Annals of Epidemiology 26 (2016) 93–99

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Original article

Spatial distribution of individuals with symptoms of depression in a periurban area in Lima: an example from Peru

Paulo Ruiz-Grosso MD [a,b], J. Jaime Miranda MD, MSc, PhD [a,c,*,1], Robert H. Gilman MD [a,d,e], Blake Byron Walker BA, MSc, PhD [f], Gabriel Carrasco-Escobar BSc [a,g], Marco Varela-Gaona BEng [a], Francisco Diez-Canseco BA, MPH [a], Luis Huicho MD, DrMed [a,c,h,i], William Checkley MD, PhD [a,d,e,j], Antonio Bernabe-Ortiz MD, MPH [a,g], CRONICAS Cohort Study Group [1]

[a] CRONICAS Centre of Excellence in Chronic Diseases, Universidad Peruana Cayetano Heredia, Lima, Peru
[b] Mental Health Working Group, Universidad Peruana Cayetano Heredia, Lima, Peru
[c] Department of Medicine, School of Medicine, Universidad Peruana Cayetano Heredia, Lima, Peru
[d] International Health Department, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD
[e] Área de Investigación y Desarrollo, Asociación Benéfica PRISMA, Lima, Peru
[f] GIS and Health Informatics Laboratory, Department of Geography, Simon Fraser University, Burnaby, British Columbia, Canada
[g] School of Public Health and Administration, Universidad Peruana Cayetano Heredia, Lima, Peru
[h] Department of Pediatrics, Instituto Nacional de Salud del Niño, Lima, Peru
[i] Department of Pediatrics, School of Medicine, Universidad Nacional Mayor de San Marcos, Lima, Peru
[j] Division of Pulmonary and Critical Care, School of Medicine, Johns Hopkins University, Baltimore, MD

ARTICLE INFO

Article history:
Received 17 June 2015
Accepted 3 November 2015
Available online 14 November 2015

Keywords:
Depression
Hotspot
Spatial clustering
Peru
Mental health

ABSTRACT

Purpose: To map the geographical distribution and spatial clustering of depressive symptoms cases in an area of Lima, Peru.
Methods: Presence of depressive symptoms suggesting a major depressive episode was assessed using a short version of the Center for Epidemiologic Studies Depression Scale. Data were obtained from a census conducted in 2010. One participant per selected household (aged 18 years and above, living more than 6 months in the area) was included. Residence latitude, longitude, and elevation were captured using a GPS device. The prevalence of depressive symptoms was estimated, and relative risks (RRs) were calculated to identify areas of significantly higher and lower geographical concentrations of depressive symptoms.
Results: Data from 7946 participants, 28.3% male, mean age 39.4 (SD, 13.9) years, were analyzed. The prevalence of depressive symptoms was 17.0% (95% confidence interval = 16.2%–17.8%). Three clusters with high prevalence of depressive symptoms (primary cluster: RR = 1.82; $P = .003$ and secondary: RR = 2.83; $P = .004$ and RR = 5.92; $P = .01$), and two clusters with significantly low prevalence (primary: RR = 0.23; $P = .016$ and secondary: RR = 0; $P = .035$), were identified. Further adjustment by potential confounders confirmed the high prevalence clusters but also identified newer ones.
Conclusions: Screening strategies for depression, in combination with mapping techniques, may be useful tools to target interventions in resource-limited areas.

Introduction

Psychiatric disorders, particularly depressive disorders, contribute close to 14% of the global burden of disease [1]. Depressive disorders are among the most disabling diseases and are among the ten leading causes of the global burden of disease [2–4]. In Peru, mental health is considered to be one of the national investigation priorities [5]. Accordingly, this research focuses on establishing a geographical profile of depression symptoms with the context of mental health.

Although person, place, and time comprise the three fundamental elements of epidemiologic research [6], many previous studies have underused an understanding of place to interrogate the underlying mechanisms of disease. Geographic information systems (GISs), including mapping and geo-spatial analysis, are

There are no conflicts of interest declared.

* Corresponding author. Armendariz 497, Miraflores, Lima 12, Peru. Tel.: +51-1-241-6978.
E-mail address: Jaime.Miranda@upch.pe (J.J. Miranda).

[1] Members of the CRONICAS Cohort Study Group, listed at the end.

http://dx.doi.org/10.1016/j.annepidem.2015.11.002
1047-2797/

powerful tools used often for the investigation of infectious disease [7–9] and less frequently for the study of nontransmissible conditions [10,11]. To date, these methodologies and types of analysis are relatively new in mental health research, with a few exceptions [12–16]. Moreover, many existing reports are based on passive epidemiologic surveillance, especially in studies of schizophrenia [17–19], and these approaches have barely been used in resource-constrained settings.

Several factors have been associated with a major depressive episode such as gender, age, education level, and even migration [20,21]. Knowing how diseases and risk factors are spatially distributed can contribute to an understanding of their dynamics. Some diseases tend to cluster in areas of high or low prevalence (hot or cold spot, respectively); once identified, the next step is to assess whether clusters can be explained by known or new-risk factors [11]. Therefore, the aim of this study was to describe by mapping and characterizing the spatial distribution of cases with high probability of major depressive episode in a resource-constrained area.

Materials and methods

Study design and setting

This is a cross-sectional study using data from a census conducted in Las Pampas de San Juan de Miraflores in Lima, Peru. The census formed the sampling framework for sex-stratified and age-stratified random selection of participants for a prospective cardiopulmonary cohort [22].

Las Pampas de San Juan de Miraflores is a high-density peri-urban area located in the south of Lima, with close to 60,000 inhabitants living in 4 km^2, a great number of whom are migrants from the Andean regions. According to the 2007 National Institute of Statistics and Informatics (INEI) census, 60% of the study area population has at least one unmet basic need, 65% has no health insurance, and 25% of households are overcrowded or otherwise inappropriate for habitation [23].

Participants and selection criteria

Only one participant per household (the household's respondent for the census) was included (male or female, aged 18 years and above, and living in the area of study for a minimum of the last 6 months, consecutively). The participant provided a voluntary self-report of depressive symptoms, alongside sociodemographic information for all household members. This ensured that the collected data about depressive symptoms were recent.

Procedures

Data were collected as part of a census conducted between January and June 2010. After consent, fieldworkers completed a standardized one-page form for household data, including contact information and sociodemographic variables for all household members. The household respondent was the only person who gave personal data about depressive symptoms.

The geographic position of the household was obtained (latitude, longitude, and elevation) using a consumer GPS device (Garmin eTrex H model) with a high-sensitivity receptor and 10-meter precision. Household locations were collected under the main entrance to each household, following the instructions in the device user guide [24]. These data were then recorded on the respondent data form and periodically uploaded to a personal computer.

Outcome variable

The variable of interest, defined as the presence of depressive symptoms consistent with the diagnosis of major depressive episode, was assessed using a short version of the Center for Epidemiologic Studies Depression Scale [25]. This version comprised only two questions and has been used for population screening. It has been reported that the ability for detecting depression cases using this scale is greater when the number of questions is reduced [26]. Our short version was previously piloted, demonstrating high reliability scores (Cronbach alpha = 0.85) when compared to the full-length version of the scale. The two questions of this study were assessed regarding the previous 7 days of the participant evaluation: (1) "I felt that I could not shake off the blues even with help from my family or friends" and (2) "I felt depressed". Responses to both questions had four options: rarely or never (less), some or a little of the time (1 or 2 days), occasionally or a moderate amount of time (3 or 4 days), and most or all the time (5 to 7 days). These categories, scored from 0 to 3 for each question, were summed to produce an overall score. Participants with 3 points or more were defined as having depressive symptoms. This cutoff value was derived from the pilot result, where the sensitivity and specificity were over 90% for detecting depressive symptoms using clinical diagnosis of major depressive episode according to DSM-IV criteria as the gold standard.

Other variables

Sociodemographic data, used as potential confounders and recorded during the census, included participants' sex, age ($<$40 or $\geq$40 years), education level (primary or less, secondary, and superior/technical), and migrant status (defined according of having been born in Lima or not).

Statistical and spatial analysis

STATA 13 for Windows (Stata Corp, College Station, Texas) was used for description of the study population. To analyze the spatial distribution, we used (1) SaTScan, a statistical package for cluster detection and modeling using the spatial scan statistic, developed by Kulldorff [11] and freely available online and (2) QGIS version 2.8 (QGIS developer team, Open Source Geospatial Foundation) to map georeferenced household points.

Descriptive statistics and the prevalence of depressive symptoms were calculated with 95% confidence intervals (CIs) in STATA. Then, spatial clustering was assessed using SaTScan. The purely spatial (no temporal variable included) binomial model for the spatial scan statistic was selected to detect clustering of high and low prevalence [27]. This iteratively tests for statistically significant clustering using a circular "scanning window". The window is repeatedly reconstructed, centered on each household coordinate, and the window radius expands at each iteration from zero until the time at which 50% of study households (default) are within the circle. As such, this method is able to detect clusters of varying locations, sizes, and densities. Furthermore, it does not require the specification of cluster size and location, therefore decreasing the probability of selection bias [28].

The resulting spatial scan statistic is based on the likelihood ratio test, calculating a relative risk (RR) to test whether positive and negative cases were randomly distributed throughout the study area. If this hypothesis is rejected, then the presence of a circular cluster of high or low prevalence is inferred. The cluster's relative risk, *P* value, radius, and centroid coordinates (geometric center) are reported as model outputs. In other words, the statistic assesses whether the relative risk of cases (those with depressive