PEER REVIEW HISTORY

This paper was submitted to a another journal from BMJ but declined for publication following peer review. The authors addressed the reviewers' comments and submitted the revised paper to BMJ Open. The paper was subsequently accepted for publication at BMJ Open.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Inequity and Vulnerability in Latin American Indigenous and Non-Indigenous Populations with Rheumatic Diseases :A Syndemic Approach
AUTHORS	Granados, Ysabel; Gastelum Strozzi, Alfonso; Alvarez-Nemegyei, Jose; Quintana, Rosana; Julian-Santiago, Flor; Santos, Ana; Guevara-Pacheco, Sergio; Loyola-Sanchez, Adalberto; Goycochea-Robles, Maria Victoria; Juarez, Vicente; Garza-Elizondo, Mario Alberto; Rueda, Juan; Burgos-Vargas, Ruben; Londoño, John; Pons-Estel, Bernardo; Pelaez-Ballestas, Ingris

VERSION 1 – REVIEW

REVIEWER	Nieto, Romina CREAR , Rosario
REVIEW RETURNED	29-Nov-2022

GENERAL COMMENTS	Congratulations on the work. The questions have been answered and the reviewers' comments have been modified. I consider that the manuscript is ready for publication.

REVIEWER	Crowson, Cynthia Mayo Clinic, Health Sciences Research
REVIEW RETURNED	17-Jan-2023

GENERAL COMMENTS	The revised text for the statistical methods is unclear, particularly for the logistic regression. It seems some numbering from the response to the reviewers was not included in the text, which makes it very difficult to interpret. The paper is difficult to read, and the network analysis is not clearly presented. I do not believe readers will find this paper clear and easy to understand. I believe the previous reviewers comments were very helpful and the paper is now better than it was originally, but it is still not clear. Regarding the network analysis, more information is required about how the total number of clusters is chosen, and about how the authors chose the clusters they decided to highlight. Also, more information is needed to interpret Figures 1 and 2. What are the x and y axes of these figures? what does the size of the dot represent?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Dr. Romina Nieto, CREAR , Hospital Escuela Eva Peron

Comments to the Author:

Congratulations on the work. The questions have been answered, and the reviewers' comments have been modified. I consider that the manuscript is ready for publication.

R: Thank you for your kind comments.

Reviewer: 2

Dr. Cynthia Crowson, Mayo Clinic

Comments to the Author:

1. The revised text for the statistical methods is unclear, particularly for the logistic regression. It seems some numbering from the response to the reviewers was not included in the text, which makes it very difficult to interpret.

> **R:** Thanks for your comments. We included all changes in the previous version (highlighted in yellow).

2. The paper is difficult to read, and the network analysis is not presented. I do not believe readers will find this paper clear and easy to understand. I believe the previous reviewers' comments were very helpful and the paper is now better than it was originally, but it is still not clear.

R: We agree. We followed your recommendation analysis section was re-rewriting

Analysis

A multi-phase analysis was performed.

Phase 1. We applied inferential statistics (i.e., bivariate analysis) to explore associations between ethnicity (Indigenous or non-Indigenous) and country of origin, and sociodemographic characteristics (i.e., age, gender, formal schooling, urban/rural residence), rheumatic diagnoses, comorbidities, disability (HAQ-DI) and levels of accessibility (partial, total and private coverage).

Phase 2. We performed simple logistic regression models to identify factors (i.e., sociodemographic, comorbidities, disability, accessibility, and joint biomechanical stress) associated with RMD diagnosis (i.e., present or absent) as a dependent variable by ethnicity (Indigenous vs. non-Indigenous). We estimated odds ratios (OR), along with 95% confidence intervals (CI 95%) and significance (p).

Phase 3. We used a network analysis approach (35) to generate groups with similar characteristics (e.g., sociodemographic, country, comorbidities, diagnoses, etc.) called clusters. These clusters helped to determine the negative characteristics associated with disease and disability using the syndemic framework. The network analysis method requires the definition of a characteristic that allows creation of connections between subjects; a measure of similarity was defined to create these. The similarity measure determined the relationships between the different subjects within the database. The measure of similarity evaluated the number of similarities between two subjects regarding the results of their evaluations. To construct the first part of the similarity measure, a vector was defined with the following variables: a) accessibility level, b) level of joint biomechanical stress, and c) urban/rural residence. Using the cosine similarity method, this vector was used to calculate a similarity index for each individual concerning the rest of the population (36). The final similarity index was obtained by applying a weighted difference by years of education between each individual.

> The similarity index was used to determine an individual's degree of similarity to the rest of the population and to build the relations between individuals. In the network definition, each individual is a node; an axis of relations is generated when the similarity index between two individuals is greater than the average of the similarity indices plus the standard deviation of the whole population (37). The network obtained is simulated in Gephi (38), and the final position of the nodes or individuals is used to define the new groups using the DBSCAN method (39).
>
> Due to the complexity of the representation of the clusters, we conducted a consensus process among all researchers to select the most relevant clusters regarding socio-economic and clinical impact, which included healthcare access, disability, educational level, and type of RMD. Selected clusters were further analyzed in network analysis, including the following factors in a hierarchical order of importance: 1) prevalence of RMD, 2) prevalence of rheumatoid arthritis (RA), and 3) the number of individuals comprising the cluster. All researchers assigned every cluster a weighted score for each of the three selected variables. Finally, six clusters were selected per group (i.e., Indigenous and non-Indigenous) according to their amount of representation of health inequity factors.
>
> Phase 4. We conducted a sensitivity analysis to confirm no biases using a randomly selected weighted sample of Indigenous/non-Indigenous populations from the three countries that studied both at the same time(Ecuador, Mexico, and Venezuela), and two countries that only had samples of Indigenous (Argentina) or non-Indigenous (Colombia) populations. The clusters obtained through this analysis were defined by factors such as living in a rural setting, lower health coverage, and greater disability, which went beyond our initial Indigenous/non-Indigenous classification and impacted the management of rheumatic diseases. These emerging differences can be used to document inequity insofar as they highlight the variables which negatively affect the health of people with RMD. **(Lines 239-286. page 13-14).**

3. Regarding the network analysis, more information is required about the number of clusters chosen and how the authors chose the clusters they decided to highlight. Also, more information is needed to

interpret Figures 1 and 2. What are the x and y axes of these figures? what does the size of the dot represent?

> **R:** Thank you for your suggestion. We followed your re-writing recommendation analysis section (Lines 239-289. pages 12-15). Besides, we added information in Figures 1 and 2.
>
>> **Circle size represents the number of individuals per cluster for visual comparison. The cluster positions are the result of the network simulation; the position of each cluster is obtained during the simulation depending on the similarity of the individuals.