# Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards

Saman Khalatbari-Soltani, Robert Cumming, Cyrille Delpierre, Michelle Kelly-Irving

Citation: Khalatbari-Soltani S, Cumming RC, Delpierre C*, et al* Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards *J Epidemiol Community Health* 2020;**74:**620-623. DOI: http://dx.doi.org/10.1136/jech-2020-214297

# Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards

Saman Khalatbari-Soltani[1,2], Robert C Cumming[3,4], Cyrille Delpierre[5,6], Michelle Kelly-Irving[5,6,7]

Affiliations

1. Faculty of Medicine and Health, The University of Sydney School of Public Health, Sydney, Australia
2. ARC Centre of Excellence in Population Ageing Research (CEPAR), The University of Sydney, Sydney, Australia
3. The University of Sydney School of Public Health, Sydney, Australia
4. ARC Centre of Excellence in Population Ageing Research (CEPAR), The University of Sydney Faculty of Medicine and Health, Sydney, Australia
5. LEASP, Université Toulouse III Paul Sabatier, Toulouse, France
6. Umr 1027, Inserm, Toulouse, France
7. IFERISS, Université de Toulouse, Toulouse, France

Correspondence to Dr Saman Khalatbari-Soltani, Faculty of Medicine and Health, The University of Sydney School of Public Health, Sydney, New South Wales 2006, Australia; saman.khalatbarisoltani@sydney.edu.au

# Abstract

Disadvantaged socioeconomic position (SEP) is widely associated with disease and mortality, and there is no reason to think this will not be the case for the newly emerged coronavirus disease 2019 (COVID-19) that has reached a pandemic level. Individuals with a more disadvantaged SEP are more likely to be affected by most of the known risk factors of COVID-19. SEP has been previously established as a potential determinant of infectious diseases in general. We hypothesise that SEP plays an important role in the COVID-19 pandemic either directly or indirectly via occupation, living conditions, health-related behaviours, presence of comorbidities and immune functioning. However, the influence of socioeconomic factors on COVID-19 transmission, severity and outcomes is not yet known and is subject to scrutiny and investigation. Here we briefly review the extent to which SEP has been considered as one of the potential risk factors of COVID-19. From 29 eligible studies that reported the characteristics of patients with COVID-19 and their potential risk factors, only one study reported the occupational position of patients with mild or severe disease. This brief overview of the literature highlights that important socioeconomic characteristics are being overlooked when data are collected. As COVID-19 spreads worldwide, it is crucial to collect and report data on socioeconomic determinants as well as race/ethnicity to identify high-risk populations. A systematic recording of socioeconomic characteristics of patients with COVID-19 will be beneficial to identify most vulnerable groups, to identify how SEP relates to COVID-19 and to develop equitable public health prevention measures, guidelines and interventions.

The socioeconomic gradient in health is ubiquitous,[1] and has been described across pathologies, in life expectancy and mortality. It is characterised by a social patterning of health, where the most socially disadvantaged are more likely to be exposed, to get sick and to die compared to their advantaged counterparts. The most commonly used individual measures of socioeconomic position (SEP) are occupation, income and education,[2] each one measuring different components and aspects of the social environment. All have well-established associations with numerous health outcomes and mortality, through various pathways such as behaviours, chemical and physical exposures or chronic stress exposures.[3] Moreover, evidence has suggested a link between disadvantaged SEP and increased risk of infectious disease in general[4 5] and immune response. People with disadvantaged SEP should be considered as high-risk populations at the time of any infectious disease outbreak[6] since their social context could affect the occurrence and severity of an infectious disease via several pathways.[4]

The outbreak of coronavirus disease 2019 (COVID-19) caused by the novel coronavirus SARS-CoV-2 was first reported on 31 December 2019, in Wuhan, China, and has since spread to 184 countries and regions globally. As of 9 April 2020, there were more than one million and half confirmed cases (n=1 511 104) and 88 338 deaths. So far, advanced age, being a man and having chronic conditions, such as obesity, diabetes, respiratory disease, kidney disease and cardiovascular diseases (CVD), have been linked with more severe COVID-19 symptoms often leading to the development of acute respiratory distress syndrome (ARDS) and progression from ARDS to death.[2 7] Other potential factors such as race/ethnicity and socioeconomic factors may also play an important role in the COVID-19 epidemic. While the economic impact of the COVID-19 outbreak and the resulting measures has raised as an important issue, we should not neglect individuals' own SEP and its effect on COVID-19 incidence, transmission, severity and mortality.

Based on our previous knowledge of social and structural determinants of health, there is no reason to think that these factors are any less important in affecting who is exposed to COVID-19, who gets sick, who is likely to need intensive care and who will die from it. Without collecting data on these social factors and making them available, we are blind to how these factors affect the epidemic and how they may be used to inform public health prevention measures. Here we attempt to briefly assess the extent to which the current data available on COVID-19 take into account socioeconomic factors. We then make the case for why social factors are likely to be involved, both directly and indirectly, in the epidemiology of COVID-19, and therefore why they need to be collected.

We identified published studies through a rapid review of PubMed from inception to 03 April 2020, with the following search terms: ('COVID-19', 'SARS-CoV-2') and ('characteristic', 'risk factors', 'socio', 'socioeconomic', 'occupation', 'education', 'income', 'wealth'). We also screened the reference lists of the eligible publications. We included studies that were written in

English and reported descriptive characteristics of COVID-19 cases or reported risk factors of COVID-19 incidence, severity and mortality. The search identified 161 articles; after title, abstract and full-text screening, 16 publications were eligible for inclusion. We identified another 13 eligible publications through reference screening. Most of the studies were from China (n=26; mostly in Wuhan), two from Singapore, one from Europe and one from the Diamond Princess Cruise, Japan. All studies included cases of COVID-19, and the sample sizes range from 18 to 72 314. Of note, in the rush to publish COVID-19 papers, it is highly possible that the same sample of patients have been reported in different articles.[8] While all included studies reported data on age, gender and comorbidities, only one study reported any indicator of SEP. In this study, Shi *et al*, using data on occupational position from 484 COVID-19 patients in Zhejiang Province of China, reported that severe cases were more likely to be agricultural workers and less likely to be self-employed than mild cases.[9] To our knowledge, no study thus far has reported data on other SEP indicators such as educational level, income and housing conditions.

One of the reasons for this is that data on individual-level SEP are not being collected. This could be due to the suggested World Health Organization standard COVID-19 case report form[1] which only asks for each patient's age, sex/gender, place where the case was diagnosed and usual place of residency. This could also be due to the fact that social measures are not considered as data of clinical interest by most clinicians. The consequence is that these data are absent from medical records, limiting the possibility of studying the evolution of diseases with regards to these determinants. However, it is critical to also consider socioeconomic factors at each phase of the epidemic to effectively interrupt human-to-human transmission chains, prevent further spread through appropriate equitable interventions, as well as to identify and better treat individuals who have greater susceptibility to becoming severe or even critically ill upon infection. The reasons why individual SEP must become 'clinical' data in the same way as age are numerous:

- A person's occupation may expose them to risk by the nature of their job. Work involving constant human contact, interaction with others or caring for people means that risk of infection spread through droplets of aerosol is higher.[10] With regards to COVID-19, occupation is likely to be a direct determinant of infection and an indirect determinant of disease severity and mortality through the relationship between occupational social class and comorbidities. For example, the direct impact of occupation on the COVID-19 incidence could be seen among workers such as retail staff, cleaners, teachers, healthcare workers or crew on board cruise ships.[11] People in disadvantaged SEP categories are more likely to be exposed to job stress including high job strain, burnout and unemployment, which may lead to disrupted immune and inflammatory system responses[12 13] as well as an increased risk for CVD.[14] So far,

both reduced immune function and the presence of CVD are known risk factors of COVID-19 severity.

- Low income might affect living conditions in many ways, such as residence in more deprived neighbourhoods and housing conditions, especially cramped or overcrowded housing, which has been associated with an increased risk of infection from numerous pathogens, such as tuberculosis[15] *Helicobacter pylori* [16] or Epstein–Barr virus.[5]
- A lower education level is indirectly associated with a number of factors that may increase the risk of developing severe forms of COVID-19, such as increased prevalence of smoking and poor nutrition, which could suppress the immune system. A recent systematic review of five retrospective or prospective studies reported that smoking is most likely associated with adverse outcomes of COVID-19.[17] Furthermore, known comorbidities for COVID-19, such as CVD, diabetes, respiratory disease and kidney disease are more prevalent among socially disadvantaged individuals than in populations with higher SEP.[18] Health literacy, which is associated with education, may play an important role in COVID-19 incidence and severity. Effective public health communication to act appropriately during an infectious disease outbreak is contingent upon people being able to access and understand the information. Individuals with limited health literacy may be more easily misguided by incorrect sources of information.[6] Furthermore, we also know that screening measures are often overlooked by disadvantaged populations with low health literacy levels, which may lead to delay in seeking care late when their sickness is worse.[19] Thus, a whole-of-society approach must include appropriate language and accessible healthcare strategies.

To be able to understand the complex and interrelated influence of socioeconomic factors on COVID-19 transmission, incidence and its health outcomes, data sources with comprehensive socioeconomic measures are needed. Some might argue that we can link people's addresses or postcodes to area-based SEP through geolocalisation, which may offer some insight into the likelihood of exposure to certain health risks, including pollution or public transport. Indeed, these variables are often used as proxies for individual SEP; however, they are not an accurate reflection of individual circumstances, could underestimate the extent of social inequalities compared to individual social measures[20] and are best used in parallel with individual-level variables to reflect geographical or aggregate-level exposures.

**Other dimensions that interact with SEP need also to be considered**

Race/ethnicity

Ethnic minorities including indigenous peoples are subject to systematic and structural discrimination, which leads to them being more socioeconomically disadvantaged and less likely to be able to seek healthcare when they need it.[21] Recent reports from health authorities and journalists suggest that COVID-19 hospitalisation and mortality rates may be higher in geographical areas or

neighbourhoods which are more socially disadvantaged or where ethnic minorities are more likely to live. The poorest department in mainland France, Seine St. Denis, with a high proportion of people from ethnic minorities living there, reported a 63% increase in mortality during the week of the 21st of March, a bigger increase than any other French department. Reports emerging from the United States also suggest that African Americans and possibly people from other ethnic minorities are over-represented among the COVID-19 deaths. These are anecdotal and unscientific reports for the moment; however, they deserve to be taken seriously and met with scientific methods and analyses to examine them thoroughly. Furthermore, within each country, attention should be paid to the specific cultural practices and needs of different populations; they may need support in sending appropriate prevention messages to their community about festivals, rituals and habits. Thus, as previously suggested by the Sendai Framework, governments should engage with migrants, people from ethnic minorities and indigenous peoples in the design and implementation of policies during a pandemic.[22]

**Gender and intersectionality**

Gender is another important aspect that deserves serious attention in the COVID-19 pandemic,[23] especially as it intersects with occupation, education and race/ethnicity. Women are more likely to have jobs or roles where they are in contact with others, as teachers, nurses, carers or retail workers. When schools closed across different countries, women are most likely to have had to take on childcare and homeschooling responsibilities. On the other hand, men are more likely than women to present with severe forms of COVID-19 and have a higher mortality rate.[7 24] The reasons for this remain to be understood and may vary between contexts; however, men tend to have a higher prevalence of the main COVID-19 risk factors. When it comes to the mitigation policies put in place to curtail the spread of disease, such as confinement or quarantine, these intersecting factors including gender, class and race/ethnicity will affect how people cope with the social, economic and psychological consequences of such measures.[25 26]

**National and local contexts**

The impact of socioeconomic determinants on the COVID-19 outbreak could vary in high-income-countries versus low- and middle-income countries (LMICs). As observed across the world, once COVID-19 cases appear initially from abroad, the community spread of disease will depend upon the specific local infrastructures and social inequalities in each context. Current public health measures of social distancing including lockdowns, quarantine and self-isolation cannot be implemented or may carry consequences in LMICs. In low-income settings where absolute poverty is a major problem, access to basic needs such as water, sanitation and food will have a big impact on how easily people can practise any physical distancing measures. Moreover, in LMICs where people are more likely to live in overcrowded households or neighbourhoods, it would be impossible to isolate older adults or vulnerable people. In many parts of the world, health is not free at the point of need, and

healthcare systems, which are stretched at the best of times, will become quickly saturated. Furthermore, while children appear to be less severely affected by COVID-19, they may be more vulnerable to the social, economic and psychological consequences of the disease and ensuing public health containment measures, especially in low-income settings.[27]

# CONCLUSION

In order to identify groups who are most likely to have poor outcomes, high-quality data on socioeconomic factors are urgently needed, which will have important implications in the development of public health measures. Social measures should be considered as clinical variables, in the same way as age or gender, and should therefore be routinely recorded in medical records. Pandemic recommendations and guidelines provided by international and national agencies need to recognise the collective contribution of the social determinants of health and their intersectionality to pandemic risk mitigation.

**Key messages**

- Prior evidence on the social and socioeconomic determinants of disease suggests that Covid-19 is likely to have a socially patterned population distribution.
- Social and socioeconomic data are not being collected and made available to understand and address Covid-19 health inequalities.
- Pandemic recommendations and guidelines need to reflect the contribution of the social determinants of health and their intersectional nature.

REFERENCES


1. ↲ Hertzman C . Putting the concept of biological embedding in historical perspective. Proc Natl Acad Sci U S A 2012;109:17160–7. doi: 10.1073/pnas.1202203109 Abstract/FREE Full TextGoogle Scholar
2. ↲ Krieger N . A glossary for social epidemiology. J Epidemiol Community Health 2001;55:693–700. doi: 10.1136/jech.55.10.693 FREE Full TextGoogle Scholar
3. ↲ Kelly-Irving M , Delpierre C . The embodiment dynamic over the life course: a case for examining cancer aetiology. In: Meloni M, Cromby J, Fitzgerald D, et al, eds. The Palgrave Handbook of Biology and Society . London: Palgrave Macmillan, 2017: 519–40.Google Scholar
4. ↲ Oestergaard LB , Schmiegelow MD , Bruun NE , et al. The associations between socioeconomic status and risk of Staphylococcus aureus bacteremia and subsequent endocarditis - a Danish nationwide cohort study. BMC Infect Dis 2017;17:589–589. doi: 10.1186/s12879-017-2691-3 Google Scholar
5. ↲ Gares V , Panico L , Castagne R , et al. The role of the early social environment on Epstein Barr virus infection: a prospective observational design using the Millennium Cohort Study. Epidemiol Infect 2017;145:3405–12. doi: 10.1017/S0950268817002515 Google Scholar
6. ↲ O'Sullivan TL , Phillips KP . From SARS to pandemic influenza: the framing of high-risk populations. Nat Hazards 2019;98:103–17. doi: 10.1007/s11069-019-03584-6 Google Scholar
7. ↲ Wu C , Chen X , Cai Y , et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. JAMA Intern Med Published 2020. doi: 10.1001/jamainternmed.2020.0994 Google Scholar
8. ↲ Bauchner H , Golub RM , Zylke J . Editorial concern: possible reporting of the same patients with COVID-19 in different reports. JAMA 2020;323:1256. doi: 10.1001/jama.2020.3980 Google Scholar
9. ↲ Shi Y , Yu X , Zhao H , et al. Host susceptibility to severe COVID-19 and establishment of a host risk score: findings of 487 cases outside Wuhan. Crit Care 2020;24:108. doi: 10.1186/s13054-020-2833-7 CrossRefPubMedGoogle Scholar
10. ↲ Rule AM , Apau O , Ahrenholz SH , et al. Healthcare personnel exposure in an emergency department during influenza season. PLoS One 2018;13:e0203223. doi: 10.1371/journal.pone.0203223 Google Scholar
11. ↲ Koh D . Occupational risks for COVID-19 infection. Occup Med 2020;70:3–5. doi: 10.1093/occmed/kqaa036 Google Scholar
12. ↲ Nakata A , Psychosocial job stress and immunity: a systematic review. Yan Q , ed. Psychoneuroimmunology . Totowa, NJ: Humana Press, 2012: 39–75. doi: 10.1007/978-1-62703-071-7_3 Google Scholar

13. ↲ Berger E , Castagne R , Chadeau-Hyam M , et al. Multi-cohort study identifies social determinants of systemic inflammation over the life course. Nat Commun 2019;10. doi: 10.1038/s41467-019-08732-x Google Scholar
14. ↲ Kivimäki M , Kawachi I . Work stress as a risk factor for cardiovascular disease. Curr Cardiol Rep 2015;17:630–630. doi: 10.1007/s11886-015-0630-8 PubMedGoogle Scholar
15. ↲ Gupta D , Das K , Balamughesh T , et al. Role of socio-economic factors in tuberculosis prevalence. Indian J Tuberc 2004;51:27–31.Google Scholar
16. ↲ Webb PM , Knight T , Greaves S , et al. Relation between infection with Helicobacter pylori and living conditions in childhood: evidence for person to person transmission in early life. BMJ 1994;308:750–3. doi: 10.1136/bmj.308.6931.750 Abstract/FREE Full TextGoogle Scholar
17. ↲ Vardavas C , Nikitara K . COVID-19 and smoking: a systematic review of the evidence. Tob Induc Dis 2020;18. doi: 10.18332/tid/119324 Google Scholar
18. ↲ Adler NE , Newman K . Socioeconomic disparities in health: pathways and policies. Health Aff (Millwood) 2002;21:60–76. doi: 10.1377/hlthaff.21.2.60 Abstract/FREE Full TextGoogle Scholar
19. ↲ Berkman ND , Sheridan SL , Donahue KE , et al. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011;155:97. doi: 10.7326/0003-4819-155-2-201107190-00005 CrossRefPubMedWeb of ScienceGoogle Scholar
20. ↲ Lamy S , Molinié F , Daubisse-Marliac L , et al. Using ecological socioeconomic position (SEP) measures to deal with sample bias introduced by incomplete individual-level measures: inequalities in breast cancer stage at diagnosis as an example. BMC Public Health 2019;19:857. doi: 10.1186/s12889-019-7220-4 Google Scholar
21. ↲ Krieger N . Discrimination and health inequities. Int J Health Serv 2014;44:643–710. doi: 10.2190/HS.44.4.b CrossRefPubMedGoogle Scholar
22. ↲ United nations office for disaster risk reduction. Sendai framework for disaster risk reduction 2015-2030 . Geneva: United Nations Office for Disaster Risk Reduction, 2015: Available https://www.unisdr.org/files/43291_sendaiframeworkfordrren.pdf Google Scholar
23. ↲ Wenham C , Smith J , Morgan R . COVID-19: the gendered impacts of the outbreak. Lancet 2020;395:846–8. doi: 10.1016/S0140-6736(20)30526-2 PubMedGoogle Scholar
24. ↲ Zhou F , Yu T , Du R , et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020;395:1054–62. doi: 10.1016/S0140-6736(20)30566-3 CrossRefPubMedGoogle Scholar

25. ↲ Littrell J . The mind-body connection. Soc Work Health Care 2008;46:17–37. doi: 10.1300/J010v46n04_02 CrossRefPubMedWeb of ScienceGoogle Scholar
26. ↲ Burton DC , Flannery B , Bennett NM , et al. Socioeconomic and racial/ethnic disparities in the incidence of bacteremic pneumonia among US adults. Am J Public Health 2010;100:1904–11. doi: 10.2105/AJPH.2009.181313 CrossRefPubMedWeb of ScienceGoogle Scholar
27. ↲ Gordon D , Nandy S . The extent, nature and distribution of child poverty in India. Indian J Hum Dev 2016;10:64–84. doi: 10.1177/0973703016654561 Google Scholar

## Footnotes

- https://apps.who.int/iris/bitstream/handle/10665/331234/WHO-2019-nCoV-SurveillanceCRF-2020.2-eng.pdf.

- Contributors SKS and MKI drafted all versions of the manuscript, SKS carried out the literature review, and RC and CD drafted the manuscript and provided edits.

## Funding

SKS is supported by the Australian Research Council Centre of Excellence in Population Ageing Research (Project number CE170100005). The funder had no role in study design, data collection, decision to publish or preparation of the manuscript. This work was also part of the EPIDEMIC project funded by the Agence National de Recherche Flash-Covid grant and Région Occitanie.