

1 *Type of the Paper: Article*

## 2 **Morbi-Mortality of the Victims of Internal Conflict** 3 **and Poor Population in The Risaralda Province,** 4 **Colombia**

5 **Rafael Rentería-Ramos**<sup>1\*</sup>, **Rafael Hurtado-Heredia**<sup>2</sup> and **B Piedad Urdinola**<sup>3</sup>.

6 <sup>1</sup> Postdoctoral Fellow. Departments of Physics and Statistics - Universidad Nacional de Colombia-Bogotá;

7 [rrenteriar@unal.edu.co](mailto:rrenteriar@unal.edu.co)

8 <sup>2</sup> Associate Professor. Department of Physics - Universidad Nacional de Colombia-Bogotá. Email:

9 [rghurtadoh@unal.edu.co](mailto:rghurtadoh@unal.edu.co)

10 <sup>3</sup> Associate Professor. Department of Statistics - Universidad Nacional de Colombia-Bogotá. Email:

11 [bpurdinolac@unal.edu.co](mailto:bpurdinolac@unal.edu.co)

12 \*Correspondence: [rrenteriar@unal.edu.co](mailto:rrenteriar@unal.edu.co); Tel.: +57-315-647-05-45

13

14 **Abstract:** Health differences across socioeconomic strata have always pointed out that poorer and  
15 minorities have higher mortality and morbidity than richer and majorities. This difference is  
16 exacerbated for particular populations such as the victims of ongoing armed conflicts, who are also  
17 much harder to quantify due to the conflict itself. This study applies network analysis to a  
18 combination of three large administrative records for the health system and mortality records in the  
19 province of Risaralda (Colombia) between 2011 and 2016. It produces the most common causes of  
20 morbi-mortality for both victims of violence and the poorest inhabitants of Risaralda, defined as  
21 those who qualify as recipients of subsidies from the Colombian welfare program called SISBEN in  
22 the categories of highest need. Both populations show high morbidity frequencies for non-  
23 communicable diseases such as Type II diabetes, hypertension, and hyperglyceridaemia, mostly  
24 associated with exposure to unhealthy lifestyles. Additionally, the mortality outcomes reflect the  
25 different lifestyles and medical treatments of both subpopulations. While the poorest replicate the  
26 same causes identified for morbidity, the victims of armed conflict die of additional causes including  
27 Type II diabetes, which reflects the even worse conditions they face.

28 **Keywords:** Morbidity; Mortality; Network Analysis; Victims of internal conflict.

29

### 30 **1. Introduction**

31 Refugees and internally displaced populations (IDPs) have much higher morbidity and mortality  
32 rates than the rest of the population for all observed conflicts in the world, as recorded from their  
33 experiences in refugee and IDP settlements [1]. These results are not surprising, as both populations  
34 are exposed to harsh physical, emotional, and economic conditions due to violence that has both  
35 direct effects on mortality [2-4] as well as indirect effects, as measured by higher infant mortality rates  
36 [5-6]. Indeed, the disruption that conflicts impose leads to extreme inequality in socioeconomic  
37 conditions between victims and the rest of the population, which in turn may be the underlying cause  
38 of those large differentials in morbidity and mortality, as this affects healthy lifestyles and all the  
39 other determinants of health [7].

40 This study aims to show the differences between the morbidity and mortality of Colombian victims  
41 of internal conflict and those of a comparable subpopulation, namely an economically depressed  
42 population targeted by government subsidy programs in Colombia. We propose that the living

43 standards of victims of internal conflict are so extreme that they lead to worse physical and  
44 psychological conditions than those of the poorest population. For that purpose, this study uses micro  
45 data from the province of Risaralda, one of the highest ranked provinces in Colombia in terms of  
46 socioeconomic conditions. Risaralda has almost universal coverage of public services (roads, water,  
47 electricity, gas, and sewage), and health services access, but it is in the midst of regions that have  
48 suffered persistent and intense armed conflict, attracting victims fleeing from high violence levels  
49 zones, as Chocó, Urabá, and the southwest provinces of Colombia. From all victims registered in the  
50 official Colombian records, 80% of them are IDPs. Hence, this study presents the results for all victims  
51 in Risaralda, knowing that in this province the majority are IDPs [8].

52  
53 This study adds to the increasing amount of literature on the application of network analysis and the  
54 relation between health and lifestyle disruption in three ways. First, we portray the different patterns  
55 between victims of violence and other subpopulations in an extreme case of endemic conflict, such  
56 as the Colombian one, where most victims of conflict are not living in settlements or safe havens, in  
57 contrast to other more intense but shorter conflicts in the world. Second, we take advantage of the  
58 new possibility of crossing individual administrative records that allowed us to combine information  
59 on access to the health system, individual medical records, vital statistics and the status of victims of  
60 the internal conflict in Colombia, a hard-to-reach subpopulation that is difficult or impossible to track  
61 with traditional surveys. We successfully crossed three official databases, all linked by the  
62 identification number of individuals, between 2011 and 2016: Base de Datos Única de Afiliados  
63 (BDUA, Unique Database of Affiliated [to the health system]), Registros Individuales de Prestación  
64 de Servicios de Salud (RIPS, Individual Records of Health Services) and Registro Unificado de  
65 Afiliación (RUAF, Unified Record of Affiliation). Yet, as with all big data, this novel dataset comes  
66 with its challenges for which the proposed methodology is the best solution we came up with.  
67 Network analyses allowed us to identify the patterns of morbidity and mortality for both victims of  
68 internal conflict and the proposed control group (i.e., target populations for subsidy programs). The  
69 differences among these subpopulations lead us to identify different causes of death and  
70 predominant diseases, confirming the hypothesis of this study. Third, an abrupt disruption in  
71 lifestyles leads to different observed health outcomes, and an example is the victims of internal  
72 conflict in Colombia.

73  
74 The data and methods used herein include a detailed characterization of diseases, for both mortality  
75 and morbidity, of subpopulations living under the extreme conditions due to displacement. To put  
76 these results into context, we compare them with those obtained for other poor populations, namely  
77 those that participate in subsidized governmental welfare programs.

78  
79 We are aware that official records, as the count of victims of the internal conflict, have several  
80 limitations. To begin with, the individuals registered in this dataset correspond to those that self-  
81 identify as victims of internal conflict and are looking for governmental support programs of any  
82 kind, which imposes a bias from the total population of victims. However, the results presented here  
83 could be considered as a lower bound of a larger issue. That is, we encourage readers to think of the  
84 population included here as the most vulnerable of all victims to the extent that despite putting their

85 anonymity in danger, they still look for social welfare help from the State in order to survive. The  
86 total number of victims is larger; yet, there may also be larger heterogeneity in their socioeconomic  
87 conditions, even though they are below the national average [9-10].  
88 The remainder of this paper is organized as follows. The next section includes the motivation and  
89 literature review. Section 3 describes the data and methods used, the following section presents the  
90 results and the final section contains the discussion and conclusions.

91

### 92 1.1. Motivation

93 Armed conflicts have several negative effects on human health and development, as recent literature  
94 has presented (for a review, see [11]). Among all, the populations affected by these conflicts, IDPs  
95 and refugees, report the worst morbidity and mortality conditions. Byarugaba in [12] presents low  
96 health results, both physical and psychological, for children and their mothers in IDP settlements in  
97 South Africa, driven by having access to limited health services. For instance, most of the causes of  
98 death for children under five years have preventable causes, including nutritional deficiencies,  
99 gastroenteritis, and dehydration. The World Health Organization [13] shows three examples of  
100 increased infant mortality during conflicts due to communicable diseases: (1) measles, tetanus, and  
101 diphtheria became epidemics in Uganda during the mid-1980s, producing infant mortality rates  
102 twice as large as the total national; (2) in Zepa, during the Bosnia and Herzegovina conflict, perinatal  
103 and childhood mortality rates doubled during the conflict; and (3) in Sarajevo, average birthweight  
104 fell by 20% in 1993 as a consequence of the doubling rates of premature births.

105 It is undeniable that the disruption created by violent conflicts affects socioeconomic conditions,  
106 which in turn shift lifestyles for all people, and it translates into worse health and mortality  
107 conditions. However, in a long and persistent conflict, such as the Colombian case, some skepticism  
108 remains about the lower life quality of the victims of the internal conflict. On the one hand, the conflict  
109 has persisted for over five decades, with a shift in the purposes of parties involved. On the other  
110 hand, the center of the conflict was itinerant around the country, which prevented enclaves of IDPs  
111 or safe havens of victims, as has occurred in other similar situations, thereby scattering victims across  
112 the country.

113

114 The Colombian conflict erupted in the 1940s with the confrontation of followers from the two main  
115 political parties, Conservatives and Liberals, in an era named “La Violencia” (The Violence), which  
116 systematically turned against unarmed civilians. By the 1960s, this conflict had evolved into a social  
117 class struggle with the eruption of communist guerrillas, as happened in many Latin American  
118 countries. By the 1980s, paramilitary groups arose as a force against leftist guerillas, and by the 1990s,  
119 armed groups found profitable violent ways for political and economic purposes, including the  
120 systematical displacement of populations to steal the land, drug production and trafficking, weapons  
121 trafficking, kidnapping, and extortion. Indeed, the most violent decade in Colombian history during  
122 the past century was the 1990s [10]. By the turn of the century, the president in place had counteracted  
123 a full military attack by guerilla groups and led a demobilization process of paramilitary groups.  
124 Afterwards, the next administration signed in November 2016 a peace agreement between the  
125 Colombian government and the largest guerrilla group, FARC (Fuerzas Armadas Revolucionarias de

126 Colombia), currently there is an ongoing negotiation process between the government and the next  
127 largest and eldest guerrilla group, ELN (Ejército de Liberación Nacional). The agreement began to be  
128 implemented in 2017, with the hope of ending the armed internal conflict [6].

129

130 Colombia has the second largest IDP in the world, with over six million people; many of them are  
131 also forced to dispossession [13]. As a result, when IDPs arrive at their destination, they live in  
132 poverty despite the social welfare programs implemented. Indeed, in Colombia for the past two  
133 decades they have had much lower socioeconomic, housing and employment conditions than before  
134 the displacement [10,13,14,9] reflected in lower sanitary conditions, lower access to health facilities  
135 and services and in turn worse health. However, no current studies have linked their health outcomes  
136 with their lifestyles, as we propose in this study. More importantly, IDPs and refugees may become  
137 a group that delays the epidemiological transition and even a focus of diminished communicable  
138 diseases that could spread to the rest of the population, as shown in the Jordanian case [16], mostly  
139 due to the lack of access to public health policies, leading to good habits, lower education, and bad  
140 nutrition [17].

141

142 Moreover, the measurement of victims of ongoing conflict is challenging. Brunborg and Urdal [3]  
143 explain the measurement complications of any conflict-related variable, and indeed the 2005  
144 Colombian Population Census shows this particular issue with its measurement of IDPs. Under the  
145 question “reason for migrating” asked to recent migrants, for those who moved within the previous  
146 five years the Census added the category “due to violence.”. The results could not be more  
147 disappointing: they produced the lowest number of IDPs among all available sources at the time,  
148 which included official registers of IDPs, data from NGOs and data from the Catholic Church. The  
149 reason is simple: in an ongoing conflict, any direct question will lead to under-registration because  
150 victims of conflict feel the need for anonymity for their own protection.

151

152 The empirical results in other contexts support the hypothesis proposed herein. Audet [18] shows the  
153 negative effects of habits or lifestyle on vulnerable populations, particularly on women, linked to  
154 obesity and chronic diseases. Brown [19] used the complex network analysis perspective to associate  
155 socioeconomic and environmental variables to nutritional lifestyle in Los Angeles, they studied the  
156 restriction imposed on fast food and its implications on illness related to obesity, finding that  
157 unhealthy cheap food and poverty are structurally related.

## 158 2. Materials and Methods

159 We linked the records for the province of Risaralda from three independent administrative databases,  
160 for all the years between 2011 and 2016, by using a unique identification number. Since data for a sole  
161 year may record a mortality shock or bust for exogenous reasons, then we use the six years pooled  
162 together.

163

164 The first dataset is *Base de Datos Única de Afiliados* (BDUA, Unique Database of Affiliated [to the health  
165 system]), which contains individual records of the affiliated users to the health system in Colombia.  
166 There are two possible ways of being part of this dataset: making payments to the system

167 (contributive regime) or being subsidized by the government (subsidize regime). Altogether, 96.6%  
168 of the population in Risaralda belong to either category. Both the victims of internal armed conflict  
169 and populations targeted for health subsidies belong to the subsidize regime; 52.2% of the population  
170 is in the contributive regime and 44.4% in the subsidize regime. Targeted populations include those  
171 who are unemployed, with no labor income and with no direct family member who could pay any  
172 contribution to the health system for them. Victims of internal armed conflict are recognized in BDUA  
173 as a distinct group number nine (9), for which the government has an obligation to provide health  
174 services as well as other health-related secondary goods. BDUA has relevant information on  
175 demographics such as age, sex and ethnicity.

176

177 The second dataset, *Registros Individuales de Prestación de Servicios de Salud* (RIPS, Individual Records  
178 of Health Services) holds the full medical records of all patients who are part of BDUA and who  
179 attended a regular medical visit, urgency care, and the follow-up of medical exams and treatments.  
180 The information holds symptoms, diagnosis and prognosis, as recorded by medical doctors. This  
181 information is essential for the morbidity profiles.

182

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

188 The dataset built from the BDUA, RIPS and RUAF databases contains 970.000 people (96.7% of the  
189 inhabitants of the province of Risaralda), with 48'422.050 diagnoses confirmed in 242'211.050  
190 procedures (used as a metric to eliminate disease suspicions) and 23.648 mortality records for the  
191 specific cause of the diagnoses. This information was obtained from 9'684.410 medical visits (general,  
192 emergency, and follow-up) of the inhabitants of Risaralda who used the medical system the years  
193 between 2011 and 2016.

194 This study uses the complex network perspective following the definitions in [20-23]. A network  
195 represents a system through its constituents and the relations between them. The network can be  
196 mapped into a graph  $G(V, E)$ , with  $V$  the set of vertices  $v \in V$ , and  $E$  the set of edges or links between  
197 vertices. This methodology has been widely used to investigate natural, social and artificial systems,  
198 as for example vector spread diseases [24-26], the link between mobility and diseases spreading  
199 [24,27], the spread of epidemics [28], and lifestyles and health [25]. In particular, studies linking  
200 mobility and epidemics are relevant for the aim of this study. Victims fleeing an armed conflict,  
201 mostly IDPs, may introduce diseases to populations in the arrival places; however, if the conflict had  
202 not existed, then this would never have occurred [16-17]. Further, their low living standards and their  
203 associated lifestyles, such as low nutritional intake, limited access to healthcare and services,  
204 exclusion from the public health system, and low sanitary conditions, lead to a particular morbidity  
205 profile [29-30].

206

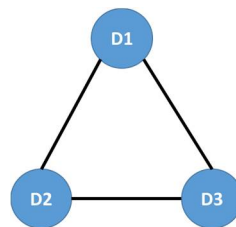
207

208 2.1 Morbidity Network of Diagnoses

209

210 We begin with the aggregate information of the complex network, namely all diagnoses in the  
 211 province of Risaralda, for both populations under study. The population who qualify for subsidy  
 212 programs in Colombia are ranked in the Identification System of Potential Beneficiaries for Welfare  
 213 Programs (*Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales, SISBEN*), classified  
 214 at levels I and II. We define the diagnoses (ICD-X) as the nodes of the network. The edges are defined  
 215 as the coupling of two diagnoses given by their co-occurrence in the records of a medical events, for  
 216 every individual in the RIPS, including the main and up to three secondary medical diagnoses, in all  
 217 kinds of medical events: regular, emergency consultations, treatments and follow-ups, as shown in  
 218 Figure 1.

219



220

221 **Figure 1.** Structure of the Morbidity Network.

222 Nodes or vertexes represent diagnoses (International Diseases Classification in its 10th version, ICD-X),  $D_1$  is the  
 223 main diagnosis, while  $D_2$  and  $D_3$  are secondary ones. Edges or links represent the co-occurrences of primary and  
 224 secondary diagnoses for every individual in any medical event: regular, emergency consultations, treatments  
 225 and follow-ups.

226 Source: Authors' scheme.

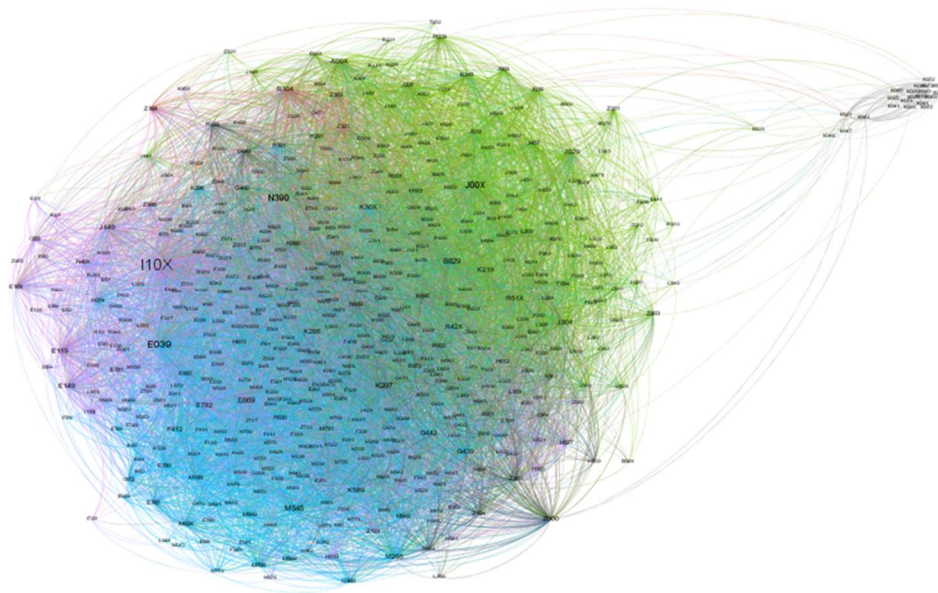
227

228 In Figure 1,  $D_1$  is the main diagnosis,  $D_2$  and  $D_3$  are the secondary diagnoses. The weight of links  $W_{ij}$   
 229 is defined as:

$$230 \quad W_{ij} = \sum_k D_{ik} \cdot D_{jk}, \quad (1)$$

231 where,  $D_{ik} = 1$  if diagnosis  $D_i$  was recorded at RIPS in the medical event  $k$ , and  $D_{ik} = 0$  otherwise.

232 The graph of the network obtained by applying the methodology described above is shown in  
 233 Figure 2.



234

235 **Figure 2.** Graph of the Morbidity Network of Diagnoses for all medical records of residents in the province of  
 236 Risaralda, years 2011 to 2016.

237 Nodes are diagnoses (International Diseases Classification in its 10th version, ICD-X) and links correspond to  
 238 pairs of diagnoses present in medical records from all kinds of medical events: regular, emergency consultations,  
 239 treatments and follow-ups. Source: Authors' own calculations.

240

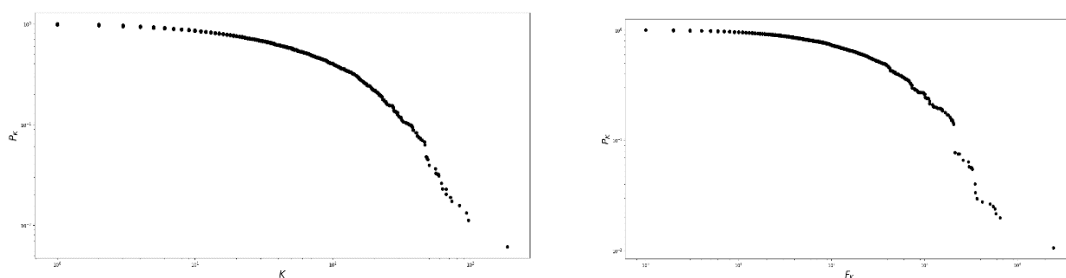
241 The degree of node  $i$ ,  $K_i$ , is the number of adjacent nodes to  $i$ ; the strength of nodes  $i$  is defined  
 242 as  $F_i = \sum_j w_{ij}$ . The diagnosis I10X, which is essential (primary) hypertension, has the largest degree  
 243 and strength, 1259 connections for 9918 medical events, followed by urinary tract infection (N390),  
 244 acute nasopharyngitis [common cold] (J00X), gastritis unspecified (K297), hypothyroidism  
 245 unspecified (E039), intestinal parasitism unspecified (B829) and low back pain (M545).

246 The degree and strength probability distributions of the network,  $p(k)$  and  $p(f)$  respectively, can be  
 247 represented using the complementary cumulative distribution function, defined in equation (2),  
 248 which produces Figure 3:

249

$$P_K = \int_k^{\infty} p(k') dk' ; F_K = \int_k^{\infty} p(f') df' \quad (2)$$

250



251

252 **Figure 3.** Nodal degree and nodal strength for the population resident in the province of Risaralda.  
253 Complementary cumulative distribution functions for the nodal degree probability distribution (left) the  
254 nodal strength (right) for the Morbidity Network of Diagnoses for the total population resident in the province  
255 of Risaralda. Years from 2011 to 2016.

256 Source: Authors' own calculations.

257  
258 Figure 3 shows the existence of three regions. First, for  $P_K$  in the interval  $[10^0, 10^{-1}]$ , several nodes  
259 have few connections and small weight. Second, between  $10^{-1}$  and  $10^{-2}$  the links strengthen. Third,  
260 for  $P_K < 10^{-2}$ , in the fat tails few nodes are highly connected, neither in degree nor in strength.

261  
262 We cut the network by selecting nodes with strength  $F_k > 10^1$ . This allows us to reduce the size of  
263 the network as well as to control spurious data from erroneous diagnoses.

264  
265  
266  
267 The network reduction provides a better understanding of the morbidity patterns in Risaralda for  
268 both subpopulations of interest. In order to detect these two subpopulations, we use subgroup  
269 cohesion, specifically cliques or sets of maximally complete subgraphs, as described in the following  
270 subsection.

271  
272 *2.2 Selection algorithm for detecting communities*

273  
274 For this study, we need to fully identify morbidity patterns with both subpopulations of interest,  
275 namely victims of internal conflict and beneficiaries of subsidy programs, SISBEN I and II. To do so,  
276 we chose the k-communities algorithm of [31-32], because it keeps the superposition of diagnoses in  
277 the subgraphs even though many individuals share the same initial diagnosis, but with different final  
278 diagnoses.

279  
280 For this algorithm, we must first quantify the total number of cliques in the network following [33]  
281 (see Figure 5, panel 1). In this case, the maximum number of diagnoses per individual medical event  
282 is  $k = 3$ , and this becomes the value  $k$  for percolation. The second step is to find the adjacent cliques,  
283 it is those that share  $k-1$  nodes. Again, we followed [32] and [31] (see Figure 4, panels 2 and 3). The  
284 final step is to join all cliques reached by the series of adjacent cliques to conform the k-community  
285 (see Figure 4, panel 4). Figure 4 shows each step of this algorithm, which allows us to distinguish  
286 between the two subpopulations of interest. For consistency purposes, it is necessary to analyze each  
287 subpopulation's subgraphs, for which we use motif analysis.



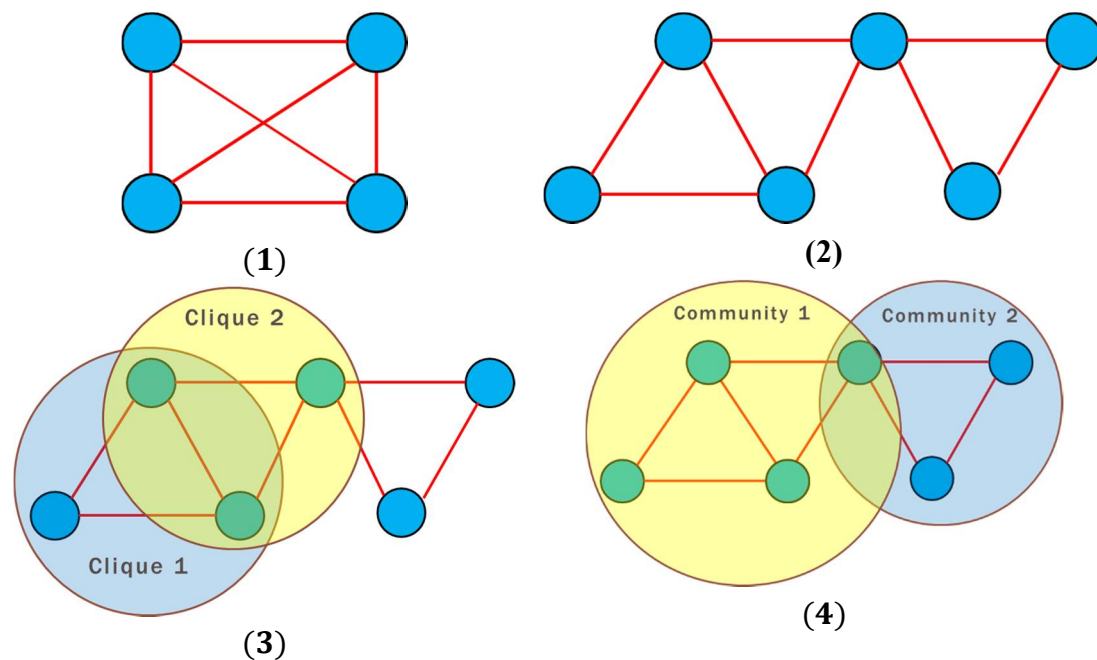


Figure 4. Detection algorithm for k-communities.

288

289 Source: [31-32].

290

### 291 2.3 Intensity Analysis and Motif Coherence

292 In complex networks, motifs are interconnected patterns with a much larger frequency than random  
 293 graphs [33]. They are common in systems studied in biology [34-35] (Green et al., 2017; Smoly et al.,  
 294 2017), ecology [36-37], among other. Motifs have intrinsic characteristics that condition the  
 295 probability of occurrence of certain values in nodes, despite their application to particular cuts of the  
 296 network [33] This allows to generate a series of trends in the network circumvent information, such  
 297 as nodes' consensus that control their flow. This characteristic is essential to associate diagnoses and  
 298 illness to lifestyles in the province of Risaralda.

299 To incorporate topological aspects of motifs in weighted networks (or strength), we use the intensity  
 300 metrics and motif coherence developed by [38]. Intensity,  $I(g)$ , for subgraph  $g$  with vertices  $V_g$  and  
 301 edges  $l_g$ , is defined as:

$$302 \quad I(g) = \left( \prod_{(kj) \in l_g} W_{kj} \right)^{1/|l_g|} \quad (3) \text{ This guarantees link qualification in motifs from } W_{kj} \text{ values, and}$$

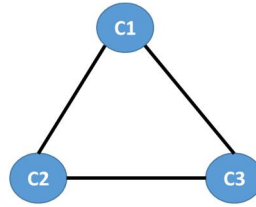
303 leads to prioritizing communities that build up morbidity. After that, we establish the coherence as  
 304  $Q(g)$ , which allows us to study the consensus between people at the edges inside motifs. Coherence  
 305 takes values near to the most important unit in its subgraph to establish the association between  
 306 subpopulations, and it is defined as the ratio between intensity,  $I(g)$ , and the geometric mean of their  
 307 weights,  $W_{kj}$ , as presented in equation (4):

$$308 \quad Q(g) = \frac{I(g)}{\sqrt[|l_g|]{\sum_{(ij) \in l_g} W_{kj}}} \quad (4)$$

### 309 2.4 Mortality Network Analysis

310 Mortality is a much simpler network than morbidity and does not need to processes all morbidity  
 311 data, as mentioned in Sections 3.1 to 3.3. The structure of this network is described in Figure 6. Node

312  $C_1$  is the main cause of death in the diagnosis, and  $C_2$  and  $C_3$  are the secondary or underlying causes  
 313 of death. Because the aim of this study is to observe the relationship between lifestyles that end up  
 314 resulting in a particular cause of death, it is not necessary to run the topological measurements for  
 315 this mortality network.



316

317 **Figure 6.** Structure of the Mortality Network.

318 Nodes or vertexes represent causes of death,  $C_1$  is the main cause, while  $C_2$  and  $C_3$  are the secondary ones. Edges  
 319 or links represent the number of cooccurrences of both primary and secondary causes for every individual.

320 Source: Authors' scheme.

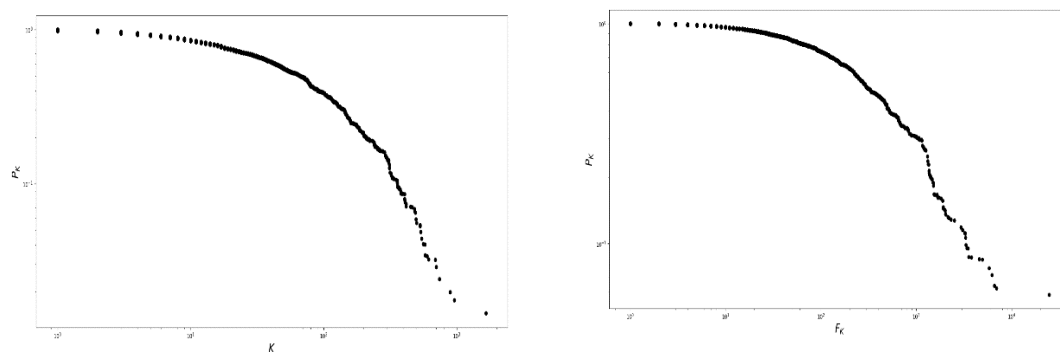
### 321 3. Results and Discussion

322 First, we present the morbidity and mortality networks for the victims of internal armed conflict, as  
 323 recorded in the BDUA dataset. Thereafter, the results for non-displaced population who qualify for  
 324 SISBEN I and II are presented.

325

#### 326 3.1. Results for the victims of Internal Armed Conflict

327 Figure 7 is the complementary cumulative distribution function that results from applying the  
 328 methodology for studying morbidity that was described above. It summarizes the complex network  
 329 of diagnoses for all the individuals included in the BDUA database that were identified as the victims  
 330 of armed conflict.



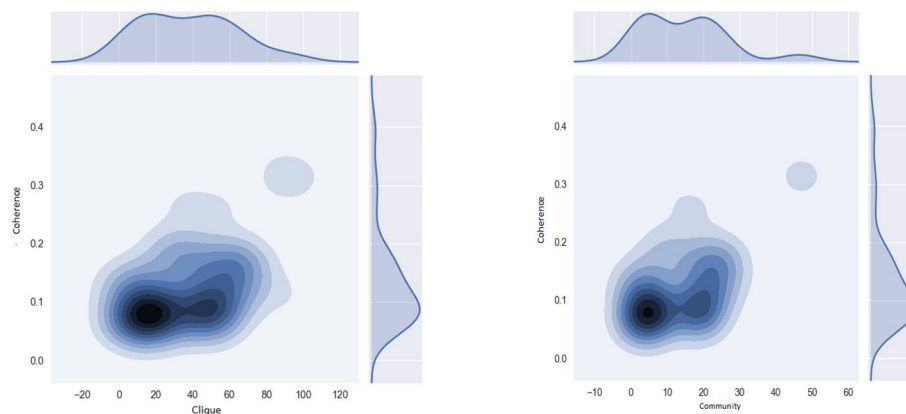
331 **Figure 7.** Nodal degree and nodal strength for the victims of the armed conflict in Colombia, resident in the  
 332 province of Risaralda.

333 Complementary cumulative distribution functions of the nodal degree,  $K$  (left), and the nodal strength  $F_K$  (right)  
 334 for the Morbidity Network of Diagnoses for the victims of the armed conflict in Colombia, resident in the  
 335 province of Risaralda. Years from 2011 to 2016.

336 Source: Authors' own calculations.

337

338 The results are similar to those of the morbidity network for the whole province of Risaralda, in the  
 339 full range of  $P_K$  for  $K_i$  and  $F_{ki}$ . The algorithm described in the section of methods allows the detection  
 340 of communities and measures of the intensity and the coherence of motifs. Figure 8 presents the  
 341 results, showing the coherence values for the two motifs selected to analyze the network diagnoses  
 342 for the victims, with cohesive subgraphs due to cliques and communities. In both cases, motifs have  
 343 low coherence values, between 0.05 and 0.1, with the exception of a small set with values close to 0.3.  
 344 The motifs with the highest coherence are the cliques labeled 80 and 82 as well as community 50  
 345 composed by the diagnoses in Figure 9 (E782, E781, I119, E119 and E039).  
 346

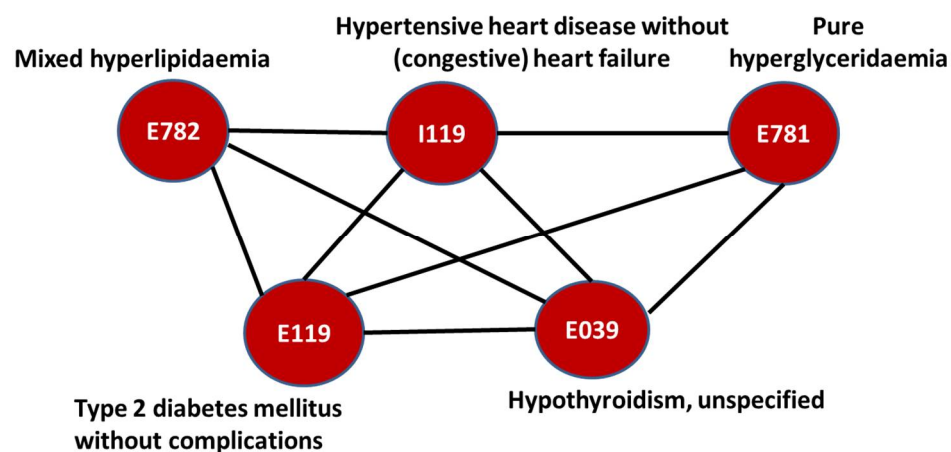


347  
 348  
 349  
 350  
 351  
 352  
 353

**Figure 8.** Motif coherence for cliques and communities.

Motif coherence for cliques (left) and for communities (right) for the Morbidity Network of Diagnoses of the victims of the armed conflict in Colombia, residents in the province of Risaralda. Years from 2011 to 2016.

Source: Authors' own calculations.

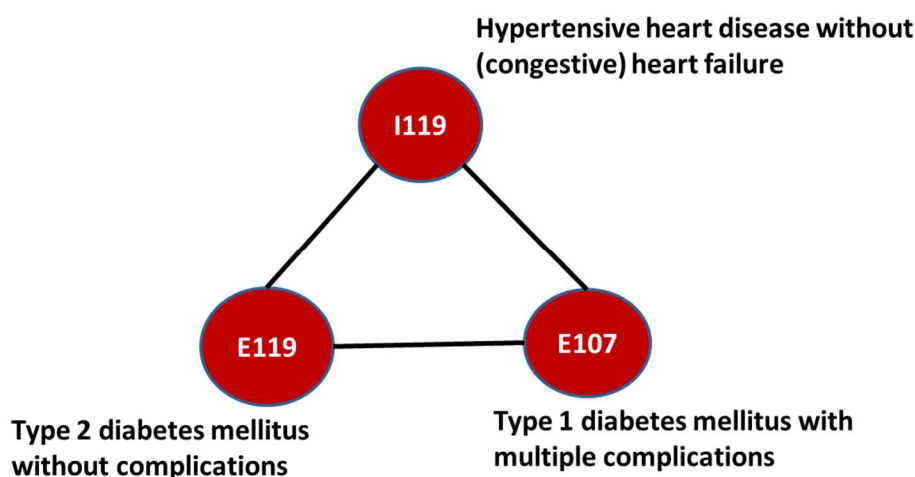


354  
 355  
 356  
 357  
 358

**Figure 9.** Community of diagnoses with the highest coherence level. Morbidity Network of Diagnoses for the victims of the armed conflict in Colombia, residents in the province of Risaralda. Years from 2011 to 2016.

Source: Authors' own calculations.

359 This community analysis evidences a relation between lifestyle and some diseases, reflected in all  
 360 portrayed illnesses except hypothyroidism unspecified (E039). For instance, mixed hyperlipidaemia  
 361 (E782) is high cholesterol and triglycerides in an individual, and pure hyperglyceridaemia (E781) is  
 362 the unusual increase of triglycerides. In most cases, the medical literature relates both diseases to  
 363 “unhealthy lifestyles”, such as sedentary behavior, diets based on a high intake of saturated fat,  
 364 mostly from animal sources and/or empty carbohydrates, and smoking and alcohol consumption  
 365 [39]. Similarly, hypertensive heart disease without (congestive) heart failure (I119) holds a larger  
 366 genetic component [40-41]; its risk factors include the intake of too much salt, saturated fats and  
 367 empty calories. The same risk factors apply to diabetes mellitus without complications, non-insulin  
 368 dependent (E119) [42-43] and this could deepen other conditions such as hypothyroidism unspecified  
 369 (E039). Moreover, all five of these diseases have obesity as a risk factor [44-45], and also have a high  
 370 prevalence in Risaralda and Colombia, according to the Ministry of Health [46].  
 371 The mortality network results are summarized in Figure 10. Unsurprisingly, there is some  
 372 overlapping of the three denoted diseases. Both diabetes mellitus without complications, non-insulin  
 373 dependent (E119) and hypertensive heart disease without (congestive) heart failure (I119) are  
 374 common to both results, denoting the final fatal outcome in both illnesses.



375

376

**Figure 10.** Clique with the highest overlapping in morbidity and mortality networks.

377

Mortality Network of Diseases for the victims of the armed conflict in Colombia, residents in the province of  
 378 Risaralda. Years from 2011 to 2016.

379

Source: Authors' own calculations.

380

381

The remaining disease in Figure 10 is diabetes mellitus with multiple complications, insulin  
 382 dependent, which is mostly due to genetic conditions; however, some Type II diabetes patients may  
 383 develop Type I diabetes mellitus if the risks of the former are not controlled, despite medical  
 384 treatment.

385

386

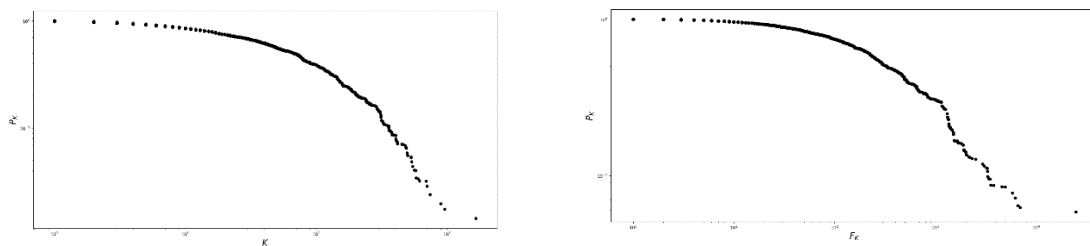
### 3.2 Results for the Population Classified as SISBEN I and II

387

Similar to the previous results for the victims of armed conflict, we identified from the BDUA  
 388 dataset those residents in the province of Risaralda who qualify for subsidies from the State

389 welfare programs, namely those recorded in the subsidiary system SISBEN I and II. The SISBEN  
 390 system is a central government program designed to properly identify the poorest households in  
 391 Colombia. The SISBEN system collects the socioeconomic characteristics of all households in  
 392 Colombia who are potentially poor. All local Majors report this information of potential  
 393 households to the National Planning Department (*Departamento Nacional de Planeación*) is in charge  
 394 of running the surveys and producing a wealth index that categorizes households into six SISBEN  
 395 levels, with I the poorest and VI the richest. As a result, this stratification allows for the allocation  
 396 of national subsidy programs. Households can check their SISBEN status and ask for a  
 397 (re)categorization when they think they were not included in the database or when sudden  
 398 economic changes make them fall into poverty. The index construction follows a secret formula  
 399 to avoid moral hazard.<sup>1</sup>

400 The results are shown in Figures 11 and 12; for simplicity, we refer to this population as SISBEN I  
 401 and II. The pattern in the degree distribution and node strength for morbidity is very much the  
 402 same as that reported for the victims of internal armed conflict. Therefore, the cut was done at the  
 403 same interval than before, the results are portrayed in Figure 12.  
 404



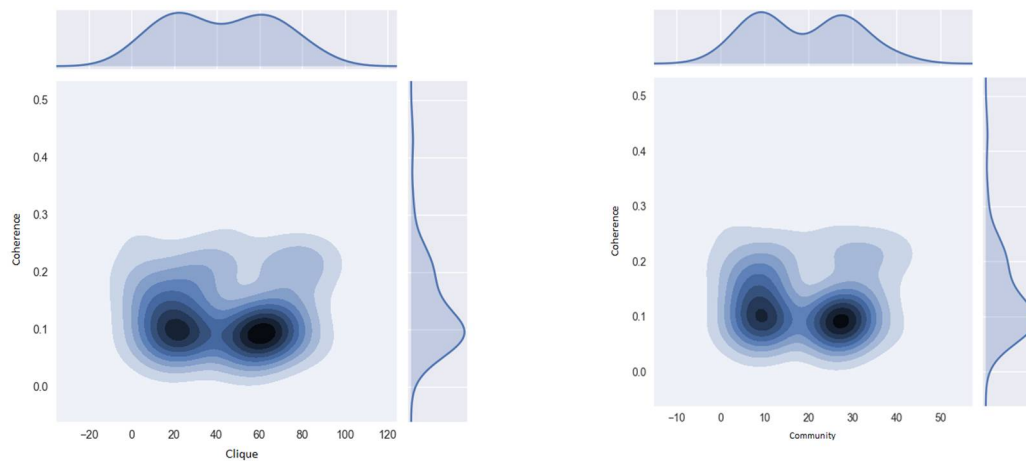
405  
 406 **Figure 11.** Nodal degree and nodal strength for the SISBEN I and II population resident in the province of  
 407 Risaralda.

408 Complementary cumulative distribution functions of the nodal degree,  $K$  (left), and the nodal strength  $F_K$  (right)  
 409 for the Morbidity Network of Diagnoses of the SISBEN I and II population, resident in the province of Risaralda.  
 410 Years from 2011 to 2016.

411 Source: Authors' own calculations.

412  
 413  
 414

<sup>1</sup> See <http://govco.co/sisben/> (accessed on September 13, 2017).

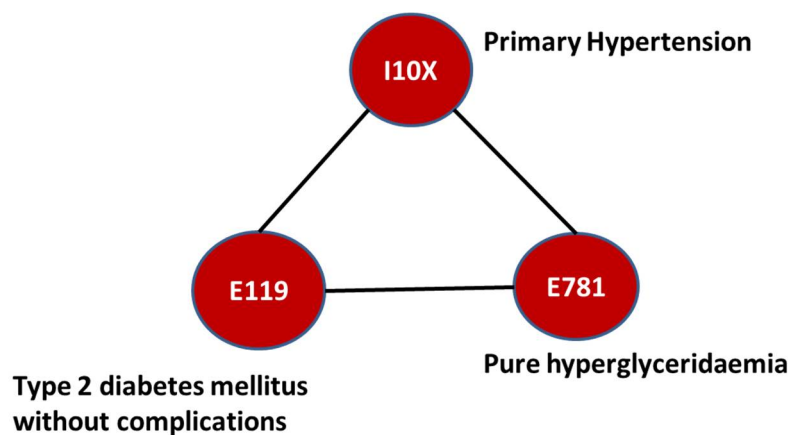


415 **Figure 12.** Motif coherence for cliques and SISBEN I and II. Motif coherence for cliques (left) and SISBEN I and  
 416 II (right) for the Morbidity Network of Diseases for victims of the armed conflict in Colombia, resident in the  
 417 province of Risaralda. Years from 2011 to 2016.

418 Source: Authors' own calculations.

419 By contrast, motif coherence levels for SISBEN I and II are dissimilar to those of the victims of the  
 420 armed conflict, despite most cliques being between 0.05 and 0.1. In fact, the highest coherence level  
 421 is 0.27 and allows the construction of Figure 13. Despite one clique having this 0.27 level, it overlaps  
 422 with its own value and replicates for the entire network, meaning it has the largest influence for most  
 423 diseases for SISBEN I and II.

424



425

426 **Figure 13.** Highest overlapping in morbidity and mortality for SISBEN I and II. Mortality Network of Diseases  
 427 from SISBEN I and SISBEN II population resident in the province of Risaralda. Years from 2011 to 2016.

428 Source: Authors' own calculations.

429

430 This motif is smaller than that of the victims of armed conflict; however, it coincides for one diagnosis,  
 431 diabetes mellitus without complications (E119). The other two diseases, pure hyperglyceridaemia  
 432 (E781) and primary hypertension (I10X), are not fatal outcomes for victims of internal armed conflict,  
 433 but the former is part of the morbidity network. In general, these results show differences between  
 434 both populations. Nonetheless, unhealthy lifestyles lead to similar but different mortality outcomes.  
 435 For instance, a hazard risk for hypertension, beyond an unhealthy lifestyle, is diabetes mellitus [39,

436 47-50]. Thus, the causes of these fatal consequences are all connected and in most cases linked to  
437 unhealthy lifestyles, as described before.

438

439 Both populations under study are, without question, highly vulnerable to poor health conditions, as  
440 a consequence of the extreme socioeconomic conditions that both face. This study's results show an  
441 unhealthy lifestyle as the common ground for the different diseases. More importantly, the  
442 differences in the final causes of death reported in each case, which result in access to health  
443 treatments and prognoses for both populations, are different.

444

445

#### 446 **4. Conclusions**

447 Despite the evidence in the literature that refugees and IDPs have worse health conditions than the  
448 rest of the population, little had been studied in terms of morbidity and its link to mortality for the  
449 victims of the internal armed conflict in Colombia. In part, this is due to the lack of reliable data.  
450 However, this paper presents enough evidence from newly released administrative records with  
451 health-related micro data to show the main morbi-mortality causes for the victims of armed conflict,  
452 as recorded in those databases. The vast majority of self-reported victims in the dataset are IDPs, 80%;  
453 however, it is a novelty to account for health issues for victims of all human rights violations due to  
454 the conflict in Colombia.

455

456 Combining micro data from administrative records is stimulating, but linked unique identification  
457 allows this process. More challenging, however, is the analysis of such large datasets. The proposed  
458 methodology, based on complex network analysis, led us to fully identify the exact morbidity and  
459 mortality patterns from an ocean of chaotic information.

460

461 To establish a reference point, we analyzed the morbi-mortality frequencies of the victims of internal  
462 armed conflict and the poorest population (SISBEN I and II) resident in the province of Risaralda in  
463 Colombia. The results proved that both subpopulations are vulnerable and suffer non-communicable  
464 diseases mostly related to unhealthy lifestyles, such as Type II diabetes, hypertension and  
465 hyperglyceridaemia. However, victims of armed conflict show mortality outcomes that prove the  
466 deepening of those conditions (e.g., Type I diabetes).

467

468 This first approach to measure the morbi-mortality profiles of the victims of armed conflict in  
469 Colombia poses interesting results that can be useful for the design of targeted public health policies,  
470 particularly in a post-conflict scenario such as the one Colombia faces. As all first approaches,  
471 however, it comes with its limitations. First, and as stated before, the administrative records do not  
472 fully account for all the victims of armed conflict. As they are probably under-reported, the results  
473 presented here are to be read carefully and can be stated as a lower bound of the total number of  
474 victims.

475

476 Second, morbi-mortality patterns differ by age and sex in most populations, and the victims of  
477 internal armed conflict may not be an exception. We expect to expand, in the near future, this kind of  
478 analysis to subdivisions of the information by sex and age groups to show a more detailed morbi-  
479 mortality map for the victims of armed conflict.

480

481 Lastly, processing massive datasets requires powerful hardware. Similar to the analysis of other large  
482 datasets, we recommend using partition data, rather than working with a much larger set such as,  
483 for instance, information on the entire Colombian territory. We suggest dividing data in a country  
484 into subnational entities such as provinces (*departamentos*) as we have done, by starting with one of  
485 the smallest provinces in Colombia, and replicating this exercise with other provinces. Researchers  
486 should also use information from multiple years to avoid misinformation from a potential morbi-  
487 mortality shock, bust or registration issue.

488

489 **Acknowledgments:** This paper was supported by Project *DESARROLLO DE CAPACIDADES CT+I PARA*  
490 *SIMULACIÓN DE POLÍTICAS PUBLICAS EN SALUD Y SEGURIDAD SOCIAL EN EL DEPARTAMENTO DE*  
491 *RISARALDA*. BPIN Code: 20150000100017 (Fundación Salutia – Centro de Investigación en Salud and  
492 Universidad Nacional de Colombia),

493

494 **Author Contributions:** Study design, Methodology, Data collection and analysis, Rafael Rentería-Ramos;  
495 Writing, Editing, Reviewing, Rafael Hurtado-Heredia and B Piedad Urdinola

496 **Conflicts of Interest:** The authors declare that there is no conflict of interest

497

498

## 499 References

- 500 1. Keely, C.; Reed, H.; Waldman, R. "Understanding Mortality Patterns in Complex Humanitarian  
501 Emergencies". In NRC (eds.) *Forced Migrat and Mortal*. National Academies Press. *Washington, DC*. 2001
- 502 2. Akresh, R.; Lucchetti, L.; Thirumurthy, H. Wars and child health: Evidence from the Eritrean-Ethiopian  
503 conflict. *J of Devel Econo*, **2012**, 99(2), 330–340. DOI:  
504 <http://doi.org/10.1016/j.jdeveco.2012.04.001>
- 505 3. Brungborg, H.; Urdal, H. The Demography of Conflict and Violence: An Introduction. *J of Peace Rese*,  
506 **2005**, 42(4), 371–374. <http://doi.org/10.1177/0022343389026001001>
- 507 4. Bundervoet, T.; Verwimp, P.; Akresh, R. Health and Civil War in Rural Burundi. *J of Hum Resou*, **44**(2),  
508 **2008**, 536–563. <http://doi.org/10.3368/jhr.44.2.536>
- 509 5. Toole, M. J.; Waldman, R. J. Refugees and displaced persons: war, hunger, and public health. *Jama*, **1993**,  
510 **270**(5), 600–605.
- 511 6. Urdinola, B. P. Could Political Violence Affect Infant Mortality? The Colombian Case. *Coyun Soc*, **2004**  
512 **(31)**, 63–79.
- 513 7. De Azevedo Barros, M. B.; Lima, M. G.; Medina, P. B.; Szwarcwald, C. L.; Malta, D. C. Social inequalities  
514 in health behaviors among Brazilian adults: National Health Survey, 2013. *International journal for*



- 515 equity in health, vol. 15, no 1, p. 148.
- 516 8. RNI. Registro Nacional de Información. Available Online:  
517 <https://cifras.unidadvictimas.gov.co/Home/Vigencia> (accessed on June 05 2017).
- 518 9. UNHCR. UNHCR country operations profile – Colombia, Available Online: <http://goo.gl/2a7r1S>,  
519 (accessed on 26 March 2015).
- 520 10. Torres, A. F.; Urdinola, B. P. Armed Conflict and Fertility in Colombia, 2000–2010. *Popul Rese and Poli*  
521 *Rev*, 2018, 1-41.
- 522 11. Blattman, C.; Miguel, E. Civil War. *J of Econ Liter*, 2010, 48(1), 3–57. [http://doi.org/10.1016/S1574-](http://doi.org/10.1016/S1574-0013(06)02023-0)  
523 [0013\(06\)02023-0](http://doi.org/10.1016/S1574-0013(06)02023-0)
- 524 12. Byarugaba, J.; Kielkowski, D. Reflections on trauma and violence-related deaths in Soweto, July 1990-  
525 June 1991. *Sou Afr Med J*, 1994, 84(9), 610-614 WHO. *World Report on Violence and Health*. Edited by E.G.  
526 Krug, L.L. Dahlberg, J.A. Mercy, A. B. Zwi and R. Lozano. WHO, Geneva, 2002. Bilak, A. *Global overview*  
527 *2015: people internally displaced by conflict and violence*. Internal Displacement Monitoring Centre,  
528 Norwegian Refugee Council, 2015.
- 529 13. Franco, S. The effects of the armed conflict on the life and health in Colombia. *Cièn & Saúid Cole*, 2005,11,  
530 1247-1258.
- 531 14. Reconciliacion Colombia. Los desplazados son más pobres que el resto de la población, Available  
532 online: <http://goo.gl/eK5F38> (accessed on 16 February 2015 )
- 533 15. Doocy, S. Lyles, E.; Robertson, T.; Akhu-Zaheya, L.; Oweis, A.; Burnham, G. Prevalence and care-seeking  
534 for chronic diseases among Syrian refugees in Jordan. *BMC Pub H*, 2015.15,1, 1097.
- 535 16. Doocy, S.; Lyles, E.; Delbiso, T. D.; Robinson, C. Internal displacement and the Syrian crisis: an analysis  
536 of trends from 2011–2014. *Con and H*, 2015, 9,1-33.
- 537 17. Audet, M.; Dumas, A.; Binette, R.; Dionne, I. Lifestyle inequalities: Explaining socioeconomic  
538 differences in preventive practices of clinically overweight women after menopause. *Qual H R*,  
539 2017,27,10, 1541-1552.
- 540 18. Brown, D.; Brewster, L.; The food environment is a complex social network. *Soc Sci & Med*, 133,  
541 2015,202-204.
- 542 19. Newman, M. *Networks: an introduction*. Second Edition, Oxford University Press, USA, 2010, ISBN-  
543 10: 0198805098
- 544 20. Strogatz, S. H. (2001). Exploring complex networks. *Nat*, 2001, 410, 268-276.
- 545 21. Dorogovtsev, S.; Mendes, J. *Evolution of Networks: From Biological Networks to the Internet and WWW*  
546 *Oxford*. Oxford University Press, 2003, ISBN-13: 9780198515906
- 547 22. Estrada, E.; Rodriguez-Velazquez, J. *Subgraph centrality in complex networks*. *Phys Rev E*, 2005,71,5,  
548 056103.
- 549 23. Bichara, D.; Castillo-Chavez, C. Vector-borne diseases models with residence times—a lagrangian  
550 perspective. *Math Bio*, 2016, 281, 128-138.
- 551 24. Pastor-Satorras, R.; Vespignani, A. Immunization of complex networks. *Phys Rev E*, 2002, 65,3, 036104.

- 552 25. Towers, S.; Brauer, F.; Castillo-Chavez, C.; Falconar, A.; Mubayi, A.; Romero-Vivas, C. Estimate of the  
553 reproduction number of the 2015 Zika virus outbreak in Barranquilla, Colombia, and estimation of the  
554 relative role of sexual transmission. *Epid*, **2016**,*17*, 50-55.
- 555 26. Balcan, D.; Colizza, V.; Gonçalves, B.; Hu, H.; Ramasco, J.; Vespignani, A. (2009). Multiscale mobility  
556 networks and the spatial spreading of infectious diseases. *Proce of the Nat Acad of Scie*, **2009**,*106*,*51*,  
557 21484-21489.
- 558 27. Brauer, F.; Castillo-Chavez, C.; Mubayi, A.; Towers, S. Some models for epidemics of vector-transmitted  
559 diseases. *Infect Dise Model*, **2016** *1*,*1*, 79-87.
- 560 28. Zuin, M.; Caserta, C.; Romanò, L.; Mele, A.; Zanetti, A.; Cannatelli, R.; Battezzati, P. Seroepidemiology  
561 of HEV and HAV in two populations with different socio-economic levels and hygienic/sanitary  
562 conditions. *Eur J of Cli Microbi & Infe Disea*, **2017**,*36*,*3*, 479-485.
- 563 29. Vyhrystyuk, M.; Rogozhnikova, T.; Semenova, G.; Shusharina, I.; Savelyeva, E. A.; Kungina, O. (2017).  
564 The Evolution of the Concept of Public Health in Modern Youth Discourse. *Perta J of Soc Sci & Hum*,  
565 **2017**,*25*,151-164
- 566 30. Palla, G.; Derényi, I.; Farkas, I.; Vicsek, T. Uncovering the overlapping community structure of complex  
567 networks in nature and society, *Nat*, **2005**, 435,814-824.
- 568 31. Fortunato, S. Community detection in graphs. *Phys Rep*, **2010**, 486,3, 75-174.
- 569 32. Kumpula, J.; Kivelä, M.; Kaski, K.; Saramäki, J. Sequential algorithm for fast clique percolation. *Phy Rev*  
570 *E*, **2008**, *78*,*2*, 026109.
- 571 33. Milo, R.; Shen-Orr, S.; Itzkovitz, S.; Kashtan, N.; Chklovskii, D.; Alon, U. Network motifs: simple  
572 building blocks of complex networks. *Sci*, **2002**,*298*, 824-827.
- 573 34. Green, S.; Şerban, M.; Scholl, R.; Jones, N.; Brigandt, I.; Bechtel, W.. Network analyses in systems  
574 biology: new strategies for dealing with biological complexity. *Syn*,**2017**, 1-27.
- 575 35. Smoly, Y.; Lerman, E.; Ziv-Ukelson, M.; Yeger-Lotem, E. MotifNet: a web-server for network motif  
576 analysis. *Bioin*, **2017**,*33*,*12*, 1907-1909.
- 577 36. Delmas, E.; Besson, M.; Brice, H.; Burkle, L.; Dalla - Riva, V.; Fortin, J.; Olesen, M. Analyzing ecological  
578 networks of species interactions. *BioRev*, **2018**,*1*-*21*, DOI: 10.1111/brv.12433
- 579 37. Rodríguez-Rodríguez, C.; Jordano, P.; Valido. Functional consequences of plant-animal interactions  
580 along the mutualism-antagonism gradient, *Eco*, **2017**,*98*,*5*, 1266-1276.
- 581 38. Onnela, J.; Saramäki, J.; Kertész, K. Intensity and coherence of motifs in weighted complex networks.  
582 *Phys. Rev. E*. **2005**, *71*, *6*, 065103.1–065103.4.
- 583 39. Co, J., Jeffrey, J., Emmett, M.; Modak, A.; Sondike, S. Obesity, hypertension and metabolic syndrome in  
584 children in West Virginia. *T W Vir Med J*, **2015**, *111*, *4*, 20-24.
- 585 40. Varan, B.; Tokel, K.; Yilmaz, G. Malnutrition and growth failure in cyanotic and acyanotic congenital  
586 heart disease with and without pulmonary hypertension. *Arch of Dise in Child*, **1999**, *81*, *1*, 49-52.
- 587 41. Taimeh, Z.; Duprez, D., Garry, D. "Heart Failure Prevention." Congestive Heart Failure and Cardiac  
588 Transplantation. Springer, *Cham*, **2017**. 267-284.

- 589 42. Powers, M.; Bardsley, J.; Cypress, M.; Duker, P.; Funnell, M.; Fischl, A.; Vivian, E. Diabetes self-  
590 management education and support in type 2 diabetes: a joint position statement of the American  
591 Diabetes Association, the American Association of Diabetes Educators, and the Academy of Nutrition  
592 and Dietetics. *T Diab Ed*, **2017**, 43, 1, 40-53.
- 593 43. Furukawa, S.; Fujita, T.; Shimabukuro, M.; Iwaki, M.; Yamada, Y.; Nakajima, Y.; Shimomura, I.  
594 Increased oxidative stress in obesity and its impact on metabolic syndrome. *T J of Clin Inv*, **2017**, 114, 12,,  
595 1752-1761.
- 596 44. Moran, P.; Zhang, J.; Gebregziabher, M.; Brownfield, L.; Davis, S.; Schreiner, D.; Ball, S. Chaos to  
597 complexity: leveling the playing field for measuring value in primary care. *J of Eva in Clin Prac*, **2017**,  
598 23(2), 430-438.
- 599 45. Ohlsson, B.; Orho-Melander, M.; Nilsson, P. Higher Levels of Serum Zonulin May Rather Be Associated  
600 with Increased Risk of Obesity and Hyperlipidemia, Than with Gastrointestinal Symptoms or Disease  
601 Manifestations. *Inter J of Mol Sci*, **2017**, 18(3), 582.
- 602 46. Ministerio de Salud (2017). Modelo Integral de Atención de Salud "MIAS", Available  
603 online: [https://www.minsalud.gov.co/Paginas/presentaciones-modelo-integral-atencion-salud-](https://www.minsalud.gov.co/Paginas/presentaciones-modelo-integral-atencion-salud-mias.aspx)  
604 [mias.aspx](https://www.minsalud.gov.co/Paginas/presentaciones-modelo-integral-atencion-salud-mias.aspx), (accessed on 16 January 2018 )
- 605 47. Retterstøl, K.; Narverud, I.; Selmer, R.; Berge, E.; Osnes, V.; Ulven, M.; Iversen, O. Severe  
606 hypertriglyceridemia in Norway: prevalence, clinical and genetic characteristics. *Lip in He and Dise*, **2017**  
607 16(1), 115.
- 608 48. Personhood, C.; Chronic, W.; Home, I.; Guideline, H.; Rousseau, H.; Water, P. Self-management of  
609 Hypertension in High-Risk Patients, *JAMA*, **2014**, 312, 8, 763-858.
- 610 49. Xu, Y.; Niu, Y.; Gao, Y.; Wang, F.; Qin, W.; Lu, Y.; Xiong, W. Borapetoside E, a Clerodane Diterpenoid  
611 Extracted from *Tinospora crispa*, Improves Hyperglycemia and Hyperlipidemia in High-Fat-Diet-  
612 Induced Type 2 Diabetes Mice. *J. Nat. Prod.*, **2017**, 80 8, 2319–2327. Brinton, E. Management of  
613 hypertriglyceridemia for prevention of atherosclerotic cardiovascular disease. *Endo and Meta Clin of NA*,  
614 **2016**, 45, 1, 185-204.