M.J. Tormo<sup>2,3,4</sup>, G. Vilagut<sup>2,5</sup>, J. Alonso<sup>2,5,6</sup>, G. Ruíz-Merino<sup>7</sup>,  
11  
12 5 T. Escámez<sup>7,8</sup>, D. Salmerón<sup>2,3,4</sup>, J. Júdez<sup>7,8</sup>, S. Martínez<sup>9</sup>, C. Navarro<sup>2,3,4</sup>13  
14  
15  
16 6  
17 7 <sup>1</sup> Unidad de Docencia, Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de  
18 8 Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud. Murcia. Spain.19 9 <sup>2</sup> CIBER de Epidemiología y Salud Pública (CIBERESP). Spain.20 10 <sup>3</sup> Servicio de Epidemiología. Consejería de Sanidad y Política Social. Murcia. Spain.21 11 <sup>4</sup> Departamento de Ciencias Sociosanitarias. Universidad de Murcia. Spain.22 12 <sup>5</sup> IMIM-Institut Hospital del Mar d'Investigacions Mèdiques. Barcelona. Spain.23 13 <sup>6</sup> Departamento de Salud y Ciencias Experimentales. Universidad Pompeu Fabra. Barcelona. Spain.24 14 <sup>7</sup> Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia. Spain.25 15 <sup>8</sup> IMIB BIOBANC-MUR. Biobanco-HUVA-AECC-FFIS. Murcia. Spain.26 16 <sup>9</sup> Instituto de Neurociencias. UMH-CSIC. Alicante. Spain.27  
28  
29 \* **Corresponding author:** Fernando Navarro-Mateu, MD, PhD. Unidad de Docencia,  
30 20 Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de Salud Mental  
31 21 y Asistencia Psiquiátrica. Servicio Murciano de Salud. c/ Lorca, nº 58. 30120-El Palmar  
32 22 (Murcia). Spain. Email: [fernando.navarro@carm.es](mailto:fernando.navarro@carm.es)33  
34 23 **Emails of all coauthors**35  
36 24  
37 25 FNM: [fernando.navarro@carm.es](mailto:fernando.navarro@carm.es)  
38 26 MJT: [mjose.tormo@carm.es](mailto:mjose.tormo@carm.es)  
39 27 GV: [gvilagut@imim.es](mailto:gvilagut@imim.es)  
40 28 JA: [jalonso@imim.es](mailto:jalonso@imim.es)  
41 29 GRM: [guadalupe.ruiz@carm.es](mailto:guadalupe.ruiz@carm.es)  
42 30 TE: [teresa.escamez@carm.es](mailto:teresa.escamez@carm.es)  
43 31 DS: [diego.salmeron@carm.es](mailto:diego.salmeron@carm.es)  
44 32 JJ: [javier.judez@ffis.es](mailto:javier.judez@ffis.es)  
45 33 SM: [smartinez@umh.es](mailto:smartinez@umh.es)  
46 34 CN: [carmen.navarro@carm.es](mailto:carmen.navarro@carm.es)47  
48 37 **KEYWORDS:** Cross-sectional survey, mental disorders, prevalence, gene-environmental  
49 38 interactions, genome, epigenome, transcriptome.50  
51 40 **Word count** (excluding title page, abstract, references, figures and tables): **5 689**



1  
2  
3 41  
4  
5  
67 **ABSTRACT (298 words)**

8  
9 **Background:** Multidisciplinary collaboration between clinicians, epidemiologists,  
10 neurogeneticists and statisticians on research projects has been encouraged to improve our  
11 knowledge of the complex mechanisms underlying the etiology and burden of mental disorders.  
12 The PEGASUS-Murcia project was designed to assess the prevalence of common mental  
13 disorders, to identify risk and protective factors and it also included the collection of biological  
14 samples to study gene-environmental interactions in the context of the World Mental Health  
15 Survey Initiative.  
16  
17  
18  
19

20 **Methods and Analysis:** The PEGASUS-Murcia project is a new cross-sectional face-to-face  
21 interview survey based on a representative sample of non-institutionalized adults in the Region  
22 of Murcia (Mediterranean Southeast, Spain). Trained lay interviewers used the latest version for  
23 use in Spain of the computer-assisted personal interview (CAPI) of the Composite International  
24 Diagnostic Interview (CIDI 3.0), specifically adapted for the project. Two biological samples of  
25 buccal mucosal epithelium will be collected from each interviewed participant, one for DNA  
26 extraction for genomic and epigenomic analyses and the other to obtain mRNA for gene  
27 expression quantification. Several quality control procedures will be implemented to assure the  
28 highest reliability and validity of the data. This paper describes the rationale, sampling methods  
29 and questionnaire content as well as the laboratory methodology.  
30  
31  
32  
33  
34  
35

36 **Ethics and dissemination:** Informed consent will be obtained from all participants and a  
37 Regional Ethics Research Committee has approved the protocol. Results will be disseminated in  
38 peer reviewed publications and presented at national and international conferences.  
39  
40  
41

42 **Discussion:** Cross-sectional studies, which combine detailed personal information with  
43 biological data, offer new and exciting opportunities to study gene-environmental interactions in  
44 the etiology of common mental disorders in representative samples of the general population. A  
45 collaborative multidisciplinary research approach offers the potential to advance our knowledge  
46 of the underlying complex interactions and this opens the field for further innovative study  
47 designs in psychiatric epidemiology.  
48  
49  
50  
51  
52

53 **KEYWORDS:** Cross-sectional survey, mental disorders, prevalence, gene-environmental  
54 interactions, genome, epigenome, transcriptome.  
55  
56  
57  
58  
59  
60

72

**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.

73

74  
75 **BACKGROUND**

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

90 **Genes and environment factors in the etiology of mental disorders.**

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

1  
2  
3 102 SLEs. On the other hand, the serotonin transporter (*SERT* or *5HTT*) gene, a key regulator of  
4  
5 103 serotonergic neurotransmission and one of the most studied genetic polymorphisms in relation  
6  
7 104 to affective disorders,<sup>27</sup> has been associated with depression,<sup>28,29</sup> neuroticism<sup>30</sup> and posttraumatic  
8  
9 105 stress disorder (PTSD).<sup>31</sup> However, these findings have not always been replicated.<sup>32-34</sup>  
10  
11 106 These inconsistencies may be explained by, at least, three different factors. Firstly, in adults,  
12  
13 107 higher levels of neuroticism are associated with an increased risk of depression,<sup>35</sup> anxiety<sup>36</sup> and  
14  
15 108 PTSD after exposure to a traumatic event<sup>37</sup> and are a powerful predictor of comorbidity between  
16  
17 109 depression and anxiety.<sup>38</sup> Neuroticism includes those personality traits that represent how some  
18  
19 110 people perceive the world around them as threatening or stressful. In addition, some personality  
20  
21 111 traits also influence the individual tendency to be potentially exposed to stressful environments.  
22  
23 112 Predisposed individuals may tend to choose environments prone to having a high risk of  
24  
25 113 exposure to stressful events. Specifically, this scenario, known as GxE correlation, may mediate  
26  
27 114 the relationship between neuroticism and specific SLEs.<sup>39</sup> Secondly, the genetic factors  
28  
29 115 influencing the level of neuroticism, including the *5-HTTLPR* polymorphism, are shared by  
30  
31 116 persons having anxious-depressive spectrum disorders.<sup>38,40</sup> Lastly, GxE interactions have been  
32  
33 117 described involving *5-HTTLPR* and depression,<sup>41</sup> anxiety<sup>42</sup> and PTSD.<sup>43</sup> Despite all of the above  
34  
35 118 evidence, genetic association and GxE interaction studies do not usually analyze or control for  
36  
37 119 the level of neuroticism in the relationship between *5-HTTLPR*, SLEs and anxious-depressive  
38  
39 120 spectrum disorders.  
40  
41 121 However, the question arising in this context is how environmental and genetic factors interact  
42  
43 122 to produce a mental disorder.<sup>21,44</sup> In recent years, increasing interest in epigenetic factors  
44  
45 123 described in other human diseases has focused on its role in mental disorders.<sup>45</sup> The study of the  
46  
47 124 epigenome, changes in gene expression by modulating the accessibility of information that  
48  
49 125 occurs without modifying the DNA sequence, suggests that, although inheritable, these changes  
50  
51 126 are not necessarily stable over the life span of individuals and can be modified under some  
52  
53 127 environmental stimuli that modulate the activity of the enzymes involved, opening new  
54  
55 128 prospects for developing therapeutic approaches based on epigenetic mechanisms.<sup>46</sup> Epigenetic  
56  
57 129 mechanisms have been associated with different mental disorders including depression,<sup>47</sup>

1  
2  
3 130 PTSD,<sup>48</sup> schizophrenia,<sup>49,50</sup> autism,<sup>49</sup> bipolar disorder<sup>50</sup> and alcohol dependence.<sup>51</sup> In fact,  
4  
5 131 epigenetic regulation of the glucocorticoid receptor signaling in neurons has been recently  
6  
7 132 shown to be the mechanism underlying GxE interactions to explain risk and resilience of PTSD  
8  
9 133 after SLE in childhood.<sup>52</sup>

10  
11 134 In order to integrate all these findings and create new opportunities and challenges offered by  
12  
13 135 the GxE interaction scenarios in the field of mental disorders, a multidisciplinary collaboration  
14  
15 136 between clinicians, epidemiologists, geneticists and statisticians offers greater  
16  
17 137 opportunities.<sup>20,23,53</sup> One of the proposed mechanisms for this collaboration includes carrying  
18  
19 138 out community psychiatric surveys and this has been facilitated by the possibility of obtaining  
20  
21 139 DNA and/or mRNA from peripheral tissues. Specifically, saliva or buccal cells offers an easy,  
22  
23 140 save, inexpensive, and non-invasive method with accumulating scientific rationale to be added  
24  
25 141 in general population surveys.<sup>54-59</sup> Changes in gene expression can also be due to transcriptional  
26  
27 142 alterations. In order to deepen the understanding of molecular mechanisms implicated in mental  
28  
29 143 disorders, it is relevant to take into account transcriptional analyses with the RNA obtained at  
30  
31 144 the same time as the DNA samples. The opportunity to get both biological samples at the same  
32  
33 145 time from saliva offers the challenge of testing the suitability of this material for transcriptional  
34  
35 146 analyses in general population surveys. Population-based surveys offer several advantages over  
36  
37 147 other study designs to contribute to the clarification of the GxE interactions in mental  
38  
39 148 disorders.<sup>44,60,61</sup> Firstly, current knowledge of genes as risk factors is based almost exclusively  
40  
41 149 on clinical and non-representative population samples. Secondly, the distribution of the gene  
42  
43 150 polymorphisms of interest in the general population has not been well investigated. Thirdly, this  
44  
45 151 type of study can provide samples for future case-control studies and can be the bases for future  
46  
47 152 longitudinal ones. Finally, hypotheses generated from epidemiologic surveys may contribute to  
48  
49 153 test new basic studies and can be considered as a complementary strategy to translational  
50  
51 154 research.

#### 52 53 54 155 **Pegasus-Murcia Project**

55  
56 156 Spain actively participated in the ESEMeD Project with a representative sample of the adult  
57  
58 157 general Spanish population (n=5473) and the results have been published in national and

1  
2  
3 158 international journals.<sup>62-67</sup> However, the sample size within most of the Autonomous  
4  
5 159 Communities in Spain was too small to be able to achieve accurate and precise estimates at the  
6  
7 160 Regional level where Health Care policies are decided. Moreover, several differences between  
8  
9 161 the Autonomous Communities in Spain in important aspects related to mental health such as  
10  
11 162 socioeconomic<sup>68</sup> and territorial inequalities in health care supply and in long-term care, access  
12  
13 163 to and use of health care facilities,<sup>69</sup> premature deaths due to alcohol consumption<sup>70</sup> and the  
14  
15 164 prevalence of psychological distress<sup>71</sup> have recently been described.

16 165 Murcia is one of the 17 Autonomous Communities of Spain. It is located in the southeast of  
17  
18 166 the country on the Mediterranean coast, with a population of 1,424,063 inhabitants at the time  
19  
20 167 of the survey (INE 2008, National Statistical Institute of Spain), almost a third of them (30.7%)  
21  
22 168 living in the capital.

23  
24 169 The PEGASUS-Murcia (“Psychiatric Enquiry to General Population in Southeast Spain-  
25  
26 170 Murcia”) project has been designed in order to obtain regional data of the prevalence, burden  
27  
28 171 and care of a representative sample of the general adult population of Murcia to allow planning  
29  
30 172 of new regional mental health policies and to compare the results with the national sample of  
31  
32 173 Spain, Europe and all other countries participating in the WMH Survey Initiative. The project  
33  
34 174 also constitutes a unique opportunity to initiate a biological bank of a well-studied  
35  
36 175 representative sample of the general population.

### 37 176 **Objectives**

38  
39 177 The PEGASUS-Murcia project is a multi-purpose, observational, cross-sectional, comparative  
40  
41 178 study of the non-institutionalized general population of Murcia Region whose objective is to  
42  
43 179 improve knowledge about common psychiatric disorders in two main areas. The first one is the  
44  
45 180 epidemiology of mental disorders and protective and risk factors in the general population of  
46  
47 181 Murcia. The specific objectives are: i) to estimate the one-month, 12-month and lifetime  
48  
49 182 prevalence of the most common mental disorders, specifically, mood and anxiety disorders, in  
50  
51 183 the general population of Murcia; ii) to assess the independent association of mood and anxiety  
52  
53 184 disorders with sociodemographic factors (gender, age, education and urban/rural location) and  
54  
55 185 selected risk factors (family history, childhood experiences, religion, partnership status and  
56  
57  
58  
59  
60

1  
2  
3 186 sexual problems, among others); iii) to assess the quality of life of persons with the most  
4  
5 187 common psychiatric disorders and to analyze how other variables (physical medical conditions  
6  
7 188 and sociodemographic factors) may influence this outcome; iv) to assess the treatment for these  
8  
9 189 disorders and to evaluate the unmet need and the quality of care received; and v) to compare our  
10  
11 190 results with those obtained from Spain, Europe and other non-European countries, including the  
12  
13 191 United States. The second area is the genetic, epigenetic and transcriptomic influences  
14  
15 192 associated with mental disorders. Its specific aims include: i) the estimation of the distribution  
16  
17 193 of different candidate genes in the general population and their association with different  
18  
19 194 psychiatric disorders; ii) the identification of sensitive alleles underlying potential GxE  
20  
21 195 interactions and the study of epigenetic mechanisms involved, specifically, DNA methylation;  
22  
23 196 and iii) the analysis of gene expression alterations through transcriptomic assays.

## 25 197 **METHODS AND ANALYSIS**

### 28 198 **Study Design**

29  
30 199 The project is a cross-sectional face-to-face interview survey based on a representative sample  
31  
32 200 of the adult and non-institutionalized general population of the Murcia Region. Those who  
33  
34 201 complete the interview will be invited to provide two biological samples from their oral mucous  
35  
36 202 membranes. The target population is defined as persons aged 18 or older residing in Murcia, not  
37  
38 203 living in institutions and with an active health card (defined as persons included in PERSAN, a  
39  
40 204 regional registry that contains all residents with a Health Card which is periodically up-dated.  
41  
42 205 Exclusion criteria are: i) Confirmed irretrievable contact errors (e.g. telephone number and/or  
43  
44 206 address); ii) Institutionalized individuals (e.g. in prison, in a hospital or in another institution) or  
45  
46 207 those living outside the Autonomous Community during the survey field work; and iii)  
47  
48 208 individuals not able to understand the Spanish language or not able to conduct the questionnaire  
49  
50 209 due to his/her physical or mental condition.

### 52 210 **Sampling plan**

53  
54 211 The geographical area of the survey is the Murcia Region and a two-stage, stratified sampling  
55  
56 212 design has been used. The primary sampling unit is the Primary Health Centre and the second is

1  
2  
3 213 the individual. The sampling frame has been PERSAN, the regional health care population  
4  
5 214 database in Murcia. Primary Health Centres have been grouped into nine strata, the current  
6  
7 215 Health Care Areas in Murcia Region. The initial sample size was 4,500 adult individuals  
8  
9 216 divided into the nine Health Care Areas with proportionate allocation. A representative sample  
10  
11 217 of two centres has been chosen in each health area, without individual participant replacement.  
12  
13 218 Selection probability for each centre was known a priori and it was proportional to the size of  
14  
15 219 the centre (% of adult individuals registered in the centre) and the proportion of adult  
16  
17 220 individuals in the centre whose place of residence was rural, semi-urban or urban. Within each  
18  
19 221 of the two selected Health Centers, a stratified random sample procedure, performed for each  
20  
21 222 combination of gender, age group (18-24, 25-34, 35-49, 50-64 and 65+) and type of residence  
22  
23 223 (rural, semi-urban and urban), constitutes a stratum and individuals have been selected using  
24  
25 224 simple random sampling.

26  
27 225 For each Health Care Area, the sample size of each stratum has been selected such that  
28  
29 226 individuals in with the same demographic characteristics had equal probability of being selected  
30  
31 227 independently of the selected centre. If a high number of those fulfilling the exclusion criteria in  
32  
33 228 one area is reached, a fixed number of additional individuals will be released (subsequent  
34  
35 229 releases), according to the number of interviews completed in the area and following the same  
36  
37 230 selection procedure within each of the centres as the ones used to select the initial release (no  
38  
39 231 new centre will be selected for these releases). Any replacement of those persons who do not  
40  
41 232 want to collaborate or who do not meet the non-eligibility criteria is not allowed.

### 42 43 233 **Survey procedures and data control**

44  
45 234 Those selected will receive no financial incentive to participate and there will be no individual  
46  
47 235 replacement procedure. Trained lay interviewers carry out the survey using the computer-  
48  
49 236 assisted personal interview (CAPI) that was programmed centrally using the Blaise software  
50  
51 237 system. This is an interviewing application developed by Statistics Netherlands (Herleen, the  
52  
53 238 Netherlands) and designed to ease the handling of elaborate skip and complex randomization  
54  
55 239 patterns and to facilitate data entry, allow the elaboration of some questions and direct the  
56  
57 240 interviewer through the questioning sequence.  
58  
59  
60



1  
2  
3 241 Periodically, the completed interviews will be submitted to the central project Data Center  
4  
5 242 (Regional Mental Health Service, Murcia-Spain) for checking and storage following a  
6  
7 243 predetermined security procedure. All raw data will be transferred to the Hospital del Mar  
8  
9 244 Medical Research Institute (IMIM) and the Department of Health Care Policy at Harvard  
10  
11 245 University, coordinating centers of the ESEMeD and WMH Survey Initiative projects  
12  
13 246 respectively, via secure websites. The database has been declared to the Spanish Data Protection  
14  
15 247 Agency.

16  
17 248 A survey firm has been contracted to undertake the fieldwork and, in order to ensure the quality  
18  
19 249 of the survey, several strategies are being implemented: i) a one week training course for all  
20  
21 250 interviewers by WHO certified trainers on the original protocol and use of the CAPI version of  
22  
23 251 the CIDI; ii) development of a written manual to standardize the interviewing procedure and all  
24  
25 252 scientific and administrative elements that could affect comparability of data; iii) regular  
26  
27 253 meetings with the survey firm to ensure adherence to the protocol and to deal with any difficulty  
28  
29 254 that may have arisen; and iv) data quality analysis to detect any inconsistencies and/or  
30  
31 255 incomplete data.

32  
33 256 The survey firm has been provided with sufficient data to allow contact with each of the  
34  
35 257 individuals of the selected sample and only after 10 unsuccessful attempts the person will be  
36  
37 258 considered as not-contactable or after confirmation that the selected person does not live at that  
38  
39 259 address and new contact information is unavailable. Several methods will be used to improve  
40  
41 260 the participation of those selected: i) an informative flyer providing general information related  
42  
43 261 to the project and giving notice of future contact will be sent by conventional post together with  
44  
45 262 an invitation letter signed by a person from the Health Care Authority; ii) a phone call to invite  
46  
47 263 them to participate in the interview process and to offer them the possibility to do the interview  
48  
49 264 either at home or in their Primary Care Center; iii) Several informative sessions for the  
50  
51 265 healthcare personnel of the Primary Care Centers will be organized to facilitate their  
52  
53 266 collaboration should the participants ask them about the project; iv) During the period when the  
54  
55 267 interviews will take place, some official posters will be put in public centres to inform people  
56  
57  
58  
59  
60

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

### 270 **The survey questionnaire**

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

#### 275 *The Composite International Diagnostic Interview (CIDI)*

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

287 **Table 1: Description of the adapted version of the World Health Organization -Composite**  
288 **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 <sup>‡</sup>	Psychosis	42 to 84	All respondents
CFQ <sup>§</sup>	Risk Factors	25	All respondents
SLE <sup>€</sup>	Risk Factors	13 to 39	All respondents
Neuroticism and Extroversion subscales <sup>¥</sup>	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 Childhood	Socio-demographics	25-36	Long path
Attention Hyperactivity	Childhood	110	Long path
Oppositional Defiant	Childhood	90	Long path and Screening
Conduct Disorder	Childhood	46	Long path and Screening
Separation Anxiety Disorder	Childhood	54	Long path
Family Burden	Risk Factors	86	Screening questions (SCR)
Quality/ Lie subscale	Functioning and physical Disorder	40	Long path
Respondent Contacts	Methodological	26	Random assignment to the beginning of the questionnaire or at the end
Interviewer Observation	Methodological	19	All respondents
		14	All respondents

289 <sup>§</sup> EQ-5D: European Quality of Life Scale; <sup>¶</sup> SF-12 v2: Short Form 12 Health Questionnaire; <sup>†</sup> Lie subscale of the abbreviated version  
 290 of the Eysenck Personality Questionnaire (EPQR-A); <sup>‡</sup> CAPE: Community Assessment of Psychic Experiences; <sup>§</sup> CFQ: Cognitive  
 291 Failure Questionnaire; <sup>€</sup> SLE: Stressful Life Events; <sup>¥</sup> Neuroticism and Extroversion subscales of the abbreviated version of the  
 292 Eysenck Personality Questionnaire (EPQR-A)

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

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

1  
2  
3 301 standard WHO protocol with a rigorous process of adaptation.<sup>73,74</sup> Several clinical reappraisal  
4  
5 302 studies have been carried out and the concordance of the CIDI version 3.0 has been evaluated in  
6  
7 303 different subgroups of WMH surveys using the Structured Clinical Interview for DSM-IV  
8  
9 304 (SCID) as the clinical gold standard and a moderate to excellent concordance has been found for  
10  
11 305 most mental disorders.<sup>75,76</sup> CIDI is available in two formats: the paper form or PAPI (Paper and  
12  
13 306 Pencil Interviewing) and the computerized form or CAPI (Computer Assisted Personal  
14  
15 307 Interviewing), designed to ease the handling of elaborate skip and complex randomization  
16  
17 308 patterns and to facilitate data entry with a resulting reduction in interview time and errors in data  
18  
19 309 collection and recording. The original Spanish CAPI version used in Spain had not been  
20  
21 310 updated since it was used in the context of the ESEMeD project almost ten years ago. Since  
22  
23 311 then, all improvements in the questionnaire have only been added to the CIDI Latin American  
24  
25 312 (LA) v20.0 version. However, due to linguistic and cultural differences in Spanish-speaking  
26  
27 313 populations, this CAPI version had to be culturally adapted for use in Spain by our research  
28  
29 314 team and this process is fully described elsewhere.<sup>77</sup>  
30  
31 315 To further shorten the length of the questionnaire, some sections were not selected for the  
32  
33 316 purposes of this project. These include Intermittent Explosive Disorder, Personality I and II,  
34  
35 317 Neurasthenia and Pre-Menstrual and Gambling sections. Some others were substituted by other  
36  
37 318 questions or questionnaires, e.g. the Tobacco Use section was simplified using some questions  
38  
39 319 obtained from the Spanish National Health Survey and the Psychosis section with the CAPE  
40  
41 320 instrument (Community Assessment of Psychic Experiences), both described below.

#### 321 *Other study instruments*

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

1  
2  
3 329 CIDI, as the latest is only used as a screening instrument in the detection of psychosis. Those  
4  
5 330 who positively answer two items of the positive dimension with a score equal or superior to 3,  
6  
7 331 have been hospitalized for psychiatric reasons and/or have received psychotropic medication  
8  
9 332 during the last year will be evaluated by a clinic psychiatrist with the module C (Psychotic  
10  
11 333 Disorders) of the SCID (Structured Clinical Interview for DSM Disorders) ; vi) a brief list of 12  
12  
13 334 stressful life events in the last 12 months was included by the combination of a List of  
14  
15 335 Threatening Experiences (LTE)<sup>88,89</sup> and the emotional and life-changing impact of each event;<sup>90</sup>  
16  
17 336 vii) the European Quality of Life Scale (EuroQol 5D)<sup>91</sup> and the Short Form 12 Health  
18  
19 337 Questionnaire (SF-12 v2);<sup>92</sup> viii) an ad-hoc questionnaire of partner violence obtained from the  
20  
21 338 Spanish National Health Survey and from the regional mental health clinical guidelines;<sup>93</sup> and,  
22  
23 339 finally, ix) some questions related to tobacco use and physical exercises from the Spanish  
24  
25 340 National Health Survey.

#### 27 341 **Questionnaire pathways**

28  
29 342 In order to optimize the duration of the interview, the WMH questionnaire was divided into two  
30  
31 343 parts with questions in Part 1 administered to all respondents and those in Part 2 only to a  
32  
33 344 subsample of individuals who followed the long path of the interview. Part 2 of the interview  
34  
35 345 includes detailed information about a wide range of aspects related to the primary disorders and  
36  
37 346 also to mental disorders of secondary interest (Table 1). The inclusion criteria for the long path  
38  
39 347 are: a) all individuals that could be considered as “high risk individuals” because they positively  
40  
41 348 answer a number of specific questions related to mood and anxiety disorders and b) a random  
42  
43 349 subsample (25%) of the respondent without symptoms (“low risk individuals”). The remaining  
44  
45 350 75% of respondents without screening symptoms not randomly selected for the long path  
46  
47 351 followed the short path. The computer, without any intervention of the interviewer,  
48  
49 352 automatically makes all these pathways. In this shorter itinerary, a specific section, that included  
50  
51 353 those questions needed to calculate some demographic indicators, substituted the sections  
52  
53 354 omitted. Moreover, two sections were only used in a percentage of the long path itinerary,  
54  
55 355 Eating Disorders (50 %) and Obsessive-Compulsive Disorder (33%).

#### 57 356 **Quality control procedures**

1  
2  
3 357 Data quality will be controlled in a number of ways to ensure that the predetermined protocol  
4  
5 358 has been followed achieving the greatest reliability and validity and these quality control  
6  
7 359 procedures will be organized and supervised by members of the coordinating centers. The  
8  
9 360 principal investigator will reviewed all responses to open-ended questions to check if narratives  
10  
11 361 excludes a clinical diagnosis of mental disorders, i.e., whether symptoms were due to a physical  
12  
13 362 illness. All these procedures will be verified by the coordinating centers and the final document  
14  
15 363 included several aspects, for example, sample releases, the duration of the interviews and the  
16  
17 364 proportion of positive responses to selected screening questions. Local members of the research  
18  
19 365 team will be responsible for verifying the informed consent forms and the quality checking  
20  
21 366 following computerized protocols. These procedures are similar to those implemented in the  
22  
23 367 ESEMeD project and are fully described elsewhere.<sup>8</sup> Briefly, they consist of checks of  
24  
25 368 individual pieces of information from the interviewees, for example, completion status,  
26  
27 369 consistency across the questionnaire, questionnaire itinerary and length of the interview, and  
28  
29 370 from the interviewers, number of disorders screened positively, verification of a random  
30  
31 371 selection of almost 1% of interviews completed by a telephone contact to confirm the interview  
32  
33 372 and some aspects related to it such as place, approximate duration and identification of the  
34  
35 373 interviewer.

#### 374 **Laboratory Methods**

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

1  
2  
3 385 Samples for DNA extraction will be collected in sterile 1.5 ml tubes. Those to be used for RNA  
4  
5 386 extraction will be harvested in dark sterile tubes containing RNA protect cell Reagent  
6  
7 387 (QIAGEN, Hilden, Germany), which provides immediate stabilization of RNA. Cells will be  
8  
9 388 thus stabilized at room temperature and can then be stored or transported at ambient temperature  
10  
11 389 prior to RNA purification. Tubes will be labeled with tags (14C.B. 40X40 type) with a specific  
12  
13 390 code for each sample and will be packaged and sent to BIOBANC-MUR (the biobank for  
14  
15 391 biomedical research network of the Region of Murcia, RD09/0076/00065, as a partner of the  
16  
17 392 Spanish National Biobanks Network; IMIB: Instituto Murciano de Investigación Biosanitaria)  
18  
19 393 according to current Spanish legislation and following the regulations of the International Air  
20  
21 394 Transport Association (IATA) on biological sample shipping.  
22  
23 395 Those sample accepted by BIOBANC-MUR will be registered using a specific biobanking  
24  
25 396 software (bio-e-bank, VITROSOFT, SL), as part of a Laboratory Integrated Management  
26  
27 397 System (LIMS). The nucleic acid extraction will be performed automatically (QIAcube system;  
28  
29 398 QIAGEN, Hilden, Germany) to minimize variability due to manual handling using QIAamp  
30  
31 399 DNA Blood Mini Kit and RNeasyPlus Mini Kit (QIAGEN, Hilden, Germany) for DNA and  
32  
33 400 RNA extraction, respectively.  
34  
35 401 QIAamp DNA Blood Mini Kit provide fast and easy method for purification of total DNA for  
36  
37 402 reliable PCR and Southern blotting from whole human blood, buffy coat, cultured cells,  
38  
39 403 lymphocytes; plasma, serum, body fluids, and buccal swabs. The synthesis of complementary  
40  
41 404 DNA (cDNA) from mRNA for expression studies will be developed for all samples by reverse  
42  
43 405 transcription using the *High Capacity cDNA Reverse Transcription Kit* (Applied Biosystems).  
44  
45 406 All processes will be performed according to the manufacturer's instructions.  
46  
47 407 Nucleic acids quantity and quality will be determined by the ratio A260/280 calculated based on  
48  
49 408 260 and 280 nm absorbance measured using a spectrophotometer.<sup>95-97</sup> The ratio A260/230 is  
50  
51 409 commonly used as a secondary indicator of nucleic acid purity<sup>98-100</sup>. The integrity of DNA will  
52  
53 410 be visualized by electrophoresis on 1% agarose gel (migration for 1 hour at 100 V) using 100 ng  
54  
55 411 of total DNA and a 23 kb DNA ladder (Lambda DNA/HindIII Marker (Thermo Fisher  
56  
57 412 Scientific) as DNA marker. All mRNA samples will be transformed into cDNA.  
58  
59  
60

1  
2  
3 413 Specially trained technicians from the BIOBANC-MUR will be used to monitor the specimen  
4  
5 414 collection by donors and to perform sample manipulations in order to minimize variability of  
6  
7 415 results and to obtain the optimal quality of nucleic acids for this and future studies. The  
8  
9 416 processed biospecimens (150  $\mu$ l of DNA and 80  $\mu$ l of cDNA) will be stored in 750  $\mu$ l  
10  
11 417 microtubes in an ultra-freezer at -80 °C located in BIOBANC-MUR.

### 12 13 418 **Statistical methods**

14  
15 419 The expected response-rate (RR) has been set to a minimum of 65%, based on a previous  
16  
17 420 regional community survey which included the donation of blood samples<sup>101,102</sup>. The response  
18  
19 421 rate will be calculated based on the proportion of people interviewed and was defined as the  
20  
21 422 number of completed interviews divided by the total number of cases minus the number of non-  
22  
23 423 eligible cases.

### 24 25 424 **Weighting procedures**

26  
27 425 Given that the interview is divided into two parts and only a portion of the sample will be  
28  
29 426 selected for the second part, two types of weightings are considered to estimate population  
30  
31 427 parameters. The first is to weight for the probability of selection for each Health Care Area,  
32  
33 428 Health Centre and demographic stratum and the second is for the random skips included in the  
34  
35 429 questionnaire. The method designed is described in Box 1.

### 36 37 38 430 **BOX 1: Weighting procedures**

#### *First weighting procedure:*

39  
40 Step 1) For each Healthcare Area  $h$ , health centre  $c$  and demographic stratum (sex, age group and type of  
41 residence), all individuals have sampling weight  $w_s = 1/p_{hc}p^1_{hcsgr}$ , where  $p_{hc}$  is the probability that  
42 the centre  $c$  was selected,  $p^1_{hcsgr} = n_{hcsgr} / N_{hcsgr}$  and  $n_{hcsgr}$  is the sample size for the demographic  
43 stratum with  $N_{hcsgr}$  individuals registered in the sampling frame.

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

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

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

59 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  $\pi_i$  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:  $w_{p2norm} = w_{p2adj} n / \sum_{i=1}^n w_{p2adj_i}$ .

431

432

### **433 Analysis of the data and forthcoming research projects**

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

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

1  
2  
3 447 earthquake that are going to be interviewed in the PEGASUS-Murcia project and without a  
4  
5 448 diagnosis of any affective and/or anxiety disorder. Recently, our research team has published a  
6  
7 449 meta-analysis of the relationship between *5-HTTLPR* polymorphism and PTSD.<sup>34</sup>  
8

## 9 450 **ETHICS AND DISSEMINATION**

10  
11 451 Eligible individuals will be asked to sign two independent informed consents to participate, the  
12  
13 452 first one to be interviewed, including the possibility of future new contacts and the second to  
14  
15 453 provide the biological samples but only those who had already completed the questionnaire.  
16  
17 454 Name and contact information will be stored separately from any information provided as part  
18  
19 455 of the study questionnaire. The Clinical Research Ethics Committee of the University Hospital  
20  
21 456 Virgen de la Arrixaca of Murcia approved the protocol and the database of personal information  
22  
23 457 has been registered with the National Data Protection Agency. Data from PEGASUS-Murcia  
24  
25 458 project will be included in the WMH Cross National Sample for international comparisons. The  
26  
27 459 study findings will be submitted to peer-reviewed journals for publication, and presented at  
28  
29 460 national and international scientific meetings.  
30  
31

## 32 461 **DISCUSSION**

33  
34 462 The epidemiology of mental illnesses is a fascinating but highly complex area of research. This  
35  
36 463 complexity is primarily due to the wide range of factors, environmental and genetic, which  
37  
38 464 combines to produce a recognized psychiatric disorder. Previous epidemiological research has  
39  
40 465 resulted in the production of a great amount of data but it has been difficult to make cross-  
41  
42 466 national comparisons due to methodological variability. The WMH Survey Initiative aimed to  
43  
44 467 address this issue by using an international standardized protocol, allowing comparisons of the  
45  
46 468 most common mental disorders and their associated factors throughout the world. Using this  
47  
48 469 study design, it therefore offers the opportunity for new surveys to be performed in the context  
49  
50 470 of an international collaborative initiative and the possibility to adapt the questionnaire  
51  
52 471 according to the specific aims of the research being undertaken. The PEGASUS-Murcia project  
53  
54 472 can be considered as an example of how the latter has been successfully achieved. It is a cross-  
55  
56 473 sectional study designed to assess the prevalence of the most frequent mental disorders and their  
57  
58  
59  
60

1  
2  
3 474 correlates in a representative sample of the general population of Murcia. Its primary strengths  
4  
5 475 are: i) the fact that it was specifically adapted to assess factors not only associated with mental  
6  
7 476 disorders but also with positive mental health in a representative sample of the general  
8  
9 477 population; ii) its context focused on regional needs where healthcare decisions are taken  
10  
11 478 regarding resource allocation and mental health planning; iii) the collection of biological  
12  
13 479 samples not only for DNA analysis but also for mRNA; iv) all the information collected in our  
14  
15 480 study, including biological samples, can be correlated with past and future health events because  
16  
17 481 all Spanish population had free access to the Healthcare System at the time of its inception and  
18  
19 482 were thus registered and provided with a unique identification number and therefore; v) finally,  
20  
21 483 the inclusion of a multidisciplinary research team is in accordance with the international  
22  
23 484 consensus regarding the need for interdisciplinary collaboration between clinicians,  
24  
25 485 epidemiologists and neuroscience researchers to increase their combined efforts to study the  
26  
27 486 complex gene-gene and gene-environmental interactions underlying mental health  
28  
29 487 disorders.<sup>23,61,103,104</sup>  
30  
31 488 Concerns have been expressed about the cost-effectiveness of psychiatric epidemiological  
32  
33 489 surveys, such as World Mental Health 2000 (WMH-2000) projects,<sup>105</sup> an example being the  
34  
35 490 rationale for starting a new psychiatric epidemiological survey in the Autonomous Community  
36  
37 491 of Murcia if Spain had already participated in the ESEMeD project. However, there are several  
38  
39 492 reasons to justify this regional initiative. Firstly, public health and healthcare agencies usually  
40  
41 493 allocate mental health resources, including human, based on data from national epidemiologic  
42  
43 494 surveys,<sup>106</sup> such as that provided by the Spanish participation in the ESEMeD Project. As  
44  
45 495 previously mentioned, the involvement of the Region of Murcia in the Spanish ESEMeD survey  
46  
47 496 did not allow evaluation of specific regional data. Nowadays, the main responsibility for  
48  
49 497 planning and management of Healthcare resources in Spain lies with the Autonomous  
50  
51 498 Communities and differences exist between them in terms of accessibility, amount of health-  
52  
53 499 care resources and political decision-making.<sup>68-71</sup> Devolution of this responsibility to Murcia  
54  
55 500 occurred in December 2001.  
56  
57  
58  
59  
60

1  
2  
3 501 Secondly, the inclusion of biological data in a well-designed multidisciplinary epidemiological  
4  
5 502 study offers great advantages in terms of a more global understanding of mental disorders.  
6  
7 503 These are complex illnesses of the brain where social, familial, psychological and biological  
8  
9 504 elements interact throughout the entire life of a person to influence his/her risk of developing a  
10  
11 505 mental health disorder. To extend our understanding of the physiopathology and epidemiology  
12  
13 506 of the more common ones (mood and anxiety), it is necessary to identify genetic loci and  
14  
15 507 polymorphic alleles and their distribution in the healthy and affected population whose function  
16  
17 508 in determining risk for, and protection against, these conditions probably depends on gene-gene  
18  
19 509 and GxE interactions. The collection of genetic material from representative samples from the  
20  
21 510 general population, well described using international diagnostic instruments such as CIDI,  
22  
23 511 offers new and different possibilities to evaluate candidate genes in non-biased samples and to  
24  
25 512 describe their distribution in the general population that may contribute to clarification of the  
26  
27 513 complexity of mental disorders.

28  
29 514 Thirdly, our project involving a multidisciplinary research team gives new opportunities to  
30  
31 515 develop different study designs that can move from descriptive to analytical epidemiology. For  
32  
33 516 example, this representative sample constitutes a good source of controls for future case-control  
34  
35 517 studies, where cases will be provided from the public health care clinics, and can be the starting  
36  
37 518 point for future cohort studies. Our project was designed to allow for all these possibilities.

### 39 519 **Limitations of the study**

40  
41 520 Currently, the main limitations of the PEGASUS-Murcia project are related to: i) the cross-  
42  
43 521 sectional design which, while it allows association studies, limits the possible causal  
44  
45 522 interpretation of the findings. However, these findings may provide new hypotheses and enable  
46  
47 523 the design of new studies; ii) not all interviewees will provide biological samples and this may  
48  
49 524 affect the representativeness of some mental disorders in future analyses. To determine if this  
50  
51 525 will result in selection bias, we will analyze whether there are distinguishing characteristics  
52  
53 526 between donors and non-donors in the distribution of mental disorders and other characteristics  
54  
55 527 of the participants; iii) the population stratification in our study which will be used for future  
56  
57 528 genetic association analyses is performed by using the stated ancestral origin by participants<sup>107</sup>

1  
2  
3 529 instead of using genetic markers; and iv) biological samples will be obtained from oral mucosal  
4  
5 530 scrapings and not from brain neurons. However, this is a general situation given the ethical  
6  
7 531 issues and difficulties in obtaining neural tissues and, in any case, gene expression does not  
8  
9 532 appear to be specific to neural tissue, at least in some genes that have ubiquitous expression, for  
10  
11 533 example, 5-HTTLPR.<sup>108-111</sup>

### 534 **Conclusions and Future Directions**

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

### 539 **Competing interests**

540 The authors declare that they have no competing interests.

### 541 **Author's contributions**

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

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

557 The PEGASUS-Murcia project is supported by the Regional Health Authorities of Murcia  
558 (“Servicio Murciano de Salud and Consejería de Sanidad y Política Social”) (Decreto n<sup>o</sup>  
559 455/2009), the “Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de  
560 Murcia” (N<sup>o</sup> Expedientes: CM0829 I and FFIDS/EMER09/14) and the “Ayudas para proyectos

1  
2  
3 561 de Investigación en Salud –ISCIH- del Plan Nacional de Investigación Científica, Desarrollo e  
4 562 Innovación Tecnológica” (PI12/00809). None of the funders had any role in the design,  
5  
6 563 analysis, interpretation of results, or preparation of this paper. The PEGASUS-Murcia project is  
7  
8 564 carried out in conjunction with the WHOWMH Survey Initiative. We thank the WMH  
9  
10 565 Coordinating Center staff at Harvard and Michigan Universities and, especially Professor Ron  
11 566 Kessler, for their assistance with the instrumentation, fieldwork and data analysis. These  
12 567 activities were supported by the United States National Institute of Mental Health  
13 568 (R01MH070884), the John D. and Catherine T. MacArthur Foundation, the Pfizer Foundation,  
14  
15 569 the U.S. Public Health Service (R13-MH066849, R01-MH069864, and R01 DA016558), the  
16  
17 570 Fogarty International Center (FIRCA R03- TW006481), the Pan American Health Organization,  
18  
19 571 the Eli Lilly & Company Foundation, Ortho-McNeil Pharmaceutical, Inc., GlaxoSmithKline,  
20 572 Bristol-Myers Squibb and Shire. A complete list of WMH publications can be found at  
21 573 <http://www.hcp.med.harvard.edu/wmh/>.

22  
23 574  
24 575  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 576  
4 577 **Figure 1: Flow chart of the PEGASUS-Murcia project**  
5 578  
6 579  
7 580  
8 581  
9 582  
10 583  
11 584  
12 585  
13 586  
14 587  
15 588  
16 589  
17 590  
18 591  
19 592  
20 593  
21 594  
22 595  
23 596  
24 597  
25 598  
26 599  
27 600  
28 601  
29 602  
30 603  
31 604  
32 605  
33 606  
34 607  
35 608  
36 609  
37 610  
38 611  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

† 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

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

612

For peer review only



## 613 References

- 614 1. Kessler RC. The global burden of anxiety and mood disorders: putting the European Study of  
615 the Epidemiology of Mental Disorders (ESEMeD) findings into perspective. *J Clin*  
616 *Psychiatry*. 2007;68 Suppl 2:10-9.:10-19.
- 617 2. Alonso J, Lepine JP. Overview of key data from the European Study of the Epidemiology of  
618 Mental Disorders (ESEMeD). *J Clin Psychiatry*. 2007;68 Suppl 2:3-9.
- 619 3. Kessler RC, Aguilar-Gaxiola S, Alonso J et al. The WHO World Mental Health (WMH)  
620 Surveys. *Psychiatrie (Stuttg)*. 2009;6:5-9.
- 621 4. Kessler RC. The World Health Organization International Consortium in Psychiatric  
622 Epidemiology (ICPE): initial work and future directions -- the NAPE Lecture 1998.  
623 Nordic Association for Psychiatric Epidemiology. *Acta Psychiatr Scand*. 1999;99:2-9.
- 624 5. Kessler RC, Angermeyer M, ANTHONY JC et al. Lifetime prevalence and age-of-onset  
625 distributions of mental disorders in the World Health Organization's World Mental  
626 Health Survey Initiative. *World Psychiatry*. 2007;6:168-176.
- 627 6. de Graaf R, Kessler RC, Fayyad J et al. The prevalence and effects of adult attention-  
628 deficit/hyperactivity disorder (ADHD) on the performance of workers: results from the  
629 WHO World Mental Health Survey Initiative. *Occup Environ Med*. 2008;65:835-842.
- 630 7. Scott KM, von Korff M, Alonso J et al. Mental-physical co-morbidity and its relationship  
631 with disability: results from the World Mental Health Surveys. *Psychol Med*.  
632 2009;39:33-43.
- 633 8. Alonso J, Angermeyer MC, Bernert S et al. Sampling and methods of the European Study of  
634 the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*.  
635 2004;8-20.
- 636 9. Nutt DJ, Kessler RC, Alonso J et al. Consensus statement on the benefit to the community of  
637 ESEMeD (European Study of the Epidemiology of Mental Disorders) survey data on  
638 depression and anxiety. *J Clin Psychiatry*. 2007;68:42-48.

- 1  
2  
3 639 10. Alonso J, Angermeyer MC, Bernert S et al. Prevalence of mental disorders in Europe:  
4  
5 640 results from the European Study of the Epidemiology of Mental Disorders (ESEMeD)  
6  
7 641 project. *Acta Psychiatr Scand Suppl.* 2004;21-27.  
8  
9 642 11. Alonso J, Angermeyer MC, Bernert S et al. 12-Month comorbidity patterns and associated  
10  
11 643 factors in Europe: results from the European Study of the Epidemiology of Mental  
12  
13 644 Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl.* 2004;28-37.  
14  
15 645 12. Alonso J, Angermeyer MC, Bernert S et al. Disability and quality of life impact of mental  
16  
17 646 disorders in Europe: results from the European Study of the Epidemiology of Mental  
18  
19 647 Disorders (ESEMeD) project. *Acta Psychiatr Scand.* 2004;420:38-46.  
20  
21 648 13. Alonso J, Angermeyer MC, Bernert S et al. Use of mental health services in Europe: results  
22  
23 649 from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project.  
24  
25 650 *Acta Psychiatr Scand Suppl.* 2004;47-54.  
26  
27 651 14. Alonso J, Buron A, Bruffaerts R et al. Association of perceived stigma and mood and  
28  
29 652 anxiety disorders: results from the World Mental Health Surveys. *Acta Psychiatr Scand.*  
30  
31 653 2008;118:305-314.  
32  
33 654 15. Alonso J, Codony M, Kovess V et al. Population level of unmet need for mental healthcare  
34  
35 655 in Europe. *Br J Psychiatry.* 2007;190:299-306.  
36  
37 656 16. Codony M, Alonso J, Almansa J et al. Perceived need for mental health care and service use  
38  
39 657 among adults in Western Europe: results of the ESEMeD project. *Psychiatr Serv.*  
40  
41 658 2009;60:1051-1058.  
42  
43 659 17. Sevilla-Dedieu C, Kovess-Masfety V, Gilbert F et al. Mental health care and out-of-pocket  
44  
45 660 expenditures in Europe: results from the ESEMeD project. *J Ment Health Policy Econ.*  
46  
47 661 2011;14:95-105.  
48  
49 662 18. Hyman SE. The millennium of mind, brain, and behavior. *Arch Gen Psychiatry.*  
50  
51 663 2000;57:88-89.  
52  
53 664 19. Kubota T, Miyake K, Hirasawa T. Epigenetic understanding of gene-environment  
54  
55 665 interactions in psychiatric disorders: a new concept of clinical genetics. *Clin*  
56  
57 666 *Epigenetics.* 2012;4:1.  
58  
59  
60

- 1  
2  
3 667 20. Merikangas KR, Zhang H, Avenevoli S et al. Longitudinal trajectories of depression and  
4  
5 668 anxiety in a prospective community study: the Zurich Cohort Study. *Arch Gen*  
6  
7 669 *Psychiatry*. 2003;60:993-1000.  
8  
9 670 21. Toyokawa S, Uddin M, Koenen KC et al. How does the social environment 'get into the  
10  
11 671 mind'? Epigenetics at the intersection of social and psychiatric epidemiology. *Soc Sci*  
12  
13 672 *Med*. 2012;74:67-74.  
14  
15 673 22. Tsuang M, Bar J, Stone W et al. Gene-environment interactions in mental disorders. *World*  
16  
17 674 *Psychiatry*. 2004;3:73-83.  
18  
19 675 23. Caspi A, Moffitt TE. Gene-environment interactions in psychiatry: joining forces with  
20  
21 676 neuroscience. *Nat Rev Neurosci*. 2006;7:583-590.  
22  
23 677 24. Caspi A, Hariri AR, Holmes A et al. Genetic sensitivity to the environment: the case of the  
24  
25 678 serotonin transporter gene and its implications for studying complex diseases and traits.  
26  
27 679 *Am J Psychiatry*. 2010;167:509-527.  
28  
29 680 25. Koenen KC, Fu QJ, Ertel K et al. Common genetic liability to major depression and  
30  
31 681 posttraumatic stress disorder in men. *J Affect Disord*. 2008;105:109-115.  
32  
33 682 26. Norris FH, Friedman MJ, Watson PJ et al. 60,000 disaster victims speak: Part I. An  
34  
35 683 empirical review of the empirical literature, 1981-2001. *Psychiatry*. 2002;65:207-239.  
36  
37 684 27. Caspi A, Hariri AR, Holmes A et al. Genetic sensitivity to the environment: the case of the  
38  
39 685 serotonin transporter gene and its implications for studying complex diseases and traits.  
40  
41 686 *Am J Psychiatry*. 2010;167:509-527.  
42  
43 687 28. Clarke H, Flint J, Attwood AS et al. Association of the 5- HTTLPR genotype and unipolar  
44  
45 688 depression: a meta-analysis. *Psychol Med*. 2010;40:1767-1778.  
46  
47 689 29. Karg K, Burmeister M, Shedden K et al. The serotonin transporter promoter variant (5-  
48  
49 690 HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic  
50  
51 691 moderation. *Arch Gen Psychiatry*. 2011;68:444-454.  
52  
53 692 30. Takano A, Arakawa R, Hayashi M et al. Relationship between neuroticism personality trait  
54  
55 693 and serotonin transporter binding. *Biol Psychiatry*. 2007;62:588-592.  
56  
57  
58  
59  
60

- 1  
2  
3 694 31. Koenen K, Kilpatrick D, Acierno R et al. Serotonin transporter gene and social support  
4  
5 695 moderate post-traumatic stress disorder in hurricane exposed adults. *American Journal*  
6  
7 696 *of Medical Genetics Part B-Neuropsychiatric Genetics*. 2006;141B:742-743.  
8  
9 697 32. Risch N, Herrell R, Lehner T et al. Interaction between the serotonin transporter gene (5-  
10  
11 698 HTTLPR), stressful life events, and risk of depression: a meta-analysis. *JAMA*.  
12  
13 699 2009;301:2462-2471.  
14  
15 700 33. Munafo MR, Durrant C, Lewis G et al. Gene X environment interactions at the serotonin  
16  
17 701 transporter locus. *Biol Psychiatry*. 2009;65:211-219.  
18  
19 702 34. Navarro-Mateu F, Escámez T, Koenen KC et al. Meta-Analyses of the 5-HTTLPR  
20  
21 703 Polymorphisms and Post-Traumatic Stress Disorder. *PLOS One*. 2013;8:e66227.  
22  
23 704 35. Kendler KS, Neale MC, Kessler RC et al. A longitudinal twin study of personality and  
24  
25 705 major depression in women. *Arch Gen Psychiatry*. 1993;50:853-862.  
26  
27 706 36. Kotov R, Gamez W, Schmidt F et al. Linking "big" personality traits to anxiety, depressive,  
28  
29 707 and substance use disorders: a meta-analysis. *Psychol Bull*. 2010;136:768-821.  
30  
31 708 37. Wu D, Yin H, Xu S et al. Risk factors for posttraumatic stress reactions among Chinese  
32  
33 709 students following exposure to a snowstorm disaster. *BMC Public Health*. 2011;11:96.  
34  
35 710 38. Hettema JM, Neale MC, Myers JM et al. A population-based twin study of the relationship  
36  
37 711 between neuroticism and internalizing disorders. *Am J Psychiatry*. 2006;163:857-864.  
38  
39 712 39. Ploubidis GB, Frangou S. Neuroticism and psychological distress: to what extent is their  
40  
41 713 association due to person-environment correlation? *European Psychiatry*. 2011;26:1-5.  
42  
43 714 40. Skelton K, Ressler KJ, Norrholm SD et al. PTSD and gene variants: new pathways and new  
44  
45 715 thinking. *Neuropharmacology*. 2012;62:628-637.  
46  
47 716 41. Surtees PG, Wainwright NW, Willis-Owen SA et al. Social adversity, the serotonin  
48  
49 717 transporter (5-HTTLPR) polymorphism and major depressive disorder. *Biol Psychiatry*.  
50  
51 718 2006;59:224-229.  
52  
53 719 42. Stein MB, Schork NJ, Gelernter J. Gene-by-environment (serotonin transporter and  
54  
55 720 childhood maltreatment) interaction for anxiety sensitivity, an intermediate phenotype  
56  
57 721 for anxiety disorders. *Neuropsychopharmacology*. 2008;33:312-319.  
58  
59  
60

- 1  
2  
3 722 43. Koenen KC, Aiello AE, Bakshis E et al. Modification of the association between serotonin  
4  
5 723 transporter genotype and risk of posttraumatic stress disorder in adults by county-level  
6  
7 724 social environment. *Am J Epidemiol.* 2009;169:704-711.  
8  
9 725 44. Galea S, Uddin M, Koenen K. The urban environment and mental disorders: Epigenetic  
10  
11 726 links. *Epigenetics.* 2011;6:400-404.  
12  
13 727 45. Esteller M. The necessity of a human epigenome project. *Carcinogenesis.* 2006;27:1121-  
14  
15 728 1125.  
16  
17 729 46. Narayan P, Dragunow M. Pharmacology of epigenetics in brain disorders. *Br J Pharmacol.*  
18  
19 730 2010;159:285-303.  
20  
21 731 47. Olsson CA, Foley DL, Parkinson-Bates M et al. Prospects for epigenetic research within  
22  
23 732 cohort studies of psychological disorder: a pilot investigation of a peripheral cell marker  
24  
25 733 of epigenetic risk for depression. *Biol Psychol.* 2010;83:159-165.  
26  
27 734 48. Koenen KC, Uddin M, Chang SC et al. SLC6A4 methylation modifies the effect of the  
28  
29 735 number of traumatic events on risk for posttraumatic stress disorder. *Depress Anxiety.*  
30  
31 736 2011;28:639-647.  
32  
33 737 49. Smith CL, Bolton A, Nguyen G. Genomic and Epigenomic Instability, Fragile Sites,  
34  
35 738 Schizophrenia and Autism. *Current Genomics.* 2010;11:447-469.  
36  
37 739 50. Dempster EL, Pidsley R, Schalkwyk LC et al. Disease-associated epigenetic changes in  
38  
39 740 monozygotic twins discordant for schizophrenia and bipolar disorder. *Hum Mol Genet.*  
40  
41 741 2011;20:4786-4796.  
42  
43 742 51. Park BY, Lee BC, Jung KH et al. Epigenetic changes of serotonin transporter in the patients  
44  
45 743 with alcohol dependence: methylation of an serotonin transporter promoter CpG island.  
46  
47 744 *Psychiatry Investig.* 2011;8:130-133.  
48  
49 745 52. Klengel T, Mehta D, Anacker C et al. Allele-specific FKBP5 DNA demethylation mediates  
50  
51 746 gene-childhood trauma interactions. *Nat Neurosci.* 2013;16:33-41.  
52  
53 747 53. van Os J, Kenis G, Rutten BP. The environment and schizophrenia. *Nature.* 2010;468:203-  
54  
55 748 212.  
56  
57  
58  
59  
60

- 1  
2  
3 749 54. Lee YH, Wong DT. Saliva: an emerging biofluid for early detection of diseases. *Am J Dent*.  
4  
5 750 2009;22:241-248.  
6  
7 751 55. Yan W, Apweiler R, Balgley BM et al. Systematic comparison of the human saliva and  
8  
9 752 plasma proteomes. *Proteomics Clin Appl*. 2009;3:116-134.  
10  
11 753 56. Ballantyne J. Validity of messenger RNA expression analyses of human saliva. *Clin Cancer*  
12  
13 754 *Res*. 2007;13:1350.  
14  
15 755 57. Park NJ, Zhou X, Yu T et al. Characterization of salivary RNA by cDNA library analysis.  
16  
17 756 *Arch Oral Biol*. 2007;52:30-35.  
18  
19 757 58. Lee YH, Zhou H, Reiss JK et al. Direct saliva transcriptome analysis. *Clin Chem*.  
20  
21 758 2011;57:1295-1302.  
22  
23 759 59. Palanisamy V, Wong DT. Transcriptomic analyses of saliva. *Methods Mol Biol*.  
24  
25 760 2010;666:43-51. doi: 10.1007/978-1-60761-820-1\_4.:43-51.  
26  
27 761 60. Merikangas KR, Risch N. Will the Genomics Revolution Revolutionize Psychiatry?  
28  
29 762 *American Journal of Psychiatry*. 2003;160:625-635.  
30  
31 763 61. Weissman MM, Brown AS, Talati A. Translational epidemiology in psychiatry: linking  
32  
33 764 population to clinical and basic sciences. *Arch Gen Psychiatry*. 2011;68:600-608.  
34  
35 765 62. Autonell J, Vila F, Pinto-Meza A et al. [One year prevalence of mental disorders  
36  
37 766 comorbidity and associated socio-demographic risk factors in the general population of  
38  
39 767 Spain. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:4-11.  
40  
41 768 63. Haro JM, Palacin C, Vilagut G et al. [Prevalence of mental disorders and associated factors:  
42  
43 769 results from the ESEMeD-Spain study]. *Med Clin (Barc)*. 2006;126:445-451.  
44  
45 770 64. Codony M, Alonso J, Almansa J et al. [Mental health care use in the Spanish general  
46  
47 771 populations. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:21-28.  
48  
49 772 65. Codony M, Alonso J, Almansa J et al. [Psychotropic medications use in Spain. Results of  
50  
51 773 the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:29-36.  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 774 66. Gabilondo A, Rojas-Farreras S, Rodriguez A et al. Use of primary and specialized mental  
4  
5 775 health care for a major depressive episode in Spain by ESEMeD respondents. *Psychiatr*  
6  
7 776 *Serv.* 2011;62:152-161.
- 8  
9 777 67. Perales J, Olaya B, Fernandez A et al. Association of childhood adversities with the first  
10  
11 778 onset of mental disorders in Spain: results from the ESEMeD project. *Soc Psychiatry*  
12  
13 779 *Psychiatr Epidemiol.* 2013;48:371-384.
- 14  
15 780 68. Borrell C, Peiro R, Ramon N et al. [Socioeconomic inequalities and health plans in the  
16  
17 781 Autonomous Communities of Spain]. *Gac Sanit.* 2005;19:277-285.
- 18  
19 782 69. Gonzalez B, Urbanos RM, Ortega P. [Public and private supply of health services by  
20  
21 783 autonomous communities in Spain]. *Gac Sanit.* 2004;18 Suppl 1:82-89.
- 22  
23 784 70. Fierro I, Yanez JL, Alvarez FJ. [Premature death and potential years of life lost due to  
24  
25 785 alcohol consumption in Spain and the different autonomous communities in 2004]. *Aten*  
26  
27 786 *Primaria.* 2010;42:95-101.
- 28  
29 787 71. Ricci-Cabello I, Ruiz-Perez I, Plazaola-Castano J et al. [Mental disease, existence of  
30  
31 788 diagnostic, use of psychotropic medication: differences by autonomous communities  
32  
33 789 under the national health survey 2006]. *Rev Esp Salud Publica.* 2010;84:29-41.
- 34  
35 790 72. Kessler RC, Ustun TB. The World Mental Health (WMH) Survey Initiative Version of the  
36  
37 791 World Health Organization (WHO) Composite International Diagnostic Interview  
38  
39 792 (CIDI). *Int J Methods Psychiatr Res.* 2004;13:93-121.
- 40  
41 793 73. Alonso J, Ferrer M, Romera B et al. The European Study of the Epidemiology of Mental  
42  
43 794 Disorders (ESEMeD/MHEDEA 2000) project: rationale and methods. *Int J Methods*  
44  
45 795 *Psychiatr Res.* 2002;11:55-67.
- 46  
47 796 74. Harkness J, Pennell B, Villar A et al. Translation Procedures and Translation Assessment in  
48  
49 797 the World Mental Health Survey Initiative. In: Kessler RC, Bedirhan Üstün T, eds. *The*  
50  
51 798 *WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of*  
52  
53 799 *Mental Disorders.*2008; New York: Cambridge University Press. World Health  
54  
55 800 Organization.

- 1  
2  
3 801 75. Haro JM, Arbabzadeh-Bouchez S, Bugha T et al. Concordance of the Composite  
4  
5 802 International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical  
6  
7 803 assessments in the WHO World Mental Health Surveys. *Int J Methods Psychiatr Res.*  
8  
9 804 2006;15:167-180.
- 10  
11 805 76. Kessler RC, Abelson J, Demler O et al. Clinical calibration of DSM-IV diagnoses in the  
12  
13 806 World Mental Health (WMH) version of the World Health Organization (WHO)  
14  
15 807 Composite International Diagnostic Interview (WMHCIDI). *Int J Methods Psychiatr*  
16  
17 808 *Res.* 2004;13:122-139.
- 18  
19 809 77. Navarro-Mateu F, Moran-Sanchez I, Alonso J et al. Cultural adaptation of the Latin  
20  
21 810 American version of the World Health Organization Composite International  
22  
23 811 Diagnostic Interview (WHO-CIDI) (v 3.0) for use in Spain. *Gac Sanit.* 2012;27:325-  
24  
25 812 331.
- 26  
27 813 78. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading  
28  
29 814 the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12:189-198.
- 30  
31 815 79. Lobo A, Saz P, Marcos G et al. [Revalidation and standardization of the cognition mini-  
32  
33 816 exam (first Spanish version of the Mini-Mental Status Examination) in the general  
34  
35 817 geriatric population]. *Med Clin (Barc).* 1999;112:767-774.
- 36  
37 818 80. Broadbent DE, Cooper PF, FitzGerald P et al. The Cognitive Failures Questionnaire (CFQ)  
38  
39 819 and its correlates. *Br J Clin Psychol.* 1982;21 (Pt 1):1-16.
- 40  
41 820 81. García Martínez J, Sánchez Cánovas J. Adaptación al cuestionario de fallos cognitivos de  
42  
43 821 Broadbent, Cooper, Fitzgerald y Parkes (CFQ, Cognitive Failure Questionnaire).  
44  
45 822 *Análisis y Modificación de Conducta.* 1994;20:727-752.
- 46  
47 823 82. Fusté-Escolano A, Ruíz J. Estructura factorial de la versión reducida del "Eysenck  
48  
49 824 Personality Profiler". *Psicothema.* 2000;12:406-411.
- 50  
51 825 83. Sandín B, Valiente RM, Olmedo M et al. Versión española del cuestionario EPQR-  
52  
53 826 Abreviado (EPQR-A) (II): Replicación factorial, fiabilidad y validez. *Revista de*  
54  
55 827 *Psicopatología y Psicología Clínica.* 2002;7:207-216.



- 1  
2  
3 828 84. Sandín B, Valiente RM, Chorot P et al. Versión española del cuestionario EPQR-Abreviado  
4  
5 829 (EPQE-A) (I): Análisis exploratorio de la estructura factorial. *Revista de Psicopatología*  
6  
7 830 *y Psicología Clínica*. 2002;7:195-205.  
8  
9 831 85. Wagnild G, Young H. Development and psychometric evaluation of the resilience scale. *J*  
10  
11 832 *Nurs Meas*. 1993;1:165-177.  
12  
13 833 86. Heilemann M, Lee K, Kury F. Psychometric properties of the spanish version of the  
14  
15 834 Resilience Scale. *J Nurs Meas*. 2003;11:61-72.  
16  
17 835 87. Stefanis NC, Hanssen M, Smirnis NK et al. Evidence that three dimensions of psychosis  
18  
19 836 have a distribution in the general population. *Psychol Med*. 2002;32:347-358.  
20  
21 837 88. Brugha T, Bebbington P, Tennant C et al. The List of Threatening Experiences: a subset of  
22  
23 838 12 life event categories with considerable long-term contextual threat. *Psychol Med*.  
24  
25 839 1985;15:189-194.  
26  
27 840 89. Brugha TS, Cragg D. The List of Threatening Experiences: the reliability and validity of a  
28  
29 841 brief life events questionnaire. *Acta Psychiatr Scand*. 1990;82:77-81.  
30  
31 842 90. van Os J, Park S, Jones P. Neuroticism, life events and mental health: evidence for person--  
32  
33 843 environment correlation. *The British Journal of Psychiatry*. 2001;178:s72-s77.  
34  
35 844 91. Badia X, Roset M, Montserrat S et al. [The Spanish version of EuroQol: a description and  
36  
37 845 its applications. European Quality of Life scale]. *Med Clin (Barc)*. 1999;112:79-85.  
38  
39 846 92. Vilagut G, Valderas JM, Ferrer M et al. [Interpretation of SF-36 and SF-12 questionnaires in  
40  
41 847 Spain: physical and mental components]. *Med Clin (Barc)*. 2008;130:726-735.  
42  
43 848 93. Subdirección General de Salud Mental. *Guía práctica clínica. Actuación en salud mental*  
44  
45 849 *con mujeres maltratadas por su pareja*. 2010; Murcia: Servicio Murciano de Salud.  
46  
47 850 94. Beskow LM, Burke W, Merz JF et al. Informed consent for population-based research  
48  
49 851 involving genetics. *JAMA*. 2001;286:2315-2321.  
50  
51 852 95. Abitan H, Bohr H, Buchhave P. Correction to the Beer-Lambert-Bouguer law for optical  
52  
53 853 absorption. *Appl Opt*. 2008;47:5354-5357.  
54  
55 854 96. Commoner B, Lipkin D. The Application of the Beer-Lambert Law to Optically Anisotropic  
56  
57 855 Systems. *Science*. 1949;110:41-43.

- 1  
2  
3 856 97. Trumbo TA, Schultz E, Borland MG et al. Applied spectrophotometry: Analysis of a  
4  
5 857 biochemical mixture. *Biochem Mol Biol Educ*. 2013.  
6  
7 858 98. Mulot C, Stucker I, Clavel J et al. Collection of human genomic DNA from buccal cells for  
8  
9 859 genetics studies: comparison between cytobrush, mouthwash, and treated card. *J*  
10  
11 860 *Biomed Biotechnol*. 2005;2005:291-296.  
12  
13 861 99. Rogers NL, Cole SA, Lan HC et al. New saliva DNA collection method compared to buccal  
14  
15 862 cell collection techniques for epidemiological studies. *Am J Hum Biol*. 2007;19:319-  
16  
17 863 326.  
18  
19 864 100. Viltrop T, Krjutskov K, Palta P et al. Comparison of DNA extraction methods for  
20  
21 865 multiplex polymerase chain reaction. *Anal Biochem*. 2010;398:260-262.  
22  
23 866 101. Valverde JC, Tormo MJ, Navarro C et al. Prevalence of diabetes in Murcia (Spain): a  
24  
25 867 Mediterranean area characterised by obesity. *Diabetes Res Clin Pract*. 2006;71:202-  
26  
27 868 209.  
28  
29 869 102. Huerta JM, Tormo MJ, Egea-Caparros JM et al. Accuracy of self-reported diabetes,  
30  
31 870 hypertension and hyperlipidemia in the adult Spanish population. DINO study findings.  
32  
33 871 *Rev Esp Cardiol*. 2009;62:143-152.  
34  
35 872 103. Hyman SE. The genetics of mental illness: implications for practice. *Bull World Health*  
36  
37 873 *Organ*. 2000;78:455-463.  
38  
39 874 104. Merikangas KR, Low NC, Hardy J. Commentary: understanding sources of complexity in  
40  
41 875 chronic diseases--the importance of integration of genetics and epidemiology. *Int J*  
42  
43 876 *Epidemiol*. 2006;35:590-592.  
44  
45 877 105. Weich S, Araya R. International and regional variation in the prevalence of common  
46  
47 878 mental disorders: do we need more surveys? *Br J Psychiatry*. 2004;184:289-290.  
48  
49 879 106. Kessler RC. Psychiatric epidemiology: selected recent advances and future directions. *Bull*  
50  
51 880 *World Health Organ*. 2000;78:464-474.  
52  
53 881 107. Mercer KB, Orcutt HK, Quinn JF et al. Acute and Posttraumatic Stress Symptoms in a  
54  
55 882 Prospective Gene x Environment Study of a University Campus Shooting. *Arch Gen*  
56  
57 883 *Psychiatry*. 2011;69:89-97.  
58  
59  
60

- 1  
2  
3 884 108. Tunbridge EM, Eastwood SL, Harrison PJ. Changed relative to what? Housekeeping genes  
4  
5 885 and normalization strategies in human brain gene expression studies. *Biol Psychiatry*.  
6  
7 886 2011;69:173-179.  
8  
9 887 109. Abuhatzira L, Shemer R, Razin A. MeCP2 involvement in the regulation of neuronal  
10  
11 888 alpha-tubulin production. *Hum Mol Genet*. 2009;18:1415-1423.  
12  
13 889 110. Bayou N, Belhadj A, Daoud H et al. Exploring the 7p22.1 chromosome as a candidate  
14  
15 890 region for autism. *J Biomed Biotechnol*. 2010;2010:423894.  
16  
17 891 111. Fatjo-Vilas M, Papiol S, Estrada G et al. Dysbindin-1 gene contributes differentially to  
18  
19 892 early- and adult-onset forms of functional psychosis. *Am J Med Genet B Neuropsychiatr*  
20  
21 893 *Genet*. 2011;156B:322-333.  
22  
23  
24  
25  
26 894  
27  
28  
29 895  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 1 STUDY PROTOCOL4  
5 2 Title: **Epidemiology and Genetics of Common Mental Disorders in the general population:**  
6  
7 3 **the PEGASUS-Murcia project**8  
9  
10 4 **Authors:** F. Navarro-Mateu<sup>1,2,\*</sup>, M.J. Tormo<sup>2,3,4</sup>, G. Vilagut<sup>2,5</sup>, J. Alonso<sup>2,5,6</sup>, G. Ruíz-Merino<sup>7</sup>,  
11  
12 5 T. Escámez<sup>7,8</sup>, D. Salmerón<sup>2,3,4</sup>, J. Júdez<sup>7,8</sup>, S. Martínez<sup>9</sup>, C. Navarro<sup>2,3,4</sup>13  
14  
15  
16 6  
17 7 <sup>1</sup> Unidad de Docencia, Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de  
18 8 Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud. Murcia. Spain.19 9 <sup>2</sup> CIBER de Epidemiología y Salud Pública (CIBERESP). Spain.20 10 <sup>3</sup> Servicio de Epidemiología. Consejería de Sanidad y Política Social. Murcia. Spain.21 11 <sup>4</sup> Departamento de Ciencias Sociosanitarias. Universidad de Murcia. Spain.22 12 <sup>5</sup> IMIM-Institut Hospital del Mar d'Investigacions Mèdiques. Barcelona. Spain.23 13 <sup>6</sup> Departamento de Salud y Ciencias Experimentales. Universidad Pompeu Fabra. Barcelona. Spain.24 14 <sup>7</sup> Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de Murcia. Spain.25 15 <sup>8</sup> IMIB BIOBANC-MUR. Biobanco-HUVA-AECC-FFIS. Murcia. Spain.26 16 <sup>9</sup> Instituto de Neurociencias. UMH-CSIC. Alicante. Spain.27  
28  
29 \* **Corresponding author:** Fernando Navarro-Mateu, MD, PhD. Unidad de Docencia,  
30 Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de Salud Mental  
31 y Asistencia Psiquiátrica. Servicio Murciano de Salud. c/ Lorca, nº 58. 30120-El Palmar  
32 (Murcia). Spain. Email: [fernando.navarro@carm.es](mailto:fernando.navarro@carm.es)33  
34 23 **Emails of all coauthors**35  
36 24  
37 25 FNM: [fernando.navarro@carm.es](mailto:fernando.navarro@carm.es)  
38 26 MJT: [mjose.tormo@carm.es](mailto:mjose.tormo@carm.es)  
39 27 GV: [gvilagut@imim.es](mailto:gvilagut@imim.es)  
40 28 JA: [jalonso@imim.es](mailto:jalonso@imim.es)  
41 29 GRM: [guadalupe.ruiz@carm.es](mailto:guadalupe.ruiz@carm.es)  
42 30 TE: [teresa.escamez@carm.es](mailto:teresa.escamez@carm.es)  
43 31 DS: [diego.salmeron@carm.es](mailto:diego.salmeron@carm.es)  
44 32 JJ: [javier.judez@ffis.es](mailto:javier.judez@ffis.es)  
45 33 SM: [smartinez@umh.es](mailto:smartinez@umh.es)  
46 34 CN: [carmen.navarro@carm.es](mailto:carmen.navarro@carm.es)  
47 35  
48 3649 37 **KEYWORDS:** Cross-sectional survey, mental disorders, prevalence, gene-environmental  
50 38 interactions, genome, epigenome, transcriptome.51 39  
52 40 **Word count** (excluding title page, abstract, references, figures and tables): **5 689**

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## 41 ABSTRACT (298 words)

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

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

59 **Ethics and dissemination:** Informed consent ~~was~~ will be obtained from all participants and [a](#)  
60 [Regional Ethics Research Committee](#) ~~the protocol was~~ has been approved [the protocol by a](#)  
61 [Regional Ethics Research Committee](#). Results will be disseminated in peer reviewed  
62 publications and presented at national and international conferences.

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

70 **KEYWORDS:** Cross-sectional survey, mental disorders, prevalence, gene-environmental  
71 interactions, genome, epigenome, transcriptome.  
72

## 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 is 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.

## 74 **BACKGROUND**

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

### 89 **Genes and environment factors in the etiology of mental disorders.**

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

1  
2  
3 101 SLEs. On the other hand, the serotonin transporter (*SERT* or *5HTT*) gene, a key regulator of  
4  
5 102 serotonergic neurotransmission and one of the most studied genetic polymorphisms in relation  
6  
7 103 to affective disorders,<sup>27</sup> has been associated with depression,<sup>28,29</sup> neuroticism<sup>30</sup> and posttraumatic  
8  
9 104 stress disorder (PTSD).<sup>31</sup> However, these findings have not always been replicated.<sup>32-34</sup>  
10  
11 105 These inconsistencies may be explained by, at least, three different factors. Firstly, in adults,  
12  
13 106 higher levels of neuroticism are associated with an increased risk of depression,<sup>35</sup> anxiety<sup>36</sup> and  
14  
15 107 PTSD after exposure to a traumatic event<sup>37</sup> and are a powerful predictor of comorbidity between  
16  
17 108 depression and anxiety.<sup>38</sup> Neuroticism includes those personality traits that represent how some  
18  
19 109 people perceive the world around them as threatening or stressful. In addition, some personality  
20  
21 110 traits also influence the individual tendency to be potentially exposed to stressful environments.  
22  
23 111 Predisposed individuals may tend to choose environments prone to having a high risk of  
24  
25 112 exposure to stressful events. Specifically, this scenario, known as GxE correlation, may mediate  
26  
27 113 the relationship between neuroticism and specific SLEs.<sup>39</sup> Secondly, the genetic factors  
28  
29 114 influencing the level of neuroticism, including the *5-HTTLPR* polymorphism, are shared by  
30  
31 115 persons having anxious-depressive spectrum disorders.<sup>38,40</sup> Lastly, GxE interactions have been  
32  
33 116 described involving *5-HTTLPR* and depression,<sup>41</sup> anxiety<sup>42</sup> and PTSD.<sup>43</sup> Despite all of the above  
34  
35 117 evidence, genetic association and GxE interaction studies do not usually analyze or control for  
36  
37 118 the level of neuroticism in the relationship between *5-HTTLPR*, SLEs and anxious-depressive  
38  
39 119 spectrum disorders.  
40  
41 120 However, the question arising in this context is how environmental and genetic factors interact  
42  
43 121 to produce a mental disorder.<sup>21,44</sup> In recent years, increasing interest in epigenetic factors  
44  
45 122 described in other human diseases has focused on its role in mental disorders.<sup>45</sup> The study of the  
46  
47 123 epigenome, changes in gene expression by modulating the accessibility of information that  
48  
49 124 occurs without modifying the DNA sequence, suggests that, although inheritable, these changes  
50  
51 125 are not necessarily stable over the life span of individuals and can be modified under some  
52  
53 126 environmental stimuli that modulate the activity of the enzymes involved, opening new  
54  
55 127 prospects for developing therapeutic approaches based on epigenetic mechanisms.<sup>46</sup> Epigenetic  
56  
57 128 mechanisms have been associated with different mental disorders including depression,<sup>47</sup>



1  
2  
3 129 PTSD,<sup>48</sup> schizophrenia,<sup>49,50</sup> autism,<sup>49</sup> bipolar disorder<sup>50</sup> and alcohol dependence.<sup>51</sup> In fact,  
4  
5 130 epigenetic regulation of the glucocorticoid receptor signaling in neurons has been recently  
6  
7 131 shown to be the mechanism underlying GxE interactions to explain risk and resilience of PTSD  
8  
9 132 after SLE in childhood.<sup>52</sup>

10  
11 133 In order to integrate all these findings and create new opportunities and challenges offered by  
12  
13 134 the GxE interaction scenarios in the field of mental disorders, a multidisciplinary collaboration  
14  
15 135 between clinicians, epidemiologists, geneticists and statisticians offers greater  
16  
17 136 opportunities.<sup>20,23,53</sup> One of the proposed mechanisms for this collaboration includes carrying  
18  
19 137 out community psychiatric surveys and this has been facilitated by the possibility of obtaining  
20  
21 138 DNA and/or mRNA from peripheral tissues (~~blood, saliva or buccal cells~~). Specifically, saliva  
22  
23 139 or buccal cells offers an easy, save, inexpensive, and non-invasive method with accumulating  
24  
25 140 scientific rationale to be added in general population surveys.<sup>54-59</sup> Changes in gene expression  
26  
27 141 can also be due to transcriptional alterations. In order to deepen the understanding of molecular  
28  
29 142 mechanisms implicated in mental disorders, it is relevant to take into account transcriptional  
30  
31 143 analyses with the RNA obtained at the same time as the DNA samples. The opportunity to get  
32  
33 144 both biological samples at the same time from saliva offers the challenge of testing the  
34  
35 145 suitability of this material for transcriptional analyses in general population surveys. Population-  
36  
37 146 based surveys offer several advantages over other study designs to contribute to the clarification  
38  
39 147 of the GxE interactions in mental disorders.<sup>44,60,61</sup> Firstly, current knowledge of genes as risk  
40  
41 148 factors is based almost exclusively on clinical and non-representative population samples.  
42  
43 149 Secondly, the distribution of the gene polymorphisms of interest in the general population has  
44  
45 150 not been well investigated. Thirdly, this type of study can provide samples for future case-  
46  
47 151 control studies and can be the bases for future longitudinal ones. Finally, hypotheses generated  
48  
49 152 from epidemiologic surveys may contribute to test new basic studies and can be considered as a  
50  
51 153 complementary strategy to translational research.

#### 52 53 54 154 **Pegasus-Murcia Project**

55  
56 155 Spain actively participated in the ESEMeD Project with a representative sample of the adult  
57  
58 156 general Spanish population (n=5473) and the results have been published in national and

1  
2  
3 157 international journals.<sup>62-67</sup> However, the sample size within most of the Autonomous  
4  
5 158 Communities in Spain was too small to be able to achieve accurate and precise estimates at the  
6  
7 159 Regional level where Health Care policies are decided. Moreover, several differences between  
8  
9 160 the Autonomous Communities in Spain in important aspects related to mental health such as  
10  
11 161 socioeconomic<sup>68</sup> and territorial inequalities in health care supply and in long-term care, access  
12  
13 162 to and use of health care facilities,<sup>69</sup> premature deaths due to alcohol consumption<sup>70</sup> and the  
14  
15 163 prevalence of psychological distress<sup>71</sup> have recently been described.

16 164 Murcia is one of the 17 Autonomous Communities of Spain. It is located in the southeast of  
17  
18 165 the country on the Mediterranean coast, with a population of 1,424,063 inhabitants at the time  
19  
20 166 of the survey (INE 2008, National Statistical Institute of Spain), almost a third of them (30.7%)  
21  
22 167 living in the capital.

23  
24 168 The PEGASUS-Murcia (“Psychiatric Enquiry to General Population in Southeast Spain-  
25  
26 169 Murcia”) project has been designed in order to obtain regional data of the prevalence, burden  
27  
28 170 and care of a representative sample of the general adult population of Murcia to allow planning  
29  
30 171 of new regional mental health policies and to compare the results with the national sample of  
31  
32 172 Spain, Europe and all other countries participating in the WMH Survey Initiative. The project  
33  
34 173 also constitutes a unique opportunity to initiate a biological bank of a well-studied  
35  
36 174 representative sample of the general population.

### 37 38 39 175 **Objectives**

40  
41 176 The PEGASUS-Murcia project is a multi-purpose, observational, cross-sectional, comparative  
42  
43 177 study of the non-institutionalized general population of Murcia Region whose objective is to  
44  
45 178 improve knowledge about common psychiatric disorders in two main areas. The first one is the  
46  
47 179 epidemiology of mental disorders and protective and risk factors in the general population of  
48  
49 180 Murcia. The specific objectives are: i) to estimate the one-month, 12-month and lifetime  
50  
51 181 prevalence of the most common mental disorders, specifically, mood and anxiety disorders, in  
52  
53 182 the general population of Murcia; ii) to assess the independent association of mood and anxiety  
54  
55 183 disorders with sociodemographic factors (gender, age, education and urban/rural location) and  
56  
57 184 selected risk factors (family history, childhood experiences, religion, partnership status and  
58  
59  
60

1  
2  
3 185 sexual problems, among others); iii) to assess the quality of life of persons with the most  
4  
5 186 common psychiatric disorders and to analyze how other variables (physical medical conditions  
6  
7 187 and sociodemographic factors) may influence this outcome; iv) to assess the treatment for these  
8  
9 188 disorders and to evaluate the unmet need and the quality of care received; and v) to compare our  
10  
11 189 results with those obtained from Spain, Europe and other non-European countries, including the  
12  
13 190 United States.

14  
15 191 The second objective area is the genetic, epigenetic and transcriptomic influences associated  
16  
17 192 with mental disorders. Its specific aims include: i) the estimation of the distribution of different  
18  
19 193 candidate genes in the general population and their association with different psychiatric  
20  
21 194 disorders; ii) the identification of sensitive alleles underlying potential GxE interactions and the  
22  
23 195 study of epigenetic mechanisms involved, specifically, DNA methylation; and iii) the analysis  
24  
25 196 of gene expression alterations through transcriptomic assays.

## 27 197 **METHODS AND ANALYSIS**

### 28 29 30 198 **Study Design**

31  
32 199 The project is a cross-sectional face-to-face interview survey based on a representative sample  
33  
34 200 of the adult and non-institutionalized general population of the Murcia Region. Those who  
35  
36 201 complete the interview will be invited to provide two biological samples from their oral mucous  
37  
38 202 membranes. The target population is defined as persons aged 18 or older residing in Murcia, not  
39  
40 203 living in institutions and with an active health card (defined as persons included in PERSAN, a  
41  
42 204 regional registry that contains all residents with a Health Card which is periodically up-dated.  
43  
44 205 Exclusion criteria are: i) Confirmed irretrievable contact errors (e.g. telephone number and/or  
45  
46 206 address); ii) Institutionalized individuals (e.g. in prison, in a hospital or in another institution) or  
47  
48 207 those living outside the Autonomous Community during the survey field work; and iii)  
49  
50 208 individuals not able to understand the Spanish language or not able to conduct the questionnaire  
51  
52 209 due to his/her physical or mental condition.

### 53 54 210 **Sampling plan**

55  
56  
57  
58  
59  
60

1  
2  
3 211 The geographical area of the survey is the Murcia Region and a two-stage, stratified sampling  
4  
5 212 design has been used. The primary sampling unit is the Primary Health Centre and the second is  
6  
7 213 the individual. The sampling frame has been PERSAN, the regional health care population  
8  
9 214 database in Murcia. Primary Health Centres have been grouped into nine strata, the current  
10  
11 215 Health Care Areas in Murcia Region. The initial sample size was 4,500 adult individuals  
12  
13 216 divided into the nine Health Care Areas with proportionate allocation. A representative sample  
14  
15 217 of two centres has been chosen in each health area, without individual participant replacement.  
16  
17 218 Selection probability for each centre was known a priori and it was proportional to the size of  
18  
19 219 the centre (% of adult individuals registered in the centre) and the proportion of adult  
20  
21 220 individuals in the centre whose place of residence was rural, semi-urban or urban. Within each  
22  
23 221 of the two selected Health Centers, a stratified random sample procedure, performed for each  
24  
25 222 combination of gender, age group (18-24, 25-34, 35-49, 50-64 and 65+) and type of residence  
26  
27 223 (rural, semi-urban and urban), constitutes a stratum and individuals have been selected using  
28  
29 224 simple random sampling.  
30  
31 225 For each Health Care Area, the sample size of each stratum has been selected such that  
32  
33 226 individuals in with the same demographic characteristics had equal probability of being selected  
34  
35 227 independently of the selected [centercentre](#). If a high number of those fulfilling the exclusion  
36  
37 228 criteria in one area is reached, a fixed number of additional individuals will be released  
38  
39 229 (subsequent releases), according to the number of interviews completed in the area and  
40  
41 230 following the same selection procedure within each of the [centerscentres](#) as the ones used to  
42  
43 231 select the initial release (no new [centercentre](#) will be selected for these releases). Any  
44  
45 232 replacement of those persons who do not want to collaborate or who do not meet the non-  
46  
47 233 eligibility criteria is not allowed.

#### 234 **Survey procedures and data control**

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

1  
2  
3 239 developed by Statistics Netherlands (Herleen, the Netherlands) and designed to ease the  
4  
5 240 handling of elaborate skip and complex randomization patterns and to facilitate data entry, allow  
6  
7 241 the elaboration of some questions and direct the interviewer through the questioning sequence.

8  
9 242 Periodically, the completed interviews will be submitted to the central project Data Center  
10  
11 243 (Regional Mental Health Service, Murcia-Spain) for checking and storage following a  
12  
13 244 predetermined security procedure. All raw data will be transferred to the Hospital del Mar  
14  
15 245 Medical Research Institute (IMIM) and the Department of Health Care Policy at Harvard  
16  
17 246 University, coordinating centers of the ESEMeD and WMH Survey Initiative projects  
18  
19 247 respectively, via secure websites. The database has been declared to the Spanish Data Protection  
20  
21 248 Agency.

22  
23 249 A survey firm has been contracted to undertake the fieldwork and, in order to ensure the quality  
24  
25 250 of the survey, several strategies are being implemented: i) a one week training course for all  
26  
27 251 interviewers by WHO certified trainers on the original protocol and use of the CAPI version of  
28  
29 252 the CIDI; ii) development of a written manual to standardize the interviewing procedure and all  
30  
31 253 scientific and administrative elements that could affect comparability of data; iii) regular  
32  
33 254 meetings with the survey firm to ensure adherence to the protocol and to deal with any difficulty  
34  
35 255 that may have arisen; and iv) data quality analysis to detect any inconsistencies and/or  
36  
37 256 incomplete data.

38  
39 257 The survey firm has been provided with sufficient data to allow contact with each of the  
40  
41 258 individuals of the selected sample and only after 10 unsuccessful attempts the person will be  
42  
43 259 considered ~~to be as not-~~uncontactable or after confirmation that the selected person does not live  
44  
45 260 at that address and new contact information is unavailable. Several methods will be used to  
46  
47 261 improve the participation of those selected: i) an informative flyer providing general  
48  
49 262 information related to the project and giving notice of future contact will be sent by  
50  
51 263 conventional post together with an invitation letter signed by a person from the Health Care  
52  
53 264 Authority; ii) a phone call ~~-to~~ invite them to participate in the interview process and to offer  
54  
55 265 them the possibility to do the interview either at home or in their Primary Care Center; iii)  
56  
57 266 Several informative sessions for the healthcare personnel of the Primary Care Centers will be  
58  
59  
60

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

## 272 **The survey questionnaire**

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

### 277 *The Composite International Diagnostic Interview (CIDI)*

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

290 **Table 1: Description of the adapted version of the World Health Organization -Composite**  
 291 **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 <sup>‡</sup>	Psychosis	42 to 84	All respondents
CFQ <sup>§</sup>	Risk Factors	25	All respondents
SLE <sup>¶</sup>	Risk Factors	13 to 39	All respondents
Neuroticism and Extroversion subscales <sup>¶</sup>	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 Childhood	Socio-demographics	25-36	Long path
Attention Hyperactivity	Risk/Protective Factors	110	Long path
Oppositional Defiant	Childhood	90	Long path and Screening
Conduct Disorder	Childhood	46	Long path and Screening
Separation Anxiety Disorder	Childhood	54	Long path
Family Burden	Childhood	86	Screening questions (SCR)
Quality/ Lie subscale	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

292 <sup>§</sup> EQ-5D: European Quality of Life Scale; <sup>\*</sup> SF-12 v2: Short Form 12 Health Questionnaire; <sup>†</sup> Lie subscale of the abbreviated version  
 293 of the Eysenck Personality Questionnaire (EPQR-A); <sup>‡</sup> CAPE: Community Assessment of Psychic Experiences; <sup>§</sup> CFQ: Cognitive  
 294 Failure Questionnaire; <sup>¶</sup> SLE: Stressful Life Events; <sup>¶</sup> Neuroticism and Extroversion subscales of the abbreviate version of the  
 295 Eysenck Personality Questionnaire (EPQR-A)

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

1  
2  
3 300  
4  
5 301 The most recent version of the CIDI (version 3.0) is the end result of a number of international  
6  
7 302 studies and adaptations made since 2000 when it was first used in WMH surveys. It was first  
8  
9 303 created in English and has been translated into more than 30 different languages using the  
10  
11 304 standard WHO protocol with a rigorous process of adaptation.<sup>73,74</sup> Several clinical reappraisal  
12  
13 305 studies have been carried out and the concordance of the CIDI version 3.0 has been evaluated in  
14  
15 306 different subgroups of WMH surveys using the Structured Clinical Interview for DSM-IV  
16  
17 307 (SCID) as the clinical gold standard and a moderate to excellent concordance has been found for  
18  
19 308 most mental disorders.<sup>75,76</sup> CIDI is available in two formats: the paper form or PAPI (Paper and  
20  
21 309 Pencil Interviewing) and the computerized form or CAPI (Computer Assisted Personal  
22  
23 310 Interviewing), designed to ease the handling of elaborate skip and complex randomization  
24  
25 311 patterns and to facilitate data entry with a resulting reduction in interview time and errors in data  
26  
27 312 collection and recording. The original Spanish CAPI version used in Spain had not been  
28  
29 313 updated since it was used in the context of the ESEMeD project almost ten years ago. Since  
30  
31 314 then, all improvements in the questionnaire have only been added to the CIDI Latin American  
32  
33 315 (LA) v20.0 version. However, due to linguistic and cultural differences in Spanish-speaking  
34  
35 316 populations, this CAPI version had to be culturally adapted for use in Spain by our research  
36  
37 317 team and this process is fully described elsewhere.<sup>77</sup>  
38  
39 318 To further shorten the length of the questionnaire, some sections were not selected for the  
40  
41 319 purposes of this project. These include Intermittent Explosive Disorder, Personality I and II,  
42  
43 320 Neurasthenia and Pre-Menstrual and Gambling sections. Some others were substituted by other  
44  
45 321 questions or questionnaires, e.g. the Tobacco Use section was simplified using some questions  
46  
47 322 obtained from the Spanish National Health Survey and the Psychosis section with the CAPE  
48  
49 323 instrument (Community Assessment of Psychic Experiences), both described below.

51  
52 324 *Other study instruments*

53  
54 325 Several other instruments were added to the original CIDI for the specific purposes of the  
55  
56 326 PEGASUS-Murcia project. These include the Spanish version of different questionnaires: i)  
57  
58 327 Mini-Mental State Examination for interviewees older than 60 years old,<sup>78,79</sup> ii) the Cognitive



1  
2  
3 328 Failure Questionnaire (CFQ),<sup>80,81</sup> iii) the Neuroticism, Extroversion and Lie subscales of the  
4  
5 329 abbreviated version of the Eysenck Personality Questionnaire (EPQR-A);<sup>82-84</sup> iv) the Resilience  
6  
7 330 Scale;<sup>85,86</sup> v) the Community Assessment of Psychic Experiences (CAPE)<sup>87</sup> to measure  
8  
9 331 attenuated psychotic symptoms in the general population instead of the Psychosis section of the  
10  
11 332 CIDI, as the latest is only used as a screening instrument in the detection of psychosis. Those  
12  
13 333 who positively answer two items of the positive dimension with a score equal or superior to 3,  
14  
15 334 have been hospitalized for psychiatric reasons and/or have received psychotropic medication  
16  
17 335 during the last year will be evaluated by a clinic psychiatrist with the module C (Psychotic  
18  
19 336 Disorders) of the SCID (Structured Clinical Interview for DSM Disorders) ; vi) a brief list of 12  
20  
21 337 stressful life events in the last 12 months was included by the combination of a List of  
22  
23 338 Threatening Experiences (LTE)<sup>88,89</sup> and the emotional and life-changing impact of each event;<sup>90</sup>  
24  
25 339 vii) the European Quality of Life Scale (EuroQol 5D)<sup>91</sup> and the Short Form 12 Health  
26  
27 340 Questionnaire (SF-12 v2);<sup>92</sup> viii) an ad-hoc questionnaire of partner violence obtained from the  
28  
29 341 Spanish National Health Survey and from the regional mental health clinical guidelines;<sup>93</sup> and,  
30  
31 342 finally, ix) some questions related to tobacco use and physical exercises from the Spanish  
32  
33 343 National Health Survey.

#### 344 **Questionnaire pathways**

345 In order to optimize the duration of the interview, the WMH questionnaire was divided into two  
346 parts with questions in Part 1 administered to all respondents and those in Part 2 only to a  
347 subsample of individuals who followed the long path of the interview. Part 2 of the interview  
348 includes detailed information about a wide range of aspects related to the primary disorders and  
349 also to mental disorders of secondary interest (Table 1). The inclusion criteria for the long path  
350 are: a) all individuals that could be considered as “high risk individuals” because they positively  
351 answer a number of specific questions related to mood and anxiety disorders and b) a random  
352 subsample (25%) of the respondent without symptoms (“low risk individuals”). The remaining  
353 75% of respondents without screening symptoms not randomly selected for the long path  
354 followed the short path. ~~All these pathways are automatically made by the computer without~~  
355 ~~any intervention of the interviewer~~The computer, without any intervention of the interviewer.

1  
2  
3 356 | automatically makes all these pathways. -In this shorter itinerary, ~~the sections omitted were~~  
4  
5 357 | ~~substituted by a specific section that included those questions needed to calculate some~~  
6  
7 358 | ~~demographic indicators~~ a specific section, that included those questions needed to calculate some  
8  
9 359 | demographic indicators, substituted the sections omitted. Moreover, two sections were only  
10  
11 360 | used in a percentage of the long path itinerary, Eating Disorders (50 %) and Obsessive-  
12  
13 361 | Compulsive Disorder (33%).

### 362 **Quality control procedures**

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

### 380 **Laboratory Methods**

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

1  
2  
3 384 | samples will be obtained only if the interviewee signs an informed consents specifically  
4  
5 385 | designed for this project based on international recommendations for population-based research  
6  
7 386 | involving genetics;<sup>94</sup> and previously approved by~~from each interviewee, one for DNA~~  
8  
9 387 | ~~extraction for genomic and epigenomic analysis and the other one to obtain mRNA for gene~~  
10  
11 388 | ~~expression quantification (transcriptomic assays t)-~~he Regional Ethics Research Committee.  
12  
13 389 | ~~These samples will be taken using swabs compatible with molecular amplification techniques,~~  
14  
15 390 | ~~as they do not interfere with the amplification process (FLOQSwabs Flocked Swabs, Copan~~  
16  
17 391 | ~~Flock Technologies srl).~~Interviewers have been trained by one of the authors (TE) to  
18  
19 392 | adequately obtain the biological sample by scraping the oral mucosa using swabs compatible  
20  
21 393 | with molecular amplification techniques, as they do not interfere with the amplification process  
22  
23 394 | (FLOQSwabs Flocked Swabs, Copan Flock Technologies srl).

24  
25 395 | Samples for DNA extraction will be collected in sterile 1.5 ml tubes. Those to be used for RNA  
26  
27 396 | extraction will be harvested in dark sterile tubes containing RNA protect cell Reagent  
28  
29 397 | (QIAGEN, Hilden, Germany), which provides immediate stabilization of RNA. Cells will be  
30  
31 398 | thus stabilized at room temperature and can then be stored or transported at ambient temperature  
32  
33 399 | prior to RNA purification. Tubes will be labeled with tags (14C.B. 40X40 type) with a specific  
34  
35 400 | code for each sample and will be packaged and sent to BIOBANC-MUR ~~Mur~~ (the biobank for  
36  
37 401 | biomedical research network of the Region of Murcia, RD09/0076/00065, as a partner of the  
38  
39 402 | Spanish National Biobanks Network; IMIB: Instituto Murciano de Investigación Biosanitaria)  
40  
41 403 | according to current Spanish legislation and following the regulations of the International Air  
42  
43 404 | Transport Association (IATA) on biological sample shipping.

44  
45 405 | Those sample accepted by BIOBANC-MUR~~ur~~ will be registered using a specific biobanking  
46  
47 406 | software (bio-e-bank, VITROSOFT, SL), as part of a Laboratory Integrated Management  
48  
49 407 | System (LIMS). The nucleic acid extraction will be performed automatically (QIAcube system;  
50  
51 408 | QIAGEN, Hilden, Germany) to minimize variability due to manual handling using QIAamp  
52  
53 409 | DNA Blood Mini Kit and RNeasyPlus Mini Kit (QIAGEN, Hilden, Germany) for DNA and  
54  
55 410 | RNA extraction, respectively.

1  
2  
3 411 | [QIAamp DNA Blood Mini Kit provide fast and easy method for purification of total DNA for](#)  
4 [reliable PCR and Southern blotting from whole human blood, buffy coat, cultured cells,](#)  
5 [lymphocytes, plasma, serum, body fluids, and buccal swabs.](#) The synthesis of complementary  
6  
7  
8  
9 414 | DNA (cDNA) from mRNA for expression studies will be developed for all samples by reverse  
10  
11 415 | transcription using the *High Capacity cDNA Reverse Transcription Kit* (Applied Biosystems).  
12  
13 416 | All processes will be performed according to the manufacturer's instructions.

14  
15 417 | ~~Nucleic acids DNA and RNA quantity and quality will be quantitated by~~  
16 ~~measuringdetermined by the ratio A260/280 calculated based on 260 and 280 nm absorbance~~  
17 ~~measured absorbance at 260/280 nm~~ using a spectrophotometer.<sup>95-97</sup> ~~The ratio Between 260 nm~~  
18 ~~and 230 nm (A260/230) absorbance~~ is commonly used as a secondary indicator of nucleic acid  
19 419 | purity<sup>98-100</sup>. The integrity of DNA will be visualized by electrophoresis on 1% agarose gel  
20  
21 420 | (migration for 1 hour at 100 V) using 100 ng of total DNA and a 23 kb DNA ladder (Lambda  
22  
23 421 | DNA/HindIII Marker (Thermo Fisher Scientific) as DNA marker. All mRNA samples will be  
24  
25 422 | transformed into cDNA.  
26  
27 423 |  
28  
29 424 |

30  
31 425 | Specially trained technicians [from the BIOBANC-MUR](#) will be used to monitor the specimen  
32  
33 426 | collection by donors and to perform sample manipulations in order to minimize variability of  
34  
35 427 | results and to obtain the optimal quality of nucleic acids for this and future studies. The  
36  
37 428 | processed biospecimens (150 µl of DNA and 80 µl of cDNA) will be stored in 750 µl  
38  
39 429 | microtubes in an ultra-freezer at -80 °C located in BIOBANC-MUR.

#### 430 | **Statistical methods**

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

#### 436 | **Weighting procedures**

437 | Given that the interview is divided into two parts and only a portion of the sample will be  
438 | 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  $w_s = 1/p_{hc} p_{hcsgr}^1$ , where  $p_{hc}$  is the probability that the centre  $c$  was selected,  $p_{hcsgr}^1 = n_{hcsgr} / N_{hcsgr}$  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/](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  $w_{adj} = w_{unadj} w_{ps}$ .

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  $\pi_i$  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:  $w_{p2norm} = w_{p2adj} n / \sum_{i=1}^n w_{p2adj_i}$ .

443

444

### 445 **Analysis of the data and forthcoming research projects**

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

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

### 463 **ETHICS AND DISSEMINATION**

464 Eligible individuals will be asked to sign two independent informed consents to participate, the  
465 first one to be interviewed, including the possibility of future new contacts and the second to  
466 provide the biological samples but only those who had already completed the questionnaire.

467 Name and contact information will be stored separately from any information provided as part  
468 of the study questionnaire. The ~~protocol was approved by the Clinical Research Ethics~~  
469 ~~Committee of the University Hospital Virgen de la Arrixaca of Murcia~~Clinical Research Ethics  
470 Committee of the University Hospital Virgen de la Arrixaca of Murcia approved the protocol  
471 and the database of personal information ~~was~~has been registered with the National Data

1  
2  
3 472 Protection Agency. Data from PEGASUS-Murcia project will be included in the WMH Cross  
4  
5 473 National Sample for international comparisons. The study findings will be submitted to peer-  
6  
7 474 reviewed journals for publication, and presented at national and international scientific  
8  
9 475 meetings.

## 11 476 **DISCUSSION**

12  
13  
14 477 The epidemiology of mental illnesses is a fascinating but highly complex area of research. This  
15  
16 478 complexity is primarily due to the wide range of factors, environmental and genetic, which  
17  
18 479 combines to produce a recognized psychiatric disorder. Previous epidemiological research has  
19  
20 480 resulted in the production of a great amount of data but it has been difficult to make cross-  
21  
22 481 national comparisons due to methodological variability. The WMH Survey Initiative aimed to  
23  
24 482 address this issue by using an international standardized protocol, allowing comparisons of the  
25  
26 483 most common mental disorders and their associated factors throughout the world. Using this  
27  
28 484 study design, it therefore offers the opportunity for new surveys to be performed in the context  
29  
30 485 of an international collaborative initiative and the possibility to adapt the questionnaire  
31  
32 486 according to the specific aims of the research being undertaken. The PEGASUS-Murcia project  
33  
34 487 can be considered as an example of how the latter has been successfully achieved. It is a cross-  
35  
36 488 sectional study designed to assess the prevalence of the most frequent mental disorders and their  
37  
38 489 correlates in a representative sample of the general population of Murcia. Its primary strengths  
39  
40 490 are: i) the fact that it was specifically adapted to assess factors not only associated with mental  
41  
42 491 disorders but also with positive mental health in a representative sample of the general  
43  
44 492 population; ii) its context focused on regional needs where healthcare decisions are taken  
45  
46 493 regarding resource allocation and mental health planning; iii) the collection of biological  
47  
48 494 samples not only for DNA analysis but also for mRNA; iv) all the information collected in our  
49  
50 495 study, including biological samples, can be correlated with past and future health events because  
51  
52 496 all Spanish population had free access to the Healthcare System at the time of its inception and  
53  
54 497 were thus registered and provided with a unique identification number and therefore; v) finally,  
55  
56 498 the inclusion of a multidisciplinary research team is in accordance with the international  
57  
58  
59  
60

1  
2  
3 499 consensus regarding the need for interdisciplinary collaboration between clinicians,  
4  
5 500 epidemiologists and neuroscience researchers to increase their combined efforts to study the  
6  
7 501 complex gene-gene and gene-environmental interactions underlying mental health  
8  
9 502 disorders.<sup>23,61,103,104</sup>

10  
11 503 Concerns have been expressed about the cost-effectiveness of psychiatric epidemiological  
12  
13 504 surveys, such as World Mental Health 2000 (WMH-2000) projects,<sup>105</sup> an example being the  
14  
15 505 rationale for starting a new psychiatric epidemiological survey in the Autonomous Community  
16  
17 506 of Murcia if Spain had already participated in the ESEMeD project. However, there are several  
18  
19 507 reasons to justify this regional initiative. Firstly, public health and healthcare agencies usually  
20  
21 508 allocate mental health resources, including human, based on data from national epidemiologic  
22  
23 509 surveys,<sup>106</sup> such as that provided by the Spanish participation in the ESEMeD Project. As  
24  
25 510 previously mentioned, the involvement of the Region of Murcia in the Spanish ESEMeD survey  
26  
27 511 did not allow evaluation of specific regional data. Nowadays, the main responsibility for  
28  
29 512 planning and management of Healthcare resources in Spain lies with the Autonomous  
30  
31 513 Communities and differences exist between them in terms of accessibility, amount of health-  
32  
33 514 care resources and political decision-making.<sup>68-71</sup> Devolution of this responsibility to Murcia  
34  
35 515 occurred in December 2001.

36  
37 516 Secondly, the inclusion of biological data in a well-designed multidisciplinary epidemiological  
38  
39 517 study offers great advantages in terms of a more global understanding of mental disorders.  
40  
41 518 These are complex illnesses of the brain where social, familial, psychological and biological  
42  
43 519 elements interact throughout the entire life of a person to influence his/her risk of developing a  
44  
45 520 mental health disorder. To extend our understanding of the physiopathology and epidemiology  
46  
47 521 of the more common ones (mood and anxiety), it is necessary to identify genetic loci and  
48  
49 522 polymorphic alleles and their distribution in the healthy and affected population whose function  
50  
51 523 in determining risk for, and protection against, these conditions probably depends on gene-gene  
52  
53 524 and GxE interactions. The collection of genetic material from representative samples from the  
54  
55 525 general population, well described using international diagnostic instruments such as CIDI,  
56  
57 526 offers new and different possibilities to evaluate candidate genes in non-biased samples and to  
58  
59  
60



1  
2  
3 527 describe their distribution in the general population that may contribute to clarification of the  
4  
5 528 complexity of mental disorders.

6  
7 529 Thirdly, our project involving a multidisciplinary research team gives new opportunities to  
8  
9 530 develop different study designs that can move from descriptive to analytical epidemiology. For  
10  
11 531 example, this representative sample constitutes a good source of controls for future case-control  
12  
13 532 studies, where cases will be provided from the public health care clinics, and can be the starting  
14  
15 533 point for future cohort studies. Our project was designed to allow for all these possibilities.

#### 16 17 534 **Limitations of the study**

18  
19 535 Currently, the main limitations of the PEGASUS-Murcia project are related to: i) the cross-  
20  
21 536 sectional design which, while it allows association studies, limits the possible causal  
22  
23 537 interpretation of the findings. However, these findings may provide new hypotheses and enable  
24  
25 538 the design of new studies; ii) not all interviewees will provide biological samples and this may  
26  
27 539 affect the representativeness of some mental disorders in future analyses. To determine if this  
28  
29 540 will result in selection bias, we will analyze whether there are distinguishing characteristics  
30  
31 541 between donors and non-donors in the distribution of mental disorders and other characteristics  
32  
33 542 of the participants; iii) the population stratification in our study which will be used for future  
34  
35 543 genetic association analyses is performed by using the stated ancestral origin by participants<sup>107</sup>  
36  
37 544 instead of using genetic markers; and iv) biological samples will be obtained from oral mucosal  
38  
39 545 scrapings and not from brain neurons. However, this is a general situation given the ethical  
40  
41 546 issues and difficulties in obtaining neural tissues and, in any case, gene expression does not  
42  
43 547 appear to be specific to neural tissue, at least in some genes that have ubiquitous expression, for  
44  
45 548 example, 5-HTTLPR.<sup>108-111</sup>

#### 46 47 549 **Conclusions and Future Directions**

48  
49 550 The PEGASUS-Murcia project is a sound bases for multidisciplinary collaborative mental  
50  
51 551 health research studies which will provide not only a huge amount of epidemiological  
52  
53 552 information but will also offer exiting opportunities to clarify the complex interactions between  
54  
55 553 genetic and environmental factors which result in a range of mental health disorders.

## 554 **Competing interests**

555 The authors declare that they have no competing interests.

## 556 **Author's contributions**

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

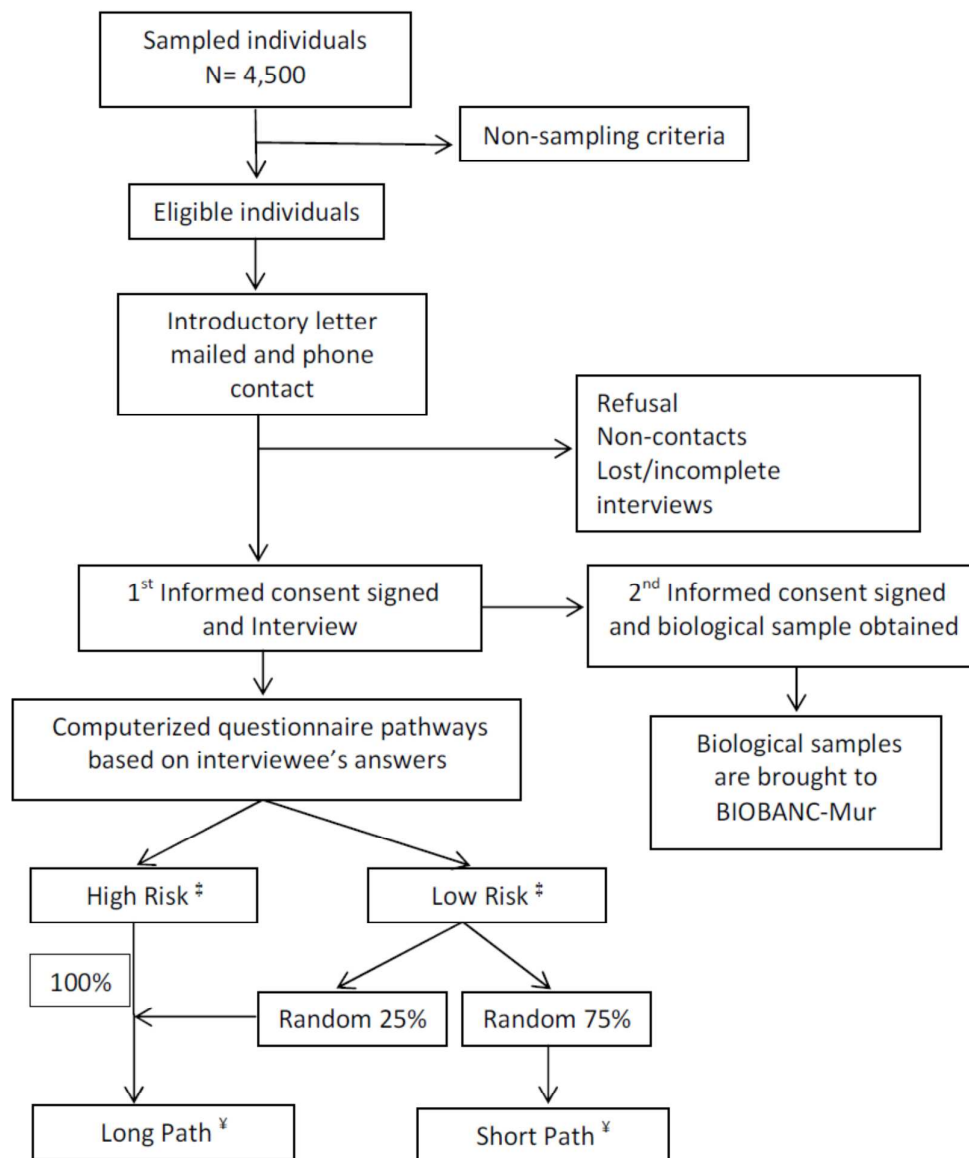
564 **Acknowledgments:** We thank Carlos Giribert Muñoz, Deputy Director of Mental  
565 Health and Psychiatric Services of Murcia for his support in developing the PEGASUS-Murcia  
566 project; Inés Morán-Sánchez, M<sup>a</sup> Luisa Pujalte and Ascensión Garriga for their contribution to  
567 the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

573 The PEGASUS-Murcia project ~~was~~ is supported by the Regional Health Authorities of Murcia  
574 (“Servicio Murciano de Salud and Consejería de Sanidad y Política Social”) (Decreto n<sup>o</sup>  
575 455/2009), the “Fundación para la Formación e Investigación Sanitarias (FFIS) de la Región de  
576 Murcia” (N<sup>o</sup> Expedientes: CM0829 I and FFIDS/EMER09/14) and the “Ayudas para proyectos  
577 de Investigación en Salud –ISCIH- del Plan Nacional de Investigación Científica, Desarrollo e  
578 Innovación Tecnológica” (PI12/00809). None of the funders had any role in the design,  
579 analysis, interpretation of results, or preparation of this paper. The PEGASUS-Murcia project is  
580 carried out in conjunction with the WHOWMH Survey Initiative. We thank the WMH  
581 Coordinating Center staff at Harvard and Michigan Universities and, especially Professor Ron  
582 Kessler, for their assistance with the instrumentation, fieldwork and data analysis. These  
583 activities were supported by the United States National Institute of Mental Health  
584 (R01MH070884), the John D. and Catherine T. MacArthur Foundation, the Pfizer Foundation,  
585 the U.S. Public Health Service (R13-MH066849, R01-MH069864, and R01 DA016558), the  
586 Fogarty International Center (FIRCA R03- TW006481), the Pan American Health Organization,  
587 the Eli Lilly & Company Foundation, Ortho-McNeil Pharmaceutical, Inc., GlaxoSmithKline,

1  
2  
3 588 Bristol-Myers Squibb and Shire. A complete list of WMH publications can be found at  
4 589 <http://www.hcp.med.harvard.edu/wmh/>.  
5  
6 590  
7 591  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For peer review only

592 **Figure 1: Flow chart of the PEGASUS-Murcia project**  
593



594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608

† The response rate is defined as:  $(\text{completed interviews}) / (\text{total released respondent sample cases} - \text{respondent nonsample cases})$ .

‡ **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

609

## 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.

- 1  
2  
3 636 10. Alonso J, Angermeyer MC, Bernert S et al. Prevalence of mental disorders in Europe:  
4  
5 637 results from the European Study of the Epidemiology of Mental Disorders (ESEMeD)  
6  
7 638 project. *Acta Psychiatr Scand Suppl.* 2004;21-27.  
8  
9 639 11. Alonso J, Angermeyer MC, Bernert S et al. 12-Month comorbidity patterns and associated  
10  
11 640 factors in Europe: results from the European Study of the Epidemiology of Mental  
12  
13 641 Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl.* 2004;28-37.  
14  
15 642 12. Alonso J, Angermeyer MC, Bernert S et al. Disability and quality of life impact of mental  
16  
17 643 disorders in Europe: results from the European Study of the Epidemiology of Mental  
18  
19 644 Disorders (ESEMeD) project. *Acta Psychiatr Scand.* 2004;420:38-46.  
20  
21 645 13. Alonso J, Angermeyer MC, Bernert S et al. Use of mental health services in Europe: results  
22  
23 646 from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project.  
24  
25 647 *Acta Psychiatr Scand Suppl.* 2004;47-54.  
26  
27 648 14. Alonso J, Buron A, Bruffaerts R et al. Association of perceived stigma and mood and  
28  
29 649 anxiety disorders: results from the World Mental Health Surveys. *Acta Psychiatr Scand.*  
30  
31 650 2008;118:305-314.  
32  
33 651 15. Alonso J, Codony M, Kovess V et al. Population level of unmet need for mental healthcare  
34  
35 652 in Europe. *Br J Psychiatry.* 2007;190:299-306.  
36  
37 653 16. Codony M, Alonso J, Almansa J et al. Perceived need for mental health care and service use  
38  
39 654 among adults in Western Europe: results of the ESEMeD project. *Psychiatr Serv.*  
40  
41 655 2009;60:1051-1058.  
42  
43 656 17. Sevilla-Dedieu C, Kovess-Masfety V, Gilbert F et al. Mental health care and out-of-pocket  
44  
45 657 expenditures in Europe: results from the ESEMeD project. *J Ment Health Policy Econ.*  
46  
47 658 2011;14:95-105.  
48  
49 659 18. Hyman SE. The millennium of mind, brain, and behavior. *Arch Gen Psychiatry.*  
50  
51 660 2000;57:88-89.  
52  
53 661 19. Kubota T, Miyake K, Hirasawa T. Epigenetic understanding of gene-environment  
54  
55 662 interactions in psychiatric disorders: a new concept of clinical genetics. *Clin*  
56  
57 663 *Epigenetics.* 2012;4:1.  
58  
59  
60

- 1  
2  
3 664 20. Merikangas KR, Zhang H, Avenevoli S et al. Longitudinal trajectories of depression and  
4  
5 665 anxiety in a prospective community study: the Zurich Cohort Study. *Arch Gen*  
6  
7 666 *Psychiatry*. 2003;60:993-1000.  
8  
9 667 21. Toyokawa S, Uddin M, Koenen KC et al. How does the social environment 'get into the  
10  
11 668 mind'? Epigenetics at the intersection of social and psychiatric epidemiology. *Soc Sci*  
12  
13 669 *Med*. 2012;74:67-74.  
14  
15 670 22. Tsuang M, Bar J, Stone W et al. Gene-environment interactions in mental disorders. *World*  
16  
17 671 *Psychiatry*. 2004;3:73-83.  
18  
19 672 23. Caspi A, Moffitt TE. Gene-environment interactions in psychiatry: joining forces with  
20  
21 673 neuroscience. *Nat Rev Neurosci*. 2006;7:583-590.  
22  
23 674 24. Caspi A, Hariri AR, Holmes A et al. Genetic sensitivity to the environment: the case of the  
24  
25 675 serotonin transporter gene and its implications for studying complex diseases and traits.  
26  
27 676 *Am J Psychiatry*. 2010;167:509-527.  
28  
29 677 25. Koenen KC, Fu QJ, Ertel K et al. Common genetic liability to major depression and  
30  
31 678 posttraumatic stress disorder in men. *J Affect Disord*. 2008;105:109-115.  
32  
33 679 26. Norris FH, Friedman MJ, Watson PJ et al. 60,000 disaster victims speak: Part I. An  
34  
35 680 empirical review of the empirical literature, 1981-2001. *Psychiatry*. 2002;65:207-239.  
36  
37 681 27. Caspi A, Hariri AR, Holmes A et al. Genetic sensitivity to the environment: the case of the  
38  
39 682 serotonin transporter gene and its implications for studying complex diseases and traits.  
40  
41 683 *Am J Psychiatry*. 2010;167:509-527.  
42  
43 684 28. Clarke H, Flint J, Attwood AS et al. Association of the 5-HTTLPR genotype and unipolar  
44  
45 685 depression: a meta-analysis. *Psychol Med*. 2010;40:1767-1778.  
46  
47 686 29. Karg K, Burmeister M, Shedden K et al. The serotonin transporter promoter variant (5-  
48  
49 687 HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic  
50  
51 688 moderation. *Arch Gen Psychiatry*. 2011;68:444-454.  
52  
53 689 30. Takano A, Arakawa R, Hayashi M et al. Relationship between neuroticism personality trait  
54  
55 690 and serotonin transporter binding. *Biol Psychiatry*. 2007;62:588-592.  
56  
57  
58  
59  
60

- 1  
2  
3 691 31. Koenen K, Kilpatrick D, Acierno R et al. Serotonin transporter gene and social support  
4 moderate post-traumatic stress disorder in hurricane exposed adults. *American Journal*  
5 692 *of Medical Genetics Part B-Neuropsychiatric Genetics*. 2006;141B:742-743.  
6 693  
7  
8  
9 694 32. Risch N, Herrell R, Lehner T et al. Interaction between the serotonin transporter gene (5-  
10 HTTLPR), stressful life events, and risk of depression: a meta-analysis. *JAMA*.  
11 695 2009;301:2462-2471.  
12 696  
13  
14  
15 697 33. Munafo MR, Durrant C, Lewis G et al. Gene X environment interactions at the serotonin  
16 transporter locus. *Biol Psychiatry*. 2009;65:211-219.  
17 698  
18  
19 699 34. Navarro-Mateu F, Escámez T, Koenen KC et al. Meta-Analyses of the 5-HTTLPR  
20 Polymorphisms and Post-Traumatic Stress Disorder. *PLOS One*. 2013;8:e66227.  
21 700  
22  
23 701 35. Kendler KS, Neale MC, Kessler RC et al. A longitudinal twin study of personality and  
24 major depression in women. *Arch Gen Psychiatry*. 1993;50:853-862.  
25 702  
26  
27 703 36. Kotov R, Gamez W, Schmidt F et al. Linking "big" personality traits to anxiety, depressive,  
28 and substance use disorders: a meta-analysis. *Psychol Bull*. 2010;136:768-821.  
29 704  
30  
31 705 37. Wu D, Yin H, Xu S et al. Risk factors for posttraumatic stress reactions among Chinese  
32 students following exposure to a snowstorm disaster. *BMC Public Health*. 2011;11:96.  
33 706  
34  
35 707 38. Hettema JM, Neale MC, Myers JM et al. A population-based twin study of the relationship  
36 between neuroticism and internalizing disorders. *Am J Psychiatry*. 2006;163:857-864.  
37 708  
38  
39 709 39. Ploubidis GB, Frangou S. Neuroticism and psychological distress: to what extent is their  
40 association due to person-environment correlation? *European Psychiatry*. 2011;26:1-5.  
41 710  
42  
43 711 40. Skelton K, Ressler KJ, Norrholm SD et al. PTSD and gene variants: new pathways and new  
44 thinking. *Neuropharmacology*. 2012;62:628-637.  
45 712  
46  
47 713 41. Surtees PG, Wainwright NW, Willis-Owen SA et al. Social adversity, the serotonin  
48 transporter (5-HTTLPR) polymorphism and major depressive disorder. *Biol Psychiatry*.  
49 714 2006;59:224-229.  
50 715  
51  
52  
53 716 42. Stein MB, Schork NJ, Gelernter J. Gene-by-environment (serotonin transporter and  
54 childhood maltreatment) interaction for anxiety sensitivity, an intermediate phenotype  
55 for anxiety disorders. *Neuropsychopharmacology*. 2008;33:312-319.  
56 717  
57  
58 718  
59  
60



- 1  
2  
3 719 43. Koenen KC, Aiello AE, Bakshis E et al. Modification of the association between serotonin  
4  
5 720 transporter genotype and risk of posttraumatic stress disorder in adults by county-level  
6  
7 721 social environment. *Am J Epidemiol*. 2009;169:704-711.  
8  
9 722 44. Galea S, Uddin M, Koenen K. The urban environment and mental disorders: Epigenetic  
10  
11 723 links. *Epigenetics*. 2011;6:400-404.  
12  
13 724 45. Esteller M. The necessity of a human epigenome project. *Carcinogenesis*. 2006;27:1121-  
14  
15 725 1125.  
16  
17 726 46. Narayan P, Dragunow M. Pharmacology of epigenetics in brain disorders. *Br J Pharmacol*.  
18  
19 727 2010;159:285-303.  
20  
21 728 47. Olsson CA, Foley DL, Parkinson-Bates M et al. Prospects for epigenetic research within  
22  
23 729 cohort studies of psychological disorder: a pilot investigation of a peripheral cell marker  
24  
25 730 of epigenetic risk for depression. *Biol Psychol*. 2010;83:159-165.  
26  
27 731 48. Koenen KC, Uddin M, Chang SC et al. SLC6A4 methylation modifies the effect of the  
28  
29 732 number of traumatic events on risk for posttraumatic stress disorder. *Depress Anxiety*.  
30  
31 733 2011;28:639-647.  
32  
33 734 49. Smith CL, Bolton A, Nguyen G. Genomic and Epigenomic Instability, Fragile Sites,  
34  
35 735 Schizophrenia and Autism. *Current Genomics*. 2010;11:447-469.  
36  
37 736 50. Dempster EL, Pidsley R, Schalkwyk LC et al. Disease-associated epigenetic changes in  
38  
39 737 monozygotic twins discordant for schizophrenia and bipolar disorder. *Hum Mol Genet*.  
40  
41 738 2011;20:4786-4796.  
42  
43 739 51. Park BY, Lee BC, Jung KH et al. Epigenetic changes of serotonin transporter in the patients  
44  
45 740 with alcohol dependence: methylation of an serotonin transporter promoter CpG island.  
46  
47 741 *Psychiatry Investig*. 2011;8:130-133.  
48  
49 742 52. Klengel T, Mehta D, Anacker C et al. Allele-specific FKBP5 DNA demethylation mediates  
50  
51 743 gene-childhood trauma interactions. *Nat Neurosci*. 2013;16:33-41.  
52  
53 744 53. van Os J, Kenis G, Rutten BP. The environment and schizophrenia. *Nature*. 2010;468:203-  
54  
55 745 212.  
56  
57  
58  
59  
60

- 1  
2  
3 746 54. Lee YH, Wong DT. Saliva: an emerging biofluid for early detection of diseases. *Am J Dent*.  
4  
5 747 2009;22:241-248.  
6  
7 748 55. Yan W, Apweiler R, Balgley BM et al. Systematic comparison of the human saliva and  
8  
9 749 plasma proteomes. *Proteomics Clin Appl*. 2009;3:116-134.  
10  
11 750 56. Ballantyne J. Validity of messenger RNA expression analyses of human saliva. *Clin Cancer*  
12  
13 751 *Res*. 2007;13:1350.  
14  
15 752 57. Park NJ, Zhou X, Yu T et al. Characterization of salivary RNA by cDNA library analysis.  
16  
17 753 *Arch Oral Biol*. 2007;52:30-35.  
18  
19 754 58. Lee YH, Zhou H, Reiss JK et al. Direct saliva transcriptome analysis. *Clin Chem*.  
20  
21 755 2011;57:1295-1302.  
22  
23 756 59. Palanisamy V, Wong DT. Transcriptomic analyses of saliva. *Methods Mol Biol*.  
24  
25 757 2010;666:43-51. doi: 10.1007/978-1-60761-820-1\_4.:43-51.  
26  
27 758 60. Merikangas KR, Risch N. Will the Genomics Revolution Revolutionize Psychiatry?  
28  
29 759 *American Journal of Psychiatry*. 2003;160:625-635.  
30  
31 760 61. Weissman MM, Brown AS, Talati A. Translational epidemiology in psychiatry: linking  
32  
33 761 population to clinical and basic sciences. *Arch Gen Psychiatry*. 2011;68:600-608.  
34  
35 762 62. Autonell J, Vila F, Pinto-Meza A et al. [One year prevalence of mental disorders  
36  
37 763 comorbidity and associated socio-demographic risk factors in the general population of  
38  
39 764 Spain. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:4-11.  
40  
41 765 63. Haro JM, Palacin C, Vilagut G et al. [Prevalence of mental disorders and associated factors:  
42  
43 766 results from the ESEMeD-Spain study]. *Med Clin (Barc)*. 2006;126:445-451.  
44  
45 767 64. Codony M, Alonso J, Almansa J et al. [Mental health care use in the Spanish general  
46  
47 768 populations. Results of the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:21-28.  
48  
49 769 65. Codony M, Alonso J, Almansa J et al. [Psychotropic medications use in Spain. Results of  
50  
51 770 the ESEMeD-Spain study]. *Actas Esp Psiquiatr*. 2007;35:29-36.  
52  
53  
54  
55  
56  
57  
58  
59  
60

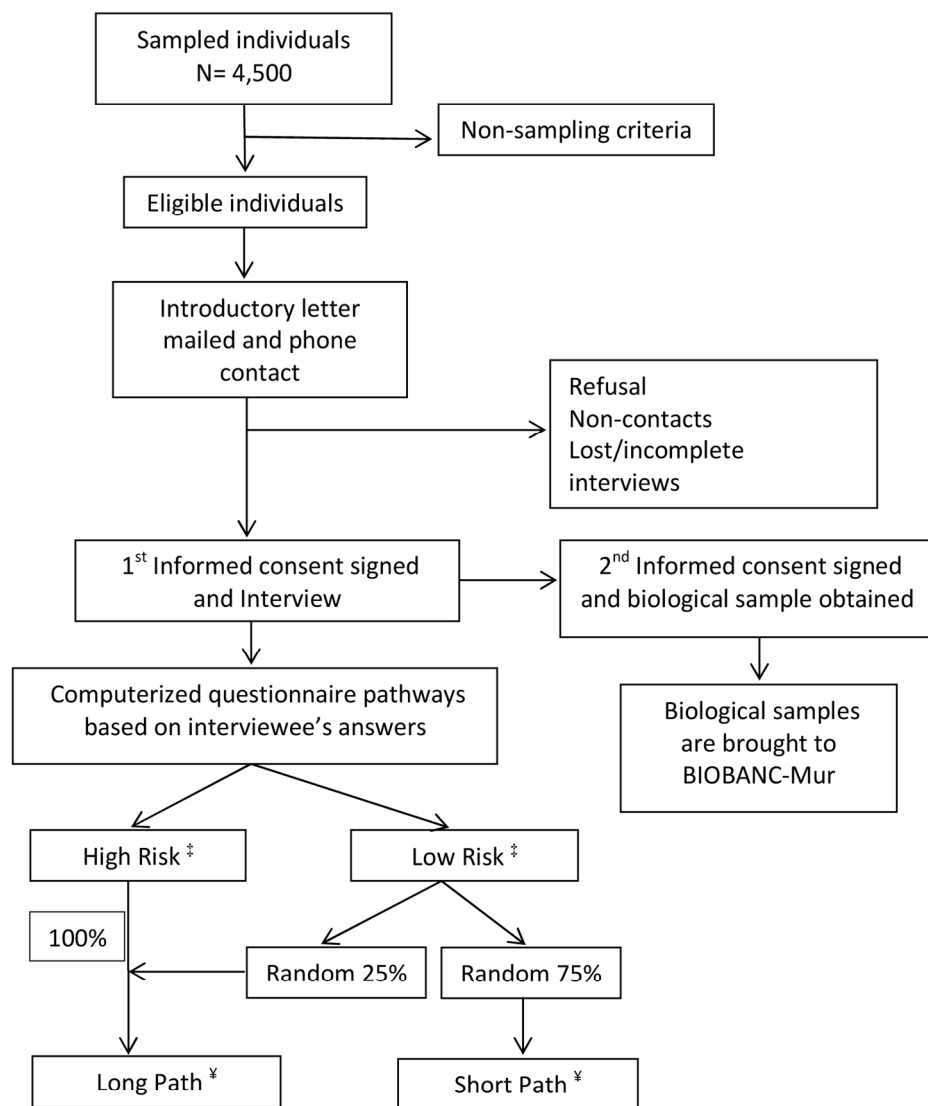
- 1  
2  
3 771 66. Gabilondo A, Rojas-Farreras S, Rodriguez A et al. Use of primary and specialized mental  
4  
5 772 health care for a major depressive episode in Spain by ESEMeD respondents. *Psychiatr*  
6  
7 773 *Serv.* 2011;62:152-161.  
8  
9 774 67. Perales J, Olaya B, Fernandez A et al. Association of childhood adversities with the first  
10  
11 775 onset of mental disorders in Spain: results from the ESEMeD project. *Soc Psychiatry*  
12  
13 776 *Psychiatr Epidemiol.* 2013;48:371-384.  
14  
15 777 68. Borrell C, Peiro R, Ramon N et al. [Socioeconomic inequalities and health plans in the  
16  
17 778 Autonomous Communities of Spain]. *Gac Sanit.* 2005;19:277-285.  
18  
19 779 69. Gonzalez B, Urbanos RM, Ortega P. [Public and private supply of health services by  
20  
21 780 autonomous communities in Spain]. *Gac Sanit.* 2004;18 Suppl 1:82-89.  
22  
23 781 70. Fierro I, Yanez JL, Alvarez FJ. [Premature death and potential years of life lost due to  
24  
25 782 alcohol consumption in Spain and the different autonomous communities in 2004]. *Aten*  
26  
27 783 *Primaria.* 2010;42:95-101.  
28  
29 784 71. Ricci-Cabello I, Ruiz-Perez I, Plazaola-Castano J et al. [Mental disease, existence of  
30  
31 785 diagnostic, use of psychotropic medication: differences by autonomous communities  
32  
33 786 under the national health survey 2006]. *Rev Esp Salud Publica.* 2010;84:29-41.  
34  
35 787 72. Kessler RC, Ustun TB. The World Mental Health (WMH) Survey Initiative Version of the  
36  
37 788 World Health Organization (WHO) Composite International Diagnostic Interview  
38  
39 789 (CIDI). *Int J Methods Psychiatr Res.* 2004;13:93-121.  
40  
41 790 73. Alonso J, Ferrer M, Romera B et al. The European Study of the Epidemiology of Mental  
42  
43 791 Disorders (ESEMeD/MHEDEA 2000) project: rationale and methods. *Int J Methods*  
44  
45 792 *Psychiatr Res.* 2002;11:55-67.  
46  
47 793 74. Harkness J, Pennell B, Villar A et al. Translation Procedures and Translation Assessment in  
48  
49 794 the World Mental Health Survey Initiative. In: Kessler RC, Bedirhan Üstün T, eds. *The*  
50  
51 795 *WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of*  
52  
53 796 *Mental Disorders.*2008; New York: Cambridge University Press. World Health  
54  
55 797 Organization.

- 1  
2  
3 798 75. Haro JM, Arbabzadeh-Bouchez S, Bugha T et al. Concordance of the Composite  
4  
5 799 International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical  
6  
7 800 assessments in the WHO World Mental Health Surveys. *Int J Methods Psychiatr Res.*  
8  
9 801 2006;15:167-180.
- 10  
11 802 76. Kessler RC, Abelson J, Demler O et al. Clinical calibration of DSM-IV diagnoses in the  
12  
13 803 World Mental Health (WMH) version of the World Health Organization (WHO)  
14  
15 804 Composite International Diagnostic Interview (WMHCIDI). *Int J Methods Psychiatr*  
16  
17 805 *Res.* 2004;13:122-139.
- 18  
19 806 77. Navarro-Mateu F, Moran-Sanchez I, Alonso J et al. Cultural adaptation of the Latin  
20  
21 807 American version of the World Health Organization Composite International  
22  
23 808 Diagnostic Interview (WHO-CIDI) (v 3.0) for use in Spain. *Gac Sanit.* 2012;27:325-  
24  
25 809 331.
- 26  
27 810 78. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading  
28  
29 811 the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12:189-198.
- 30  
31 812 79. Lobo A, Saz P, Marcos G et al. [Revalidation and standardization of the cognition mini-  
32  
33 813 exam (first Spanish version of the Mini-Mental Status Examination) in the general  
34  
35 814 geriatric population]. *Med Clin (Barc).* 1999;112:767-774.
- 36  
37 815 80. Broadbent DE, Cooper PF, FitzGerald P et al. The Cognitive Failures Questionnaire (CFQ)  
38  
39 816 and its correlates. *Br J Clin Psychol.* 1982;21 (Pt 1):1-16.
- 40  
41 817 81. García Martínez J, Sánchez Cánovas J. Adaptación al cuestionario de fallos cognitivos de  
42  
43 818 Broadbent, Cooper, Fitzgerald y Parkes (CFQ, Cognitive Failure Questionnaire).  
44  
45 819 *Análisis y Modificación de Conducta.* 1994;20:727-752.
- 46  
47 820 82. Fusté-Escolano A, Ruíz J. Estructura factorial de la versión reducida del "Eysenck  
48  
49 821 Personality Profiler". *Psicothema.* 2000;12:406-411.
- 50  
51 822 83. Sandín B, Valiente RM, Olmedo M et al. Versión española del cuestionario EPQR-  
52  
53 823 Abreviado (EPQR-A) (II): Replicación factorial, fiabilidad y validez. *Revista de*  
54  
55 824 *Psicopatología y Psicología Clínica.* 2002;7:207-216.

- 1  
2  
3 825 84. Sandín B, Valiente RM, Chorot P et al. Versión española del cuestionario EPQR-Abreviado  
4  
5 826 (EPQE-A) (I): Análisis exploratorio de la estructura factorial. *Revista de Psicopatología*  
6  
7 827 *y Psicología Clínica*. 2002;7:195-205.  
8  
9 828 85. Wagnild G, Young H. Development and psychometric evaluation of the resilience scale. *J*  
10  
11 829 *Nurs Meas*. 1993;1:165-177.  
12  
13 830 86. Heilemann M, Lee K, Kury F. Psychometric properties of the spanish version of the  
14  
15 831 Resilience Scale. *J Nurs Meas*. 2003;11:61-72.  
16  
17 832 87. Stefanis NC, Hanssen M, Smirnis NK et al. Evidence that three dimensions of psychosis  
18  
19 833 have a distribution in the general population. *Psychol Med*. 2002;32:347-358.  
20  
21 834 88. Brugha T, Bebbington P, Tennant C et al. The List of Threatening Experiences: a subset of  
22  
23 835 12 life event categories with considerable long-term contextual threat. *Psychol Med*.  
24  
25 836 1985;15:189-194.  
26  
27 837 89. Brugha TS, Cragg D. The List of Threatening Experiences: the reliability and validity of a  
28  
29 838 brief life events questionnaire. *Acta Psychiatr Scand*. 1990;82:77-81.  
30  
31 839 90. van Os J, Park S, Jones P. Neuroticism, life events and mental health: evidence for person--  
32  
33 840 environment correlation. *The British Journal of Psychiatry*. 2001;178:s72-s77.  
34  
35 841 91. Badia X, Roset M, Montserrat S et al. [The Spanish version of EuroQol: a description and  
36  
37 842 its applications. European Quality of Life scale]. *Med Clin (Barc)*. 1999;112:79-85.  
38  
39 843 92. Vilagut G, Valderas JM, Ferrer M et al. [Interpretation of SF-36 and SF-12 questionnaires in  
40  
41 844 Spain: physical and mental components]. *Med Clin (Barc)*. 2008;130:726-735.  
42  
43 845 93. Subdirección General de Salud Mental. *Guía práctica clínica. Actuación en salud mental*  
44  
45 846 *con mujeres maltratadas por su pareja*. 2010; Murcia: Servicio Murciano de Salud.  
46  
47 847 94. Beskow LM, Burke W, Merz JF et al. Informed consent for population-based research  
48  
49 848 involving genetics. *JAMA*. 2001;286:2315-2321.  
50  
51 849 95. Abitan H, Bohr H, Buchhave P. Correction to the Beer-Lambert-Bouguer law for optical  
52  
53 850 absorption. *Appl Opt*. 2008;47:5354-5357.  
54  
55 851 96. Commoner B, Lipkin D. The Application of the Beer-Lambert Law to Optically Anisotropic  
56  
57 852 Systems. *Science*. 1949;110:41-43.  
58  
59  
60

- 1  
2  
3 853 97. Trumbo TA, Schultz E, Borland MG et al. Applied spectrophotometry: Analysis of a  
4  
5 854 biochemical mixture. *Biochem Mol Biol Educ*. 2013.  
6  
7 855 98. Mulot C, Stucker I, Clavel J et al. Collection of human genomic DNA from buccal cells for  
8  
9 856 genetics studies: comparison between cytobrush, mouthwash, and treated card. *J*  
10  
11 857 *Biomed Biotechnol*. 2005;2005:291-296.  
12  
13 858 99. Rogers NL, Cole SA, Lan HC et al. New saliva DNA collection method compared to buccal  
14  
15 859 cell collection techniques for epidemiological studies. *Am J Hum Biol*. 2007;19:319-  
16  
17 860 326.  
18  
19 861 100. Viltrop T, Krjutskov K, Palta P et al. Comparison of DNA extraction methods for  
20  
21 862 multiplex polymerase chain reaction. *Anal Biochem*. 2010;398:260-262.  
22  
23 863 101. Valverde JC, Tormo MJ, Navarro C et al. Prevalence of diabetes in Murcia (Spain): a  
24  
25 864 Mediterranean area characterised by obesity. *Diabetes Res Clin Pract*. 2006;71:202-  
26  
27 865 209.  
28  
29 866 102. Huerta JM, Tormo MJ, Egea-Caparros JM et al. Accuracy of self-reported diabetes,  
30  
31 867 hypertension and hyperlipidemia in the adult Spanish population. DINO study findings.  
32  
33 868 *Rev Esp Cardiol*. 2009;62:143-152.  
34  
35 869 103. Hyman SE. The genetics of mental illness: implications for practice. *Bull World Health*  
36  
37 870 *Organ*. 2000;78:455-463.  
38  
39 871 104. Merikangas KR, Low NC, Hardy J. Commentary: understanding sources of complexity in  
40  
41 872 chronic diseases--the importance of integration of genetics and epidemiology. *Int J*  
42  
43 873 *Epidemiol*. 2006;35:590-592.  
44  
45 874 105. Weich S, Araya R. International and regional variation in the prevalence of common  
46  
47 875 mental disorders: do we need more surveys? *Br J Psychiatry*. 2004;184:289-290.  
48  
49 876 106. Kessler RC. Psychiatric epidemiology: selected recent advances and future directions. *Bull*  
50  
51 877 *World Health Organ*. 2000;78:464-474.  
52  
53 878 107. Mercer KB, Orcutt HK, Quinn JF et al. Acute and Posttraumatic Stress Symptoms in a  
54  
55 879 Prospective Gene x Environment Study of a University Campus Shooting. *Arch Gen*  
56  
57 880 *Psychiatry*. 2011;69:89-97.  
58  
59  
60

- 1  
2  
3 881 108. Tunbridge EM, Eastwood SL, Harrison PJ. Changed relative to what? Housekeeping genes  
4  
5 882 and normalization strategies in human brain gene expression studies. *Biol Psychiatry*.  
6  
7 883 2011;69:173-179.  
8  
9 884 109. Abuhatzira L, Shemer R, Razin A. MeCP2 involvement in the regulation of neuronal  
10  
11 885 alpha-tubulin production. *Hum Mol Genet*. 2009;18:1415-1423.  
12  
13 886 110. Bayou N, Belhadj A, Daoud H et al. Exploring the 7p22.1 chromosome as a candidate  
14  
15 887 region for autism. *J Biomed Biotechnol*. 2010;2010:423894.  
16  
17 888 111. Fatjo-Vilas M, Papiol S, Estrada G et al. Dysbindin-1 gene contributes differentially to  
18  
19 889 early- and adult-onset forms of functional psychosis. *Am J Med Genet B Neuropsychiatr*  
20  
21 890 *Genet*. 2011;156B:322-333.  
22  
23  
24  
25  
26  
27  
28  
29 891  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



† The response rate is defined as:  $(\text{completed interviews}) / (\text{total released respondent sample cases} - \text{respondent nonsample cases})$ .

‡ 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)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



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

	Item No	Recommendation	Pages
<b>Title and abstract</b>	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
<b>Introduction</b>			
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
<b>Methods</b>			
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) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —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 <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	8-9
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —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) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	

Continued on next page

<b>Results</b>			<b>Pages</b>
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) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	-
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time <i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure <i>Cross-sectional study</i> —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	-
<b>Discussion</b>			
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	-
<b>Other information</b>			
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](http://www.strobe-statement.org).