ABSTRACT

Health differences across socioeconomic strata have always pointed out that poorer and minorities have higher mortality and morbidity than richer and majorities. This difference is exacerbated for particular populations such as the victims of ongoing armed conflict, who are also much harder to quantify due to the conflict itself. This study uses complex network analysis applied to a combination of three large administrative records for the health system and mortality records in the province of Risaralda (Colombia) between 2011 and 2016. We estimate the most common causes of morbi-mortality for both victims of violence and the poorest inhabitants of Risaralda, defined as those who qualify as recipients of subsidies from the Colombian welfare program, called SISBEN, in the categories of those with the highest need, levels I and II. Both populations show high morbidity frequencies for non-communicable diseases such as Type II diabetes, hypertension and hyperglyceridaemia, mostly associated with exposure to unhealthy lifestyles. However, these
mortality outcomes reflect the different treatments and lifestyles of both subpopulations. While the poorest replicate the same causes identified for morbidity, the victims of armed conflict die of additional causes including Type I diabetes, which reflects the even worse conditions they face.

1. Data and Methods

As mentioned above, we linked the records from three independent administrative databases for the province of Risaralda, for all the years between 2011 and 2016, by using a unique identification number. Using a sole year is a mistake for this study, as a particular year may record a mortality shock or bust for exogenous reasons. Hence, we use these six years pooled together. The first dataset is *Base de Datos Única de Afiliados* (BDUA, Unique Database of Affiliated [to the health system]), which contains individual records affiliated to the health system in Colombia. Second, *Registro Individuales de Prestación de Servicios de Salud* (RIPS, Individual Records of Health Services) holds the full medical records of all patients who are part of BDUA and who attended a regular medical visit, urgency care, and the follow-up of medical exams and treatments. The information holds symptoms, diagnosis and prognosis, as recorded by medical doctors. This information is essential for the morbidity profiles.

Finally, *Registro Unificado de Afiliación* (RUAF, Unified Record of Affiliation) holds vital records, including birth and death certificates that aim to be universal. One of the advantages of the vital registration system in Colombia is that death records include the main cause of death as well as the underlying causes of death, recorded under the International Diseases Classification in its 10th version (ICD-X).
2. Morbidity Network Analysis

We begin with the full information of the complex network, namely all diagnoses in the Risaralda province, for both populations under study. The population who qualify for subsidy programs in Colombia are ranked in the Identification System of Potential Beneficiaries for Welfare Programs (*Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales, SISBEN*) classified at levels I and II. To define this network’s nodes, we use both main and secondary medical diagnoses for each individual in RIPS. The edges are the relationships between the main and secondary diagnoses for each individual. As these relationships have a very high frequency, regardless of whether it is a main or a secondary diagnosis, the edges do not have any direction resulting in a building scheme, as shown in Figure 1.

![Figure 1. Structure for Building the Morbidity Network](image)

For Figure 1, D_1 is the main diagnosis, while D_2 and D_3 are the secondary diagnoses, all following the ICD-X, as stated before. The strength of links $W_{ij}$ is defined as

$$W_{ij} = \sum_{i,j} D_i \cdot D_j, \text{ with } D_i \neq D_j \quad (1)$$
where \( n \) is the total number of nodes in the network, namely the number of different diagnoses recorded in RIPS (\( D_i \neq D_j \)); \( D_i \cdot D_j \) is the relation between each record for primary diagnoses \( i \) and secondary diagnoses \( j \). The total complex network from applying the above-described methodology to the data at hand, for all the years of study, results in Figure 2. The principal metric of nodes with the highest connectivity is in Table 1.

Figure 2. Graph of the Morbidity Network of Diagnoses in Risaralda, 2011–2016.
*Nodes are diagnoses and links are medical records from all kinds of medical visits: regular, emergency consultations, treatments and follow-ups.

3. Selection Algorithm for Detecting the Subpopulation

For this study, we need to fully identify both subpopulations of interest, namely victims of internal conflict and beneficiaries of subsidy programs, SISBEN I and II. To do so, we chose the k-communities algorithm of Fortunato (2010) and Palla et al. (2005), because this keeps the superposition of diagnoses in
the subgraphs even though many individuals share the same initial diagnosis, but with different final diagnoses.

**Figure 3. Detection Algorithm for the Detection of k-Communities**

Source: Fortunato (2010) and Palla et al. (2005)

4. **Intensity Analysis and Motif Coherence.**

Motifs are interconnected patterns in complex networks with a much larger frequency than random networks (Milo, 2002). They are common in biology (Green et al., 2017; Smoly et al., 2017) and ecology (Delmas et al., 2017; Rodríguez-Rodríguez et al., 2017), among other applications. Motifs have intrinsic characteristics that condition the probability of the occurrence of certain values in nodes, despite their application to particular cuts of the network (Milo, 2002). This permits them to generate a series of trends in the network circumvent information, such as nodes’ consensus that control their
flow. This characteristic is essential to associate diagnoses and illness related to lifestyles in the province of Risaralda.

Intensity, $I(g)$, for subgraph $g$ with vertices $V_g$ and edges $l_g$, as the geometric mean of their weights, or strengths, $W_{ij}$, is described in equation (1):

$$I(g) = \left( \prod_{(ij) \in E_g} W_{ij} \right)^{\frac{1}{|E_g|}}$$

The coherence $Q(g)$, which allows us to study the consensus between people at the edges inside motifs. Coherence takes values near to the most important unit in its subgraph to establish the association between subpopulations, and it is defined as the ratio between intensity, $I(g)$, and the geometric mean of their weights, or strengths, $W_{ij}$, as presented in equation (2):

$$Q(g) = \frac{I_g}{\Sigma_{(ij) \in E_g} W_{ij}}$$

**Results**

**Results for the Victims of Internal Armed Conflict**

![Image](image.png)

*Figure 4. Community with the Highest Coherence Level. Morbidity Network of the Victims of Conflict in Risaralda, 2011–2016*
Figure 5. Motifs’ Coherence in Cliques (left) and Victims of Conflict (right) for the Morbidity Network in Risaralda, 2011–2016

Figure 6. Clique with the Highest Overlapping in Morbidity and Mortality. Mortality Network of the Victims of Conflict in Risaralda, 2011–2016
Results for the Population Classified as SISBEN I and II

Figure 7. Highest Overlapping in Morbidity and Mortality. SISBEN I and II Mortality Network in Risaralda, 2011–2016

Figure 8. Motifs’ Coherence in Cliques (left) and SISBEN I and II (right) for the Morbidity Network in Risaralda, 2011–2016
REFERENCES


• Reconciliacion Colombia. 2015. Los desplazados son más pobres que el resto de la población, 16 February 2015, available at: http://goo.gl/eK5F38
• United Nations High Commissioner for Refugees-UNHCR. 2015. UNHCR country operations profile – Colombia, available at: http://goo.gl/2a7r1S, downloaded on 26 March 2015
• Urdinola, B. Piedad, Torres Avilés, Francisco, & Velasco, Jairo Alexander. (2017). The Homicide Atlas in Colombia: Contagion and Under-Registration for