SDH mathematical model

This manuscript ([permalink]) was automatically generated from [juanfernandezgracia/SDH_epi_modeling_manu@9bbe93d] on May 19, 2020.

Authors

- **Juan Fernández-Gracia**
  XXXX-XXXX-XXXX-XXXX · juanfernandezgracia · JuanFerGra
  IFISC (CSIC-UIB), Palma de Mallorca, Spain · Funded by UIB postdoct program

- **Juan M. Pericàs**

  Infectious Disease Department, Hospital Clínic de Barcelona, Barcelona, Spain; Vall d'Hebron Institute for Research (VHIR), Barcelona, Spain

- **Víctor M. Eguíluz**

  IFISC (CSIC-UIB), Palma de Mallorca, Spain

- **Joan Benach**

  GREDS-EMCONET, Department of Political and Social Sciences, UPF, Barcelona, Catalonia, Spain

Abstract

Social determinants of health are typically not taken into account when building mathematical models of infectious diseases. Here we address this issue by first analysing data on social determinants and their relation to COVID-19 prevalence and then by developing mechanistic models that take into account social determinants explicitly, in order to exemplify how to do it and their consequences.

Introduction

Social determinants of health are defined as the economic and social conditions that influence individual and group differences in health status.

The general idea behind this work is that social determinants of health are crucial for the understanding of the different impacts of an epidemic on different social groups. It is clear during the present pandemic at the level of countries, that healthy ones are able to apply social distancing and hygiene measures that are much more difficult -or even impossible- in developing countries. This is also true at lower scales as within countries or even within cities (**cite papers inequalities and COVID-19 in the US**). Many mathematical modeling efforts have been made for the forecasting of epidemics, but only few have incorporated social determinants of health into them explicitly.

Theoretical framework

One point to take into account is that deaths in residencies for the elder have their own logic and thus should be treated differently in the analysis. *Health care workers pose a similar challenge in the analysis and modeling. (???)* Regarding the rest of the population one can divide the social determinants impacting their risk during an epidemic at least into the following classes and subclasses:

1. **Inequality**. Indices that reflect material deprivation.

- *Housing*. Overcrowding is a driver of contagion.
- *Unemployment*. Unemployment leads to lower income levels and thus material deprivation, but also to lower mobility.
- *Occupation*. Depending on the type of job, working from home can be possible or not.

2. **Population density**. A crowded living environment leads also to an enhanced frequency of close contacts, fostering contagion.
3. **Public transport**. The use of public transport modes is also a strong driver of close contacts, leading again to high contagion risks.

Data analysis (WORK IN PROGRESS)

The approach is to analise the data for the social determinants of health that we want to consider at a certain spatial aggregation scale and find which ones and how correlate with the impact of the COVID-19 epidemic. For that we will explore different methods of dimensionality reduction (principal component analysis, tsne...).

The data is divided into two groups:

1. Epidemiological data.
   - Prevalence (can be taken at different times).

- Deaths (can be taken at different times).
- Observed mortality rate (dividing deaths by prevalence).
- Seroprevalence (**GET DATA FROM ONGOING STUDY IN SPAIN**).

2. Social determinants of health.
   - Population.
   - Population density.
   - Sex.
   - Age.
   - Share of migrant population.
   - Housing.
     - Number of persons per home.
     - Size of homes.
   - Occupation.
   - Unemployment.
   - Income.
   - Mobility.

Spain at the level of povinces (PRELIMINARY RESULTS)

For this aggregation scale we used the following datasets:

1. Epidemiological data from Escovid19data, which offers the data by provinces in Spain..

   - Prevalence (can be taken at different times).
   - Deaths (can be taken at different times).
   - Observed mortality rate (dividing deaths by prevalence).

2. Social determinants of health.

   -Population from INE. From the census 2011 at the level of censal sections and aggregated up to provinces. The level of census sections is very detailed, but due to statistical secret some variables are not given and thus we to make some assumptions on them. I am taking them as 0 for now. Ideally I would change this dataset for one at the level of provinces with no missing data.

   - Age also from INE. From the census 2011 at the level of censal sections and aggregated up to provinces. We count the percentage of people over 64 for this variable.
   - Share of migrant population also from INE. We count the share of foreigners.
   - Housing.
     - Number of persons per house also from INE. We divide the population of a province by the number of homes that are there. Note that this does not tell if the home is big or small for the number of people living there.
   - Unemployment also from INE, this time directly at the level of provinces
   - Mobility taken from the mobility declared for work in the census 2011. This data contains mobility between municipalities (declared residence and declared work place). We aggregated the data at the level of provinces and defined three indicators of mobility.
     - Intra-mobility, taken as the number of people moving between municipalities of the same province.
     - In-mobility, taken as the number of people coming to work at the focal province.
     - Out-mobility, taken as the number of people going to work to some other province.
   - Deprivation index from this paper. The data for the deprivation index is available here.

Note that, from the variables that we listed in the previous section, there are some missing in the present version of this analysis. These ones will be added and are *population density, sex, size of* *homes, income* and *occupation*. Note also that mobility here is only between municipalities for work as described in the census of 2011 and that there is no distinction made between public and private transport.

We applied principal component analysis (PCA) to the determinants of health, leaving out all of the epidemiological indicators, to later assess their relation.

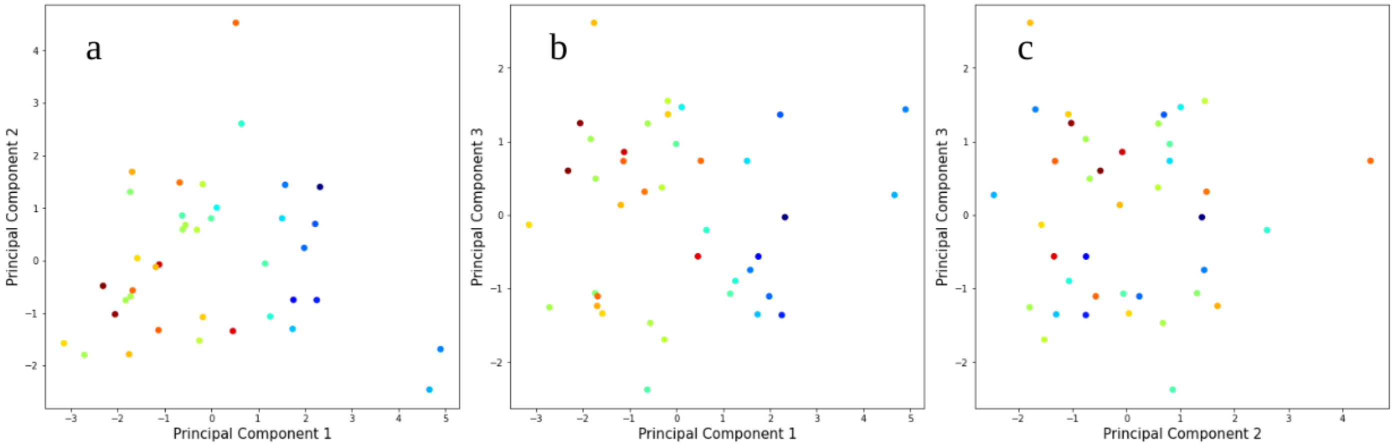

Figure 1: Plot of the data for social determinants for the provinces of Spain as a function of the principal components. The color indicates the prevalence of confirmed COVID-19 cases.

We see from Fig. 1 that there is correlation between the principal components and the prevalence of confirmed cases. Let's dig deeper in that relation. First let's see the weight of each variable for the principal components in Tab. 1.

Table 1: Principal components.

	Var. explained	Pop.	Pop. over 64	Migrants	People/home	Unemployment	Intra mob.	In mob.	Out mob.	Depriv. index
PC1	38%	0.11	-0.45	0.04	0.47	0.44	-0.23	-0.40	-0.39	-0.06
PC2	22%	0.53	-0.12	0.26	-0.21	-0.30	0.54	-0.19	-0.41	-0.08
PC3	15%	-0.01	-0.30	0.73	0.21	-0.22	-0.21	0.41	0.16	0.23

As the first prioncipal component seems to be the one correlating most with prevalence of COVID-19 confirmed cases by province, let's do scatter plots of incidence and the most relevant social determinants for that component.

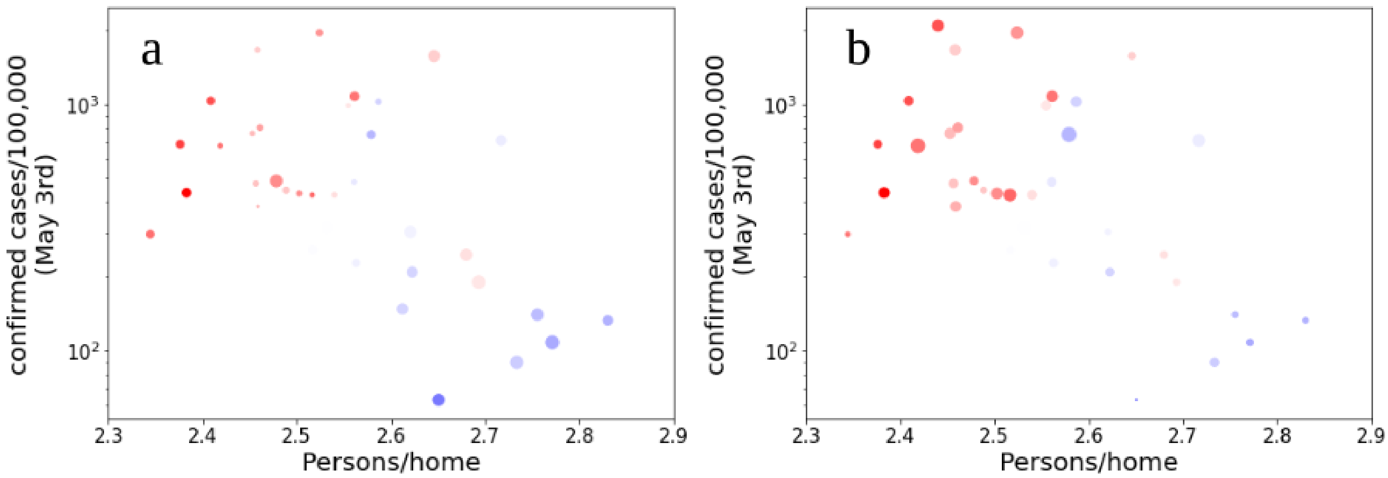

Figure 2: Prevalence of COVID-19 for spanish provinces by May 3rd (y-axis) *vs*. number of persons per home. The color codes for the share of people over 64 for that province (increasing from blue to red). The size of the circles codes for (a) unemployment rate and (b) in-mobility (out-mobility is very similar). The larger the circle, the larger the variable it represents.

From Fig. 2 we can identify several trends. First of all the more persons per home, the less prevalence. I think we should include a measure of the size of homes, as this result is very puzzling. Second, the share of people over 64 is a strong driver of prevalence of the disease. Third, the more unemployment, the less prevalence. *Might be because of less mobility?* Fourth, the more mobility, the more incidence.

Catalonia at the level of ABS (TO DO)

References