


Low-income neighbourhood was a key determinant of severe COVID-19 incidence during the first wave of the epidemic in Paris

Anne-Sophie Jannot ^{1,2,3}, Hector Countouris,¹ Alexis Van Straaten,¹ Anita Burgun,^{1,2,3} Sandrine Katsahian,^{1,2,3} Bastien Rance,^{1,2,3} AP-HP/Universities/ Inserm COVID-19 research collaboration

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/jech-2020-216068>).

¹Hôpital Européen Georges Pompidou, Service d'informatique médicale, biostatistiques et santé publique, AP-HP, Paris, France

²Université de Paris, Paris, France

³Centre de Recherche des Cordeliers, Inserm, Paris, France

Correspondence to

Dr Anne-Sophie Jannot, Hôpital Européen Georges Pompidou, Service d'informatique médicale, biostatistiques et santé publique, AP-HP, Paris 75015, Île-de-France, France; annesophie.jannot@aphp.fr

Received 17 November 2020

Accepted 15 June 2021

Published Online First

30 June 2021

ABSTRACT

Background Previous studies have demonstrated that socioeconomic factors are associated with COVID-19 incidence. In this study, we analysed a broad range of socioeconomic indicators in relation to hospitalised cases in the Paris area.

Methods We extracted 303 socioeconomic indicators from French census data for 855 residential units in Paris and assessed their association with COVID-19 hospitalisation risk.

Findings The indicators most associated with hospitalisation risk were the third decile of population income (OR=9.10, 95% CI 4.98 to 18.39), followed by the primary residence rate (OR=5.87, 95% CI 3.46 to 10.61), rate of active workers in unskilled occupations (OR=5.04, 95% CI 3.03 to 8.85) and rate of women over 15 years old with no diploma (OR=5.04, 95% CI 3.03 to 8.85). Of note, population demographics were considerably less associated with hospitalisation risk. Among these indicators, the rate of women aged between 45 and 59 years (OR=2.17, 95% CI 1.40 to 3.44) exhibited the greatest level of association, whereas population density was not associated. Overall, 86% of COVID-19 hospitalised cases occurred within the 45% most deprived areas.

Interpretation Studying a broad range of socioeconomic indicators using census data and hospitalisation data as a readily available and large resource can provide real-time indirect information on populations with a high incidence of COVID-19.

INTRODUCTION

The COVID-19 outbreak has been raging worldwide since late 2019, killing thousands and disrupting health systems with hundreds of hospitalised patients. During the first wave in France, 30 000 deaths occurred and more than 30 000 patients were hospitalised at its peak.

Some neighbourhoods had a higher incidence than others. From a public health perspective, highlighting neighbourhoods at high risk for severe COVID-19 infections is particularly useful to improve resource allocation. These neighbourhoods could be identified through census data linked to patient localisation, considering that individual outcomes, such as COVID-19 infection, are ultimately a product of context. Indeed, varying trends within countries in many of the core determinants

of health, such as socioeconomic position, suggest a rich field for examining both the impact of social determinants on COVID-19 and the extent to which they depend on context.¹ Numerous ecological studies have already been performed on COVID-19 and wide inequities have been shown.²⁻⁴

These inequities are multifactorial, involving geographical, ethnical and socioeconomic factors, and no specific socioeconomic factors have been highlighted as particularly important regarding COVID-19 burden.⁵ Studying a very broad range of socioeconomic indicators by using all available census indicators with no a priori selection provides a global picture of neighbourhoods at risk and would allow us to highlight unexpected associations between COVID-19 incidence and some socioeconomic factors. This data-mining strategy has become common since the first genome-wide association studies⁶ and has been adapted to many kinds of data, such as phenotypes⁷ and medications.⁸ This type of study would facilitate the identification of specific environmental characteristics strongly associated with risk of COVID-19 with no a priori hypotheses regarding the type of associated indicator.

In this paper, our objective is to analyse to what extent a broad range of census indicators were associated with the incidence of hospitalisation for COVID-19 in the Paris area during the first wave of the epidemic.

MATERIALS AND METHODS

Our analysis follows the recommendations provided by RECORD (REporting of studies Conducted using Observational Routinely collected health Data) (online supplemental table S2).

Data extraction

New patients who tested positive by PCR from 1 February until 4 June, hospitalised in one of the hospitals of Assistance Publique - Hôpitaux de Paris (AP-HP) and living in Paris were aggregated at the Ilots Regroupés pour l'Information Statistique (IRIS) Paris division level, which is the smallest geographical division in France with 2000 inhabitants, on average, to generate the data set for this study. Patient addresses reported by patients when hospitalised were localised through an open application programming interface available from



© Author(s) (or their employer(s)) 2021. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Jannot A-S, Countouris H, Van Straaten A, et al. *J Epidemiol Community Health* 2021;**75**:1143–1146.

the French government website. Given the centralised health-care organisation during the COVID-19 epidemic, COVID-19-infected patients were mostly hospitalised at AP-HP hospitals, and private clinics did not treat patients with COVID-19.

For each IRIS unit, we retrieved all information available as open data at this IRIS scale from 2016 census data,⁹ including 71 indicators on population demographics and distribution, 51 on housing, 86 on professional activity, 55 on family types, 19 on education and 21 on income, for a total of 303 indicators.

Methods

High-risk and low-risk IRIS units were defined using the first decile of this proportion. We then assessed the association of a given socioeconomic indicator above the median value with having a high risk of COVID-19 for each of the 303 socioeconomic indicators using χ^2 test. P values for the different tests performed were represented through a Manhattan plot using the type of indicator as the variable group.

RESULTS

As of 3 June 2020 (the end of the first wave of the epidemic), 3400 patients have been hospitalised in one of the 855 residential areas (online supplemental figure S1), with a median number of hospitalised cases per residential area of 3 (IQR: 1–5). According to 2016 census data, these 855 Paris residential areas had a population of 2 145 817 individuals for a total area of 71.5 km², that is, 30 011 inhabitants/km². These 3400 patients had a median age of 71 years (IQR: 57–83) and 55% were men. The median number of days of hospitalisation was 12 (IQR: 5–34). In total, 20% were hospitalised in the intensive care unit and 7% died 30 days after admission. The first decile of the COVID-19 hospitalisation rate among Paris residential areas was equal to 3 per 1000 inhabitants.

The indicators most strongly associated with this first decile were related to income and occupational status (figure 1 and online supplemental table S3). The indicator most significantly associated with hospitalisation risk was income corresponding to the third decile of the population (OR=9.10, 95% CI 4.98 to 18.39) (figure 2 and online supplemental figure S2). Overall, 86% of COVID-19 hospitalised cases occurred within the 45% most deprived areas. Regarding population indicators, the most significantly associated indicators concerned population activities. Specifically, the rate of workers in blue-collar unskilled occupations was strongly associated with high-risk areas (OR=4.38, 95% CI 2.67 to 7.51), and the rate of highly qualified workers was strongly associated with low-risk areas (OR=0.19, 95% CI 0.10 to 0.31). Population demographics were much less associated with hospitalisation risk. Among these factors, the highest association was observed for the rate of women aged between 45 and 59 years (OR=2.17, 95% CI 1.40 to 3.44). Of note, population density was not associated with hospitalisation risk (OR=1.1, 95% CI 0.72 to 1.69). In addition, male sex (OR=1.27, 95% CI 0.83 to 1.95), age 80 years or older (OR=0.76, 95% CI 0.49 to 1.16) and age 65–79 years old (OR=1.05, 95% CI 0.69 to 1.61) were not associated with hospitalisation risk. The rate of immigrants (OR=3.20, 95% CI 2.01 to 5.25), rate of active workers in unskilled occupations (OR=5.04, 95% CI 3.03 to 8.85) and rate of unemployment (OR=4.69, 95% CI 2.84 to 8.14) were strongly associated with hospitalisation risk. The most strongly associated housing indicators were the primary residence rate (OR=5.87, 95% CI 3.46 to 10.61) and public low-rent housing rate (OR=4.69, 95% CI 2.84 to 8.14). The most associated education indicator was the rate of women over 15 years old with no diploma (OR=5.04, 95% CI 3.03 to 8.85). The two most associated household indicators were the rate of household heads with an unskilled occupation (OR=4.38, 95% CI 2.67 to

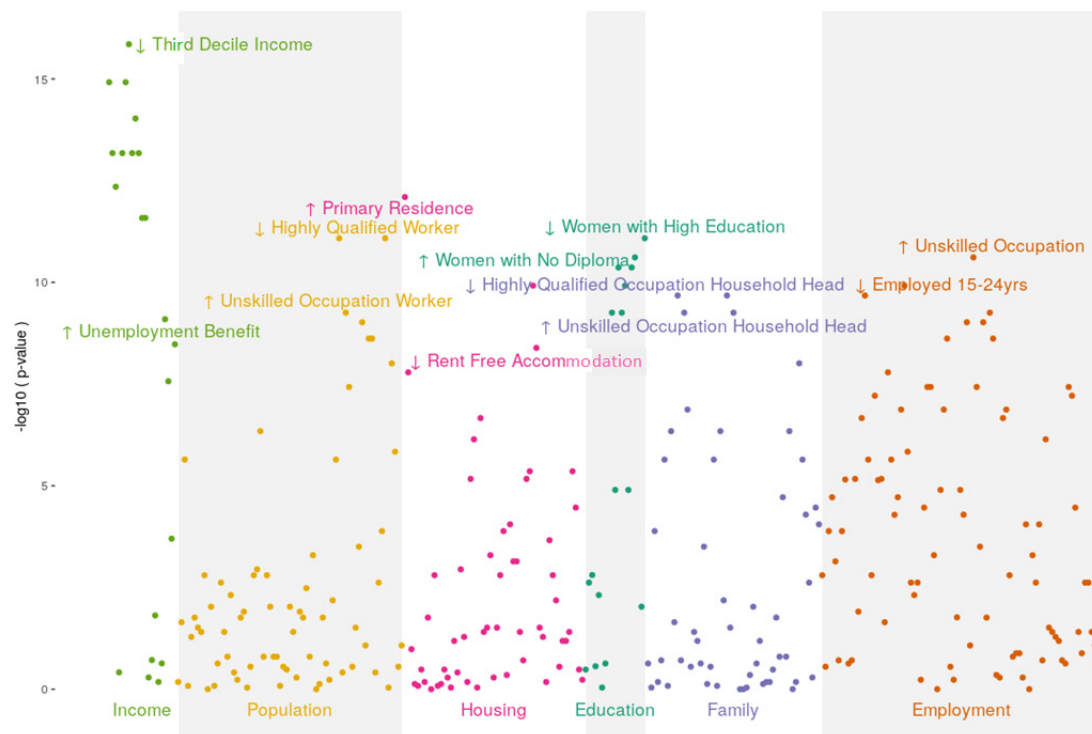


Figure 1 Manhattan plot representing the p value of the 303 association tests between COVID-19 hospitalisation rate and each socioeconomic factor.

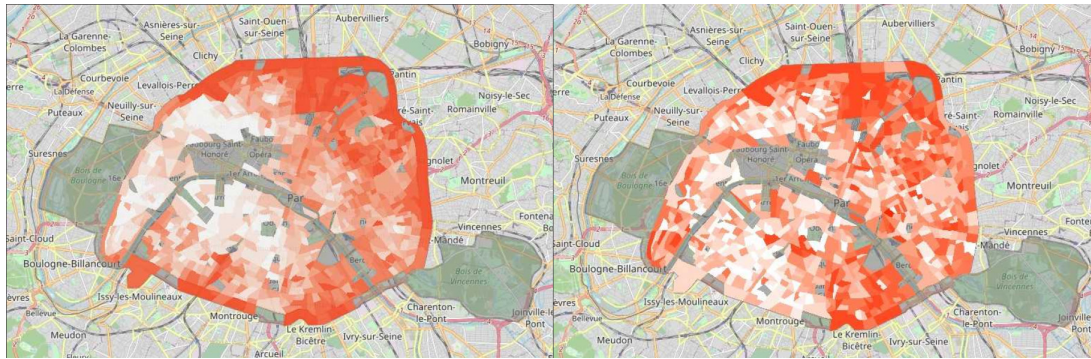


Figure 2 Gradient of the income third decile (left) and COVID-19 hospitalised cases >3 per 1000 inhabitants (right) in Paris residential areas. Non-residential areas are not covered by a red colour gradient.

7.51) and the rate of single-parent families (OR=3.84, 95% CI 2.37 to 6.46).

DISCUSSION

We show that many socioeconomic indicators are strongly associated with areas of high incidence of hospitalised patients with COVID-19 in Paris. Overall, 86% of COVID-19 hospitalised cases occurred within the 45% most deprived areas. This finding is consistent with the idea that COVID-19 occurred as a syndemic—a synergistic pandemic that interacts with and exacerbates a person's existing non-communicable diseases and social conditions.⁵

The subgroup highlighted as over-represented in the high-risk area, that is, the population with unskilled occupations living in deprived areas, is not at high risk for severe COVID-19 because these people are still active, but are likely to have a high rate of infection. Indeed, lower-paid workers, particularly in the service sector (eg, food, cleaning or delivery services), are much more likely to be designated as key workers and thus are required to go to work and rely on public transport to do so. All of these factors increase their exposure to the virus⁵ and virus transmission in their neighbourhood and thus indirectly contributing to an increased hospitalisation rate.

What is already known on this subject

- ▶ Among more than 300 indicators, income was the most associated indicator, with 45% of the most deprived areas accounting for 86% of the COVID-19 high incidence of hospitalised cases areas in Paris.
- ▶ Occupational status, immigration status, unemployment status and living in public low-rent housing were also strongly associated with hospitalisation rate, whereas population density was not associated.

What this study adds

- ▶ Studying a broad range of socioeconomic indicators using census data and hospitalisation data as a readily available and large resource is of particular interest in such a period of hospital overcrowding as it provides real-time indirect information on populations with a high incidence of COVID-19.

The strong association between income level and high-risk areas has to be faced with the low level of inequalities in France compared with the level worldwide.¹⁰ The French healthcare system is ranked among the best in terms of waiting times, results and benefits,¹¹ which renders these results unexpected. Indeed, socioeconomic studies performed in other countries, including the USA and the UK, did not show such strong associations, with an OR of approximately 3 for poverty in Chicago¹² and approximately 2 in England.¹³

Population density in residential areas is high, but the IQR is also high. Therefore, the lack of association with population density is not explained by a homogeneous population density. In Paris, population density is not strongly correlated with population wealth. This notion is in contrast to most other large cities, where areas with high densities correspond to low-income neighbourhoods.^{12 14} This notion might explain why population density was associated with high incidence of COVID-19 in other areas but not in Paris.

The indicators were based on census data, which might differ during the study period because a large exodus from Paris was observed when lockdown started, especially among students.¹⁵ This exodus could explain why the area with a high proportion of the 15–24 years age group was associated with low risk, as this age group was one of the largest groups to leave Paris. Therefore, there was reduced risk of COVID-19 transmission in the neighbourhood of their primary residence.

To conclude, studying a broad range of socioeconomic indicators using census data and hospitalisation data as a readily available and large resource is informative for better resource allocation.

Acknowledgements We would like to acknowledge and thank the EDS AP-HP COVID-19 consortium integrating the AP-HP Health Data Warehouse, as well as all the AP-HP staff and volunteers who contributed to the implementation of the EDS-COVID database and the operating solutions for this database.

Collaborators AP-HP/Universities/Inserm COVID-19 research collaboration: Pierre-Yves Ancel (APHP Paris University Center: Local CDW coordinator), Alain Bauchet (AP-HP Saclay University Beeker Nathanael AP-HP Paris University Center: Data scientist), Vincent Benoit (WIND Department APHP Greater Paris University Hospital: Data engineer), Mélodie Bernaux (Strategy and transformation department, APHP Greater Paris University Hospital: Medical coordination of data analysis), Ali Bellamine (WIND Department APHP Greater Paris University Hospital: Data engineer, data scientist), Romain Bey (WIND Department APHP Greater Paris University Hospital: Data engineer, data scientist, regulatory assessment), Aurélie Bourmaud (APHP Paris University North: Local CDW coordinator), Stéphane Bréant (WIND Department APHP Greater Paris University Hospital: Coordination of clinical research informatics), Fabrice Carrat (APHP Sorbonne University Caucheteux Charlotte Université Paris-Saclay, Inria, CEA: Data integration and analysis), Julien Champ (INRIA Sophia-Antipolis – ZENITH team, LIRMM, Montpellier, France: Data integration and analysis), Sylvie Cormont (WIND Department APHP Greater Paris

University Hospital: Data standardisation), Christel Daniel (WIND Department APHP Greater Paris University Hospital UMRS1142 INSERM: Medical director of data standardisation and clinical research inform), Julien Dubiel (WIND Department APHP Greater Paris University Hospital: Data engineer), Catherine Ducloux (APHP Paris Seine Saint Denis University Hospital: Local CDW coordinator), Loïc Esteve (SED/SIERRA, Inria Centre de Paris: Data engineer, data scientist), Marie Frank (APHP Saclay University: Local CDW coordinator), Nicolas Garcelon (Imagine Institute: Data engineer, data scientist), Alexandre Gramfort (Université Paris-Saclay, Inria, CEA: Data engineer, data scientist), Nicolas Griffon (WIND Department APHP Greater Paris University Hospital, UMRS1142 INSERM: Data standardisation), Olivier Grisel (Université Paris-Saclay, Inria, CEA: Data engineer, data scientist), Martin Guilbaud (WIND Department APHP Greater Paris University Hospital: Data engineer), Claire Hassen-Khodja (Direction of the Clinical Research and Innovation, AP-HP: Medical coordination of data-driven research), François Hemery (APHP Henri Mondor University Hospital: Local CDW coordinator), Martin Hilka (WIND Department APHP Greater Paris University Hospital: Director of Big data platform), Jerome Lambert (APHP Paris University North: Local CDW coordinator), Richard Layese (APHP Henri Mondor University Hospital), Judith Leblanc (Clinical Research Unit, Saint Antoine Hospital, APHP Greater Paris University Hospital: Data scientist), Léo Lebouter (WIND Department APHP Greater Paris University Hospital: Data engineer), Guillaume Lemaître (Université Paris-Saclay, Inria, CEA: Data engineer, data scientist), Damien Leprovost (Clevy.io: Data engineer, data scientist), Ivan Lerner (Department of Biomedical Informatics, HEGP, APHP Greater Paris University Hospital: Data engineer, data scientist), Sallah Kankoe Levi (APHP Paris University North), Aurélien Maire (WIND Department APHP Greater Paris University Hospital: Data engineer), Marie-France Mamzer (President of the AP-HP IRB), Patricia Martel (APHP Saclay University: Data scientist), Arthur Mensch (ENS, PSL University: Data engineer, data scientist), Thomas Moreau (Université Paris-Saclay, Inria, CEA: Data engineer, data scientist), Antoine Neuraz (Department of Biomedical Informatics, HEGP, APHP Greater Paris University Hospital: Data engineer, data scientist), Nina Orlova (WIND Department APHP Greater Paris University Hospital: Data engineer, data scientist), Hélène Ravera (WIND Department APHP Greater Paris University Hospital: Data engineer), Antoine Rozes (APHP Sorbonne University), Elisa Salamanca (WIND Department APHP Greater Paris University Hospital: Director of the Data & Innovation department), Arnaud Sandrin (WIND Department APHP Greater Paris University Hospital: Director of the National Rare Diseases Database), Patricia Serre (WIND Department APHP Greater Paris University Hospital: Data engineer, data standardisation), Xavier Tannier (Sorbonne University: Data engineer, data scientist), Jean-Marc Treluyer (APHP Paris University Center: Local CDW coordinator), Damien Van Gysel (APHP Paris University North: Local CDW coordinator), Gaël Varoquaux (Université Paris-Saclay, Inria, CEA, Montréal Neurological Institute, McGill University: Data engineer, data scientist), Jill Jen Vie (Sequel, Inria Lille: Data engineer, data scientist), Maxime Wack (Department of Biomedical Informatics, HEGP, APHP Greater Paris University Hospital: Data engineer, data scientist), Perceval Wajsburt (Sorbonne University: Data engineer, data scientist), Demian Wassermann (Université Paris-Saclay, Inria, CEA: Data engineer, data scientist), Eric Zapletal (Department of Biomedical Informatics, HEGP, APHP Greater Paris University Hospital: Data engineer), Collégiales of AP-HP: anesthésie-réanimation, médecine intensive réanimation, infectiologie, virologie, nutrition.

Contributors A-SJ designed the study. HC, BR and AB developed the framework for GeoCancer, the pipeline that was reused to integrate the geodata of patients with COVID-19 in this study. HC and BR performed the geolocalisation of patients' residential address and the data aggregation. A-SJ, HC and BR had full access to the aggregated data used for this study and take responsibility for the integrity of the data. A-SJ did the analyses and takes responsibility for the accuracy of the data analysis. AVS extracted the clinical data of patients with COVID-19 and participated in their analysis. A-SJ drafted the paper with the help of BR, AB and SK. Data were collected from Assistance Publique - Hôpitaux de Paris. All authors critically revised the manuscript for important intellectual content and gave final approval for the version to be published.

Funding HC received funding from Canceropole Ile-de-France to develop the geolocalisation method used for this study. A-SJ received funding for this project from the COVID-19 sponsorship (mécénat 'Collecte Crise COVID-19') of AP-HP Centre Université de Paris. A-SJ and AVS received funding for this project from Fondation de France as part of the alliance 'Tous unis contre le virus' (Fondation de France, AP-HP, Institut Pasteur).

Map disclaimer The depiction of boundaries on this map does not imply the expression of any opinion whatsoever on the part of BMJ (or any member of its group) concerning the legal status of any country, territory, jurisdiction or area or of its authorities. This map is provided without any warranty of any kind, either express or implied.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study was approved by the Scientific and Ethical Committee of AP-HP (Assistance Publique - Hôpitaux de Paris) (authorisation number IRB 00011591).

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

This article is made freely available for use in accordance with BMJ's website terms and conditions for the duration of the covid-19 pandemic or until otherwise determined by BMJ. You may use, download and print the article for any lawful, non-commercial purpose (including text and data mining) provided that all copyright notices and trade marks are retained.

ORCID iD

Anne-Sophie Jannot <http://orcid.org/0000-0002-8001-8539>

REFERENCES

- Kaplan GA. What's wrong with social epidemiology, and how can we make it better? *Epidemiol Rev* 2004;26:124–35.
- Patel AP, Paranjpe MD, Kathiresan NP. Socioeconomic deprivation, and hospitalization for COVID-19 in English participants of a National Biobank [Internet]. *Epidemiology*, 2020. Available: <http://medrxiv.org/lookup/doi/10.1101/2020.04.27.20082107>
- Bilal U, Barber S, Diez-Roux AV. Spatial Inequities in COVID-19 outcomes in Three US Cities [Internet]. *Epidemiology*, 2020. Available: <http://medrxiv.org/lookup/doi/10.1101/2020.05.01.20087833>
- Das A, Ghosh S, Das K, et al. Modeling the effect of area deprivation on COVID-19 incidences: a study of Chennai megacity, India. *Public Health* 2020;185:266–9.
- Bambra C, Riordan R, Ford J, et al. The COVID-19 pandemic and health inequalities. *J Epidemiol Community Health* 2020;74:964–8.
- Buniello A, MacArthur JAL, Cerezo M, et al. The NHGRI-EBI GWAS catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 2019;47:D1005–12.
- Denny JC, Ritchie MD, Basford MA, et al. TheWAS: demonstrating the feasibility of a genome-wide scan to discover gene-disease associations. *Bioinformatics* 2010;26:1205–10.
- Neuraz A, Lerner I, Digan W, et al. Natural language processing for rapid response to emergent diseases: case study of calcium channel blockers and hypertension in the COVID-19 pandemic. *J Med Internet Res* 2020;22:e20773.
- Documentation sur les bases de données infracommunales à l'IRIS. Available: <https://www.insee.fr/fr/information/2383389>
- Income inequality. Available: <https://data.oecd.org/inequality/income-inequality.htm>
- Euro health consumer index 2018. Available: <https://healthpowerhouse.com/publications/>
- Maroko AR, Nash D, Pavilonis BT. COVID-19 and inequity: a comparative spatial analysis of new York City and Chicago hot spots. *J Urban Health* 2020;97:461–70.
- Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 2020;584:430–6.
- Friesen J, Pelz PF. COVID-19 and slums: a pandemic highlights gaps in knowledge about urban poverty. *JMIR Public Health Surveill* 2020;6:e19578.
- Retour partiel des mouvements de population avec Le déconfinement. Available: <https://www.insee.fr/fr/statistiques/4635407>