

RESEARCH

Open Access



Comparison of geographical and individual deprivation index to assess the risk of Sars-CoV-2 infection and disease severity: a retrospective cohort study

Sara Mazzilli^{1,2*}, Gianluca Papatatto^{2,3}, Antonio Chieti⁴, Anna Maria Nannavecchia⁴, Lucia Bisceglia⁴, Pier Luigi Lopalco⁵, Domenico Martinelli⁶ and Lara Tavoschi²

Abstract

Background It has been shown that COVID-19 affects people at socioeconomic disadvantage more strongly. Previous studies investigating the association between geographical deprivation and COVID-19 outcomes in Italy reported no differences in case-hospitalisation and case-fatality. The objective of this research was to compare the usefulness of the geographic and individual deprivation index (DI) in assessing the associations between individuals' deprivation and risk of Sars-CoV-2 infection and disease severity in the Apulia region from February to December 2020.

Methods This was a retrospective cohort study. Participants included individuals tested for SARS-CoV-2 infection during the study period. The individual DI was calculated employing polychoric principal component analysis on four census variables. Multilevel logistic models were used to test associations between COVID-19 outcomes and individual DI, geographical DI, and their interaction.

Results In the study period, 139,807 individuals were tested for COVID-19 and 56,475 (43.5%) tested positive. Among those positive, 7902 (14.0%) have been hospitalised and 2215 (4.2%) died. During the first epidemic wave, according to the analysis done with the individual DI, there was a significant inversely proportional trend between the DI and the risk of testing positive. No associations were found between COVID-19 outcomes and geographic DI. During the second wave, associations were found between COVID-19 outcomes and individual DI. No associations were found between the geographic DI and the risk of hospitalisation and death. During both waves, there were no association between COVID-19 outcomes and the interaction between individual and geographical DI.

Conclusions Evidence from this study shows that COVID-19 pandemic has been experienced unequally with a greater burden among the most disadvantaged communities. The results of this study remind us to be cautious about using geographical DI as a proxy of individual social disadvantage because may lead to inaccurate assessments. The geographical DI is often used due to a lack of individual data. However, on the determinants of health and health inequalities, monitoring has to have a central focus. Health inequalities monitoring provides evidence on who is being left behind and informs equity-oriented policies, programmes and practices. Future research and data collection should focus on improving surveillance systems by integrating individual measures of inequalities into national health information systems.

*Correspondence:

Sara Mazzilli

sara.mazzilli@sns.it

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords Socioeconomic disadvantage, Deprivation index, Sars-CoV-2, COVID-19, Apulia, Italy

Introduction

The newly emerged virus SARS-CoV-2 was initially reported in China in December 2019 [1]. On February 20, 2020, the first major COVID-19 outbreak in Europe was detected in the Lombardy region, Italy [2]. On March 11, 2020, WHO declared the SARS-CoV-2 outbreak a pandemic [3]. As of August 30, 2023, the WHO reports a total of 770,085,713 confirmed cases and 6,956,173 confirmed deaths worldwide [4]; while the Italian NHS reported 26,175,146 confirmed cases, and 190,644 deaths [5]. During the first pandemic year, the temporal course of the epidemic in Italy was characterized by 3 distinct phases: the first epidemic wave from March to June 2020, followed by a summer period with a relatively low incidence, and a second wave that started in September and peaked in November 2020 [4].

Evidence shows that males, aged over 65 and smoking patients might face a greater risk of developing more critical or lethal conditions if infected with SARS-CoV2. Comorbidities, such as hypertension, diabetes, cardiovascular disease or respiratory diseases, could also greatly affect the prognosis of COVID-19 patients [6]. A study carried out in a large community cohort has also shown associations between adverse lifestyles and a higher risk of COVID-19 [7]. However, it is now well known that lifestyle plays a mediating role in the relationship between socioeconomic position (SEP) and health [8]. Literature reports that the wide inequalities seen in infection, hospitalization and mortality rates between population groups were mostly driven by social factors overlaid on biological risks [9, 10]. Multiple mechanisms explain the increased impact of COVID-19 among people with a higher level of socio-economic disadvantage, but in summary, unfavourable social determinant of health (SDH) and associated higher rates of chronic disease [11] increased their risk of poor outcomes from COVID-19 and poorer access to health services for treatment and vaccination [12]. While existing predisposing susceptibility to COVID-19 is a product of pre-existing SDH, there is growing evidence that the ability of disadvantaged groups to adhere to public health and social measures that reduce viral transmission and to deal with the aftermath of the pandemic, is also negatively affected by unfavourable SDH [13]. Difficult living and working conditions make adherence to preventive measures more difficult for disadvantaged populations, thus increasing their exposure to the risk of infection [14]. For instance, most deprived individuals usually carry out manual labour professions or work in the informal sector and

therefore, resulting in limited opportunities for working from home. Health and socioeconomic inequality mutually influence each other by triggering and feeding vicious circles [12]. The economic fallouts from the pandemic are hitting disadvantaged population groups harder. The pandemic itself has accentuated these already existing social and health inequalities, widening the gap among individuals with different SEP. This is a crucial point for the present but also for the future. Assessment and mitigation of SDH cannot be neglected in a pandemic response and prevention programme.

The available Italian deprivation index (DI) is a multidimensional measure of the disadvantage in the ownership of both social and material resources among residents in each census sections, which are comparable to neighbourhoods as described elsewhere [15, 16]. Mateo-Urdiales et al. investigated the association between deprivation and COVID-19 outcomes in Italy during pre-lockdown, lockdown and post-lockdown periods using the Italian DI as a contextual measure of deprivation. No differences in case-hospitalisation and case-fatality according to deprivation were observed [17]. Similarly, in a study on the city of Barcelona, the increase in hospitalisation and mortality rates was not significant among people living in areas characterised by higher geographical DI [18]. However, the use of geographical DI measures as a proxy for the level of individual social disadvantage is subject to potential ecological bias that can arise from attributing a collective measure to an individual.

The overall objective of this research was to compare the geographic and individual DI in assessing the associations between individuals' SEP and risk of Sars-CoV-2 infection and disease severity in the Apulia region from February to December 2020.

Methods

Study setting

The Italian 'Servizio Sanitario Nazionale' (SSN) was introduced in 1978 to ensure that healthcare is accessible to all Italian citizens without socio-economic barriers, according to a principle of horizontal equity [19]. The system is organized into three levels: national, regional, and local. The national level is responsible for establishing the general objectives and fundamental principles of the NHS. The nineteen regions and two autonomous provinces (R&AP) are then responsible for organizing and delivering health care [20]. In this scenario, through ministerial decrees, the Ministry of Health has taken the lead in the fight against the COVID-19 epidemic. Then,

the R&AP were in charge of organizing and implementing the monitoring and prevention strategy at the local level based on national guidelines. This study is carried out in the Apulia region (Dimension: 19.540 km²), which is a region in southern Italy as shown in red in Additional file 1: Figure S1. As of December 31, 2019, the population of the Apulia region was 3,953,305. The region is administratively divided into six provinces, each with its own LHA. In 2019, 51.4% of the population in Apulia were females, and 3.4% were foreigners. The average age of the population was 45.4 years, with individuals aged over 65 accounting for 23.1% of the entire population. Additionally, 40.4% of the population had at least one chronic disease, with the most common conditions being hypertension (18.8%), arthrosis and arthritis (17.0%), allergic diseases (11.8%), and osteoporosis (9.4%) [21]. In Apulia region, the hospital network comprises 33 public and 26 private institutions, providing a total of 11,565 beds for ordinary inpatient care, equating to 2.93 beds per 1000 people.

Pandemic stages and preventive measures

In Italy, after the detection of the first locally acquired Covid-19 case in Lombardy on February 20, 2020, the number of cases increased greatly in the following weeks, although unevenly among regions, forcing the government to adopt unprecedented restrictive measures. In particular, during the 1st year of the pandemic, the following phases can be defined: 1. first comprehensive national lockdown from 9 March to 3 May (closure of schools and most workplaces, and the implementation of quarantines, border closings, and restriction on public gatherings) 2., [22] gradual reopening phase from 4 May to 14 June (restrictions were gradually rolled back) 3., [23] Few restrictions from 15 June to 7 October (Covid-19 incidence remained low) 4., [24] new restrictions from 8 October to 5 November [25], and 5. lockdowns on a regional basis from 6 November until the end of 2020 (closure of regional borders) [26]. During the first epidemic wave, the number of diagnostic tests (Polymerase Chain reaction—PCR) available was limited, while availability increased considerably during the second wave [27].

Study design and data sources

This is a retrospective cohort study conducted by analysing and merging the following electronic health records from the Apulia regional health-care information systems: (i) laboratory registry of individuals tested for Sars-CoV-2 infection; (ii) registry of COVID-19 confirmed cases; (iii) healthcare workers' registry; (iv) 2011 census dataset. Figure 1. explains how the final dataset was obtained. In the study, a COVID-19 confirmed case

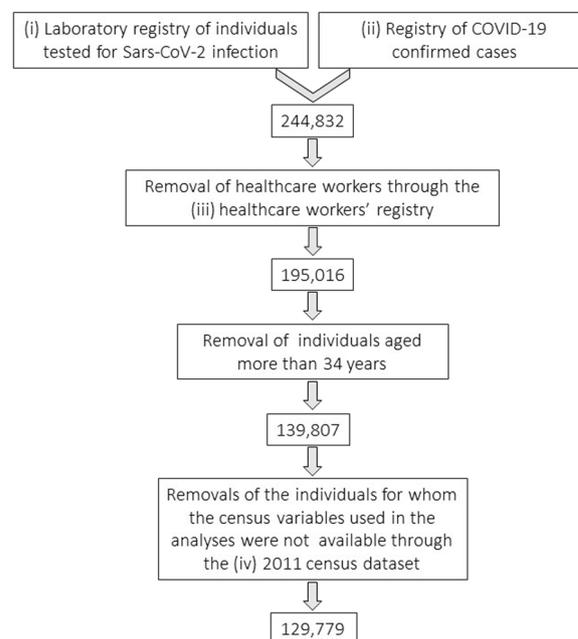


Fig. 1 Data linkage to obtain the study cohort

was defined as an individual who resulted positive to a PCR test. The study period goes from 20 February to 31 December 2020. The cohort in this study includes all the individuals tested for SARS-CoV-2 infection during the study period in the Apulia region, in which a total of 3.95 million inhabitants were registered as of December 31, 2019 [28]. Since the census data date back to 2011, we decided to exclude persons under 35 years of age, for whom part of the census data (i.e. education) may have changed. Individuals tested from 20 February 2020 to 31 May 2020 were included in the first epidemic wave, while those tested from 16 September 2020 to 31 December 2020 were included in the second epidemic wave [4].

The laboratory registry tracked: identification code; age; sex; province of residence; date of the test; PCR test result. The COVID-19 confirmed cases registry reported the following information: identification code, COVID-19 related hospitalization; admission to the intensive care unit (ICU) due to COVID-19 and death due to COVID-19. From the census variables available, the following were selected to perform this study: identification code; education; the number of family members; citizenship; family type; and employment status.

Charlson comorbidities index, extensively described previously [29], and the geographical DI [15, 16] were added to the characteristics of individuals in the final dataset. The Italian geographical DI is the sum of the z score of five simple indicators: x_1 : % of population among 15 and 60 years old with education equal to or less than

elementary school; x_2 : % of the active population unemployed or seeking their first job; x_3 : % of housing occupied for rent; x_4 : % of single-parent families (and consisting of a single-family unit) with children under 18 years old; x_5 : population density (occupants per 100 m²).

Through the use of the healthcare workers' registries, healthcare workers were removed from the dataset as their risk of acquisition of infection was strongly related to their professional exposure to the virus.

The merging of the data information system has been carried out by the regional health agency through a numeric identification code. All the health data used in the study were anonymous.

Definition of the individual deprivation index through the census data

For the definition of the individual DI, polychoric principal component analysis (PCA) was employed, and four census variables were used: citizenship, family type, employment status, and education. Table 1 reports the value attributed to each individual variable category.

Statistical analysis

Generalized linear logistic regression models (GLM) were used to test associations between COVID-19 outcomes (tested positive, being hospitalised due to COVID-19, being admitted in ICU due to COVID-19 and death due to COVID-19) and geographical or individual DI. Age, sex and Charlson comorbidity index were included in the model as covariates. Multilevel logistic models were used to test associations between COVID-19 outcomes and PCA individual DI, geographical DI, and their interaction. Census sections were used as clusters for the model, each of which was assigned a geographical DI. Age, sex and Charlson comorbidity index were included in the model as covariates. To build the model, we followed the following steps also explained in Additional file 1: Figure S2: Preliminary step: Data preparation

(centring of variables); Step1: Construction of an empty model to assess the variation of log-odds from cluster to cluster. This model did not contain the covariates. Step2: Assessment of the variation of lower-level effects from cluster to cluster. In order to perform the assessment, first, a constrained intermediate model was run, secondly, an augmented intermediate model was run and finally both were compared by performing a likelihood-ratio test. Both models contained the covariates, geographical and PCA individual DI. The augmented intermediate included also the residual term associated with the PCA individual DI, thereby estimating the random slope. Step3: Construction of a final model adding the cross-level interaction. The methodology is described in detail in the work of Morselli et al. [30]. When the variation of log-odds from cluster to cluster was not significant, the results of one-level regression analyses were considered.

The statistical analysis has been carried out separately for the first epidemic wave and the second epidemic wave.

R (version 4.2.0) was used for all statistical analysis and a p-value of 0.05 was applied for testing statistical significance.

Results

In the study period 129,779 individuals aged more than 34 (49.1% female and 50.9% male) were tested for COVID-19: 7,160 (5.6%) during the first wave, 20,523 (15.9%) during the summer and 102,096 (78.4%) during the second wave (Additional file 1: Figure S3). In our study population the median age was 60. 24,466 (18.9%) individuals lived in a census area whose geographical DI was 1, 23,594 (18.2%) in a census area whose geographical DI was 2, 25,508 (19.6%) in a census area whose geographical DI was 3, 27,676 (21.3%) in a census area whose geographical DI was 4, 28,497 (20.5%) in a census area whose geographical DI was 5. Table 2 describes study population characteristics by epidemic wave.

Table 1 Census variable used to define the individual DI

Citizenship	Family Type	Employment Status	Education
1—Italian	1—Couple without children	1—Employed	1—Academic Diploma 2nd Level/Academy of Fine Arts Diploma/Master degree
2—Foreigner	2—Single-person households	2—Recipient(s) of one or more pensions due to previous employment or investment income	2—Academic 1st Level Diploma/Bachelor degree
	3—Couple with children	3—Students /Housewives/In other status	3—High school graduation
	4—Single parent	4—First-time job seekers	4—Middle school graduation
	5—Households with 2 or more families	5—Unemployed	5—Elementary school diploma
			6—Illiterates/Literate without educational qualification

Table 2 Description of the study population by wave of epidemics. Proportion calculated only on complete data

	1st Wave						2nd Wave													
	Negative		Positive		Hospitalization		ICU		Death		Negative		Positive		Hospitalization		ICU		Death	
	N %	N %	N %	N %	N %*	N %*	N %	N %	N %	N %*	N %	N %	N %	N %	N %	N %	N %	N %	N %	N %*
Sex																				
Male	2,279 (63.3%)	1,322 (36.7%)	750 (56.7%)	165 (12.5%)	273 (20.7%)	24,075 (48.7%)	25,356 (51.3%)	3,753 (14.8%)	718 (2.8%)	1,027 (4.1%)										
Female	2,200 (63.6%)	1,259 (36.4%)	575 (45.7%)	69 (5.5%)	223 (17.7%)	25,197 (47.9%)	27,428 (52.1%)	2,542 (9.3%)	323 (1.2%)	636 (2.3%)										
NA	73 (73.0%)	27 (27.0%)	13 (48.1%)	3 (11.1%)	8 (29.6%)	28 (70.0%)	12 (30.0%)	2 (16.7%)	0 (0.0%)	1 (8.3%)										
Age group																				
35–69	2,316 (62.2%)	1,410 (37.8%)	536 (38.0%)	97 (6.9%)	73 (5.2%)	32,982 (45.7%)	39,241 (54.3%)	2,773 (7.1%)	480 (1.2%)	307 (0.8%)										
> 69	2,236 (65.1%)	1,198 (34.9%)	802 (66.9%)	140 (11.7%)	431 (36.0%)	16,318 (54.6%)	13,555 (45.4%)	3,524 (26.0%)	561 (4.1%)	1,357 (10.0%)										
Comorbidity index																				
0	3,658 (61.2%)	2,321 (38.8%)	1,144 (49.3%)	214 (9.2%)	395 (17.0%)	44,636 (47.2%)	49,996 (52.8%)	5,523 (11.0%)	927 (1.9%)	1,379 (2.8%)										
1	666 (74.8%)	224 (25.2%)	150 (67.0%)	17 (7.6%)	81 (36.2%)	3,512 (62.0%)	2,148 (38.0%)	615 (28.6%)	98 (4.6%)	224 (10.4%)										
2	190 (79.8%)	48 (20.2%)	38 (79.2%)	5 (10.4%)	26 (54.2%)	745 (68.7%)	340 (31.3%)	128 (37.6%)	13 (3.8%)	50 (14.7%)										
3	29 (90.6%)	3 (9.4%)	3 (100.0%)	0 (0.0%)	2 (66.7%)	67 (82.7%)	14 (17.3%)	4 (28.6%)	1 (7.1%)	1 (7.1%)										
NA	9 (42.9%)	12 (57.1%)	3 (25.0%)	1 (8.3%)	0 (0.0%)	340 (53.3%)	298 (46.7%)	27 (9.1%)	2 (0.7%)	10 (3.4%)										
Geographical deprivation index																				
1	893 (63.0%)	525 (37.0%)	267 (50.9%)	45 (8.6%)	107 (20.4%)	10,954 (58.5%)	7,761 (41.5%)	898 (11.6%)	140 (1.8%)	234 (3.0%)										
2	834 (63.4%)	481 (36.6%)	256 (53.2%)	52 (10.8%)	93 (19.3%)	9,839 (53.4%)	8,576 (46.6%)	1,005 (11.7%)	162 (1.9%)	282 (3.3%)										
3	889 (61.8%)	549 (38.2%)	269 (49.0%)	40 (7.3%)	93 (16.9%)	9,752 (48.5%)	10,357 (51.5%)	1,215 (11.7%)	185 (1.8%)	325 (3.1%)										
4	919 (65.3%)	489 (34.7%)	255 (52.1%)	51 (10.4%)	97 (19.8%)	9,472 (42.6%)	12,784 (57.4%)	1,435 (11.2%)	237 (1.9%)	355 (2.8%)										
5	1,013 (64.2%)	564 (35.8%)	291 (51.6%)	49 (8.7%)	114 (20.2%)	9,277 (41.1%)	13,302 (58.9%)	1,741 (13.1%)	316 (2.4%)	467 (3.5%)										
NA	4 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	6 (27.3%)	16 (72.7%)	3 (18.8%)	1 (6.2%)	1 (6.2%)										
PCA individual deprivation index																				
1	812 (59.8%)	546 (40.2%)	232 (42.5%)	47 (8.6%)	47 (8.6%)	13,129 (51.0%)	12,611 (49.0%)	923 (7.3%)	155 (1.2%)	113 (0.9%)										
2	527 (59.8%)	355 (40.2%)	150 (42.3%)	25 (7.0%)	23 (6.5%)	7,764 (44.0%)	9,884 (56.0%)	713 (7.2%)	129 (1.3%)	86 (0.9%)										
3	938 (65.7%)	490 (34.3%)	245 (50.0%)	58 (11.8%)	85 (17.3%)	10,045 (50.0%)	10,028 (50.0%)	1,345 (13.4%)	266 (2.7%)	385 (3.8%)										
4	1,098 (63.8%)	622 (36.2%)	358 (57.6%)	59 (9.5%)	187 (30.1%)	9,922 (48.6%)	10,481 (51.4%)	1,604 (15.3%)	261 (2.5%)	532 (5.1%)										
5	1,177 (66.4%)	595 (33.6%)	353 (59.3%)	48 (8.1%)	162 (27.2%)	8,440 (46.3%)	9,792 (53.7%)	1,709 (17.5%)	229 (2.3%)	547 (5.6%)										

* The percentage of hospitalization, ICU and deaths are referred to people positive to the COVID-19

Comorbidity index: 0: no comorbidity, 4: highest level of comorbidity. Index of geographical and PCA individual deprivation: 1: lowest level of deprivation, 5: highest level of deprivation

PCA Individual deprivation index

The census variables selected to build the individual DI were available for a subpopulation of 129,779 individuals. We have retained two principal components in our PCA: the variance explained by the principal component 1 was 34.1% and the variance explained by the principal component 2 was 25.5% (Additional file 1: Figure S4). The coefficients forming these components can be found in Additional file 1: Table S1.

Outcomes of interest by type of deprivation index

Figure 2 shows the counts of positive tests, COVID-19 hospitalisations, intensive care unit admissions and deaths among the DI groups per 100,000 tests or positive tests during the first or second wave respectively.

1st epidemic wave

According to the results of the logistic GLMs, during the first wave, the risk of testing positive for Sars-CoV-2 infection was not significantly different in people with a level of geographical deprivation higher than 1 when compared with the less deprived population group. The models that used PCA individual ID reported that in the more deprived population, the odds of testing positive were lower compared those estimated in the less deprived population, with ORs lower than one. Considering both DI indicators, the ORs of being hospitalised, admitted to the ICU or dying when positive did not exhibit significant differences between the subgroups of the population with the highest DI and those with the lowest DI. (Fig. 3, Table 3).

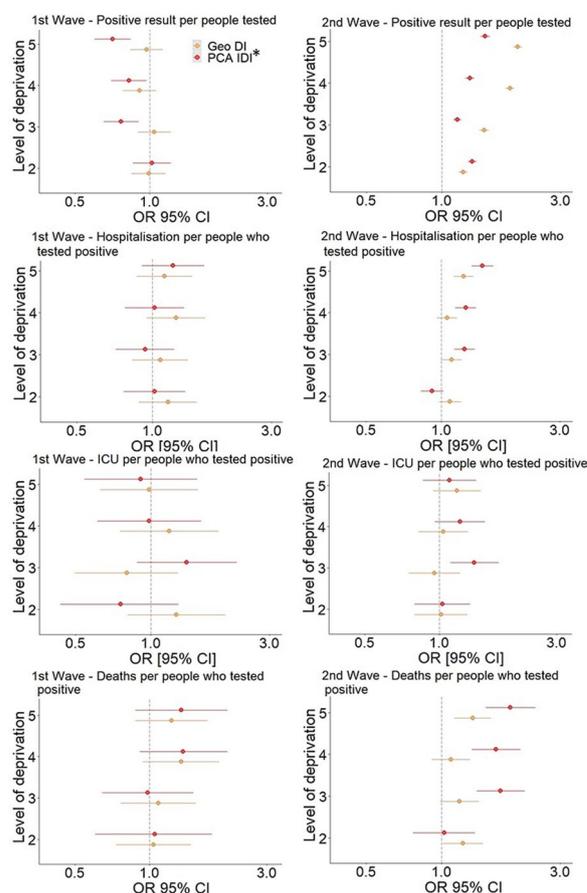


Fig. 3 Adjusted ORs per increase in the geographical, PCA individual DI for the counts of positive tests, hospitalisations, ICU admissions, and deaths. *IDI: individual deprivation index

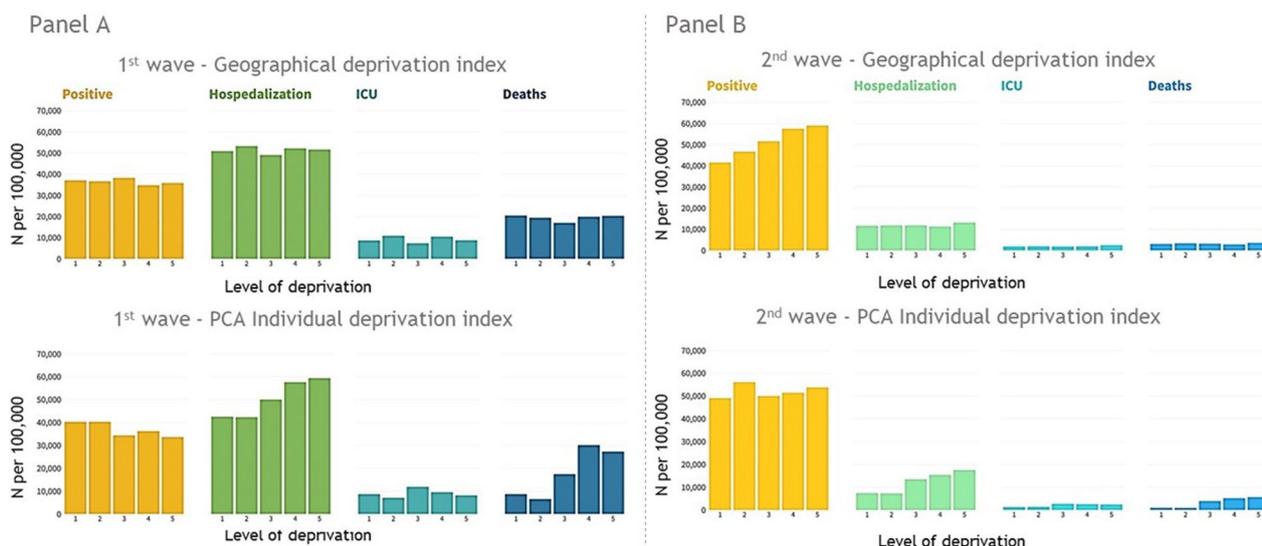


Fig. 2 Counts per 100,000 of positive tests, hospitalizations, ICU admissions, and deaths per DI level. A: 1st wave (20 February 2020 to 31 May 2020). B: 2nd wave (16 September 2020 to 31 December 2020)

Table 3 Associations between geographical and individual DI and risk of Sars-CoV-2 infection and disease severity

		2nd Wave																			
		Positive per tested		Hospitalisations per positive tests		ICU per positive tests		Deaths per positive tests		Geo DI		Positive per tested		Hospitalisations per positive tests		ICU per positive tests		Deaths per positive tests			
Geo DI	OR	CI 95%	OR	CI 95%	OR	CI 95%	OR	CI 95%	OR	CI 95%	OR	CI 95%	OR	CI 95%	OR	CI 95%	OR	CI 95%	OR	CI 95%	
1	Ref.		Ref.		Ref.		Ref.		Ref.		1	Ref.		Ref.		Ref.		Ref.		Ref.	
2	0.99	(0.84–1.16)	1.16	(0.89–1.51)	1.27	(0.81–1.99)	1.04	(0.73–1.48)	1.22	(1.17–1.27)	2	1.22	(1.17–1.27)	1.08	(0.98–1.20)	1.02	(0.79–1.30)	1.22	(1.01–1.47)		
3	1.04	(0.90–1.21)	1.08	(0.83–1.39)	0.80	(0.50–1.28)	1.09	(0.77–1.54)	1.48	(1.42–1.54)	3	1.48	(1.42–1.54)	1.10	(0.99–1.21)	0.96	(0.75–1.21)	1.18	(0.99–1.42)		
4	0.91	(0.78–1.06)	1.24	(0.95–1.63)	1.19	(0.75–1.87)	1.34	(0.94–1.92)	1.88	(1.80–1.95)	4	1.88	(1.80–1.95)	1.06	(0.96–1.16)	1.04	(0.83–1.31)	1.09	(0.91–1.30)		
5	0.97	(0.84–1.13)	1.12	(0.87–1.45)	0.98	(0.62–1.55)	1.23	(0.88–1.73)	2.02	(1.94–2.10)	5	2.02	(1.94–2.10)	1.23	(1.12–1.35)	1.18	(0.95–1.47)	1.34	(1.05–1.58)		
PCA Individual DI																					
1	ref		ref		ref		ref		ref		1	ref		ref		ref		ref		ref	
2	1.02	(0.86–1.22)	1.02	(0.77–1.36)	0.76	(0.44–1.30)	1.05	(0.61–1.81)	1.33	(1.28–1.38)	2	1.33	(1.28–1.38)	0.92	(0.83–1.36)	1.03	(0.79–1.33)	1.02	(0.77–1.36)		
3	0.77	(0.65–0.90)	0.94	(0.72–1.23)	1.40	(0.88–2.21)	0.98	(0.64–1.51)	1.16	(1.11–1.20)	3	1.16	(1.11–1.20)	1.24	(1.13–1.38)	1.39	(1.11–1.74)	1.73	(1.38–2.16)		
4	0.83	(0.70–0.97)	1.02	(0.78–1.34)	0.98	(0.61–1.59)	1.37	(0.91–2.06)	1.30	(1.25–1.35)	4	1.30	(1.25–1.35)	1.26	(1.14–1.62)	1.21	(0.96–1.54)	1.65	(1.32–2.07)		
5	0.71	(0.60–0.84)	1.21	(0.91–1.61)	0.91	(0.54–1.53)	1.34	(0.88–2.06)	1.49	(1.43–1.56)	5	1.49	(1.43–1.56)	1.46	(1.32–1.84)	1.10	(0.86–1.41)	1.89	(1.50–2.37)		

2nd epidemic wave

According to the results of the logistic GLMs, during the 2nd epidemic wave the risk of testing positive for Sars-CoV-2 infection was significantly higher in people with a level of deprivation higher than 1 when compared with the less deprived population group, with generally homogeneous results disregarding the DI used in the model. For what concern the odds of being hospitalised and dying if positive, considering the geographical DI, the subgroup with deprivation level 5 has a higher OR of being hospitalised (OR: 1.23; 95% CI 1.12–1.35) and dying (OR: 1.34; 95% CI 1.05–1.58) if positive than the subgroup with deprivation level 1. Calculating GLMs with the PCA individual DI, the OR of being hospitalised and dying if positive increased with increasing individual deprivation level also in population subgroups with a DI less than 5. The models that used individual ID reported that as deprivation increased, the risk of hospitalisation and death increased, which was not evident in the models that used geographic ID. The OR of being admitted to the ICU when positive in the subgroups of the population with the highest DI was not significantly different from the subgroup of the population with the lowest DI, irrespective of the type of DI used (Fig. 3, Table 3).

Multilevel logistic modelling

1st epidemic wave

The proportion of the between-cluster variation (ICC) in the probability of being positive if tested during the first wave was 9.9 (Table 4). This indicates that 9.9% of the chances was explained by between the census section differences. The deviance of the augmented intermediated model was significantly lower than the deviance of the constrained model. Table 4 reports multilevel logistic regression results. The model showed that higher level PCA individual DI were significantly associated with lower probability of being positive if tested. Additional file 1: Figure S5, Panel A shows the interaction between geographical and individual DI in the prospective prediction of being positive if tested. Although the association between geographical and individual DI was not significant, the coefficient (β) was negative. The proportion of the between-cluster variation in the probability of being hospitalised, admitted to ICU and dying if positive during the first wave was respectively 5.4, 4.9 and 2.1. Therefore, the results for the one-level regression model were considered.

2nd epidemic wave

The proportion of the between-cluster variation in the probability of being positive if tested during the first wave was 15.2. This indicates that 15.2% of the chances

Table 4 Results of multilevel logistic regression or GLM models including geographical and individual DI interaction

ICC for empty models	1st Wave		2nd Wave	
	Outcome: being positive if tested			
	0.10		0.15	
	β	SE	β	SE
PCA Individual DI	−0.12***	0.03	0.02°	0.01
Geo DI	−0.01	0.02	0.04***	0.01
Female	−0.02	0.05	−0.01	0.01
Age	0.01**	0.00	−0.01***	0.00
Charlson comorbidity index 1	−0.66***	0.09	−0.61***	0.03
Charlson comorbidity index 2	−0.96***	0.17	−0.89***	0.07
Charlson comorbidity index 3	−1.69**	0.62	−1.77***	0.30
Geo x Individual DI	0.00	0.01	0.00	0.00

*** p-value < 0.001; ** p-value = 0.001; * p-value = 0.01; ° p-value = 0.05

was explained by between the census section differences. The deviance of the augmented intermediated model was significantly lower than the deviance of the constrained model. The model shows that both higher level PCA individual DI or geographical DI were significantly associated with higher probability of being positive if tested (Table 4). Additional file 1: Figure S5, Panel B shows the interaction between geographical and individual DI in the prospective prediction of being positive if tested. Although the association between geographical and individual DI was not significant, the coefficient (β) was positive. The proportion of the between-cluster variation in the probability of being hospitalised, admitted to ICU and dying if positive during the first wave was respectively 1.3, 2.4 and 1.9. Therefore, the results for the one-level regression model were considered.

Discussion

This study reports results from a large population-wide cohort of people tested for COVID-19 in the Apulia region, Italy, during the first and second wave of the pandemic in 2020. To our knowledge, this was the first study to investigate the role of individual DI on COVID-19 outcomes in Italy. Previous studies were limited to the use of an area-based DI or the individual component of socioeconomic deprivation as a proxy of individual deprivation [17, 31–33]. Although for the first wave, the geographic DI and the individual DI calculated in this study show no differences in the non-significance of the association with COVID-19 outcomes, the results were different for the second wave.

According to our findings, the association between the risk of testing positive for Sars-CoV-2 and the level of socioeconomic deprivation in the Apulia region changed

between the first and second waves. According to the analysis done with the PCA individual DI, during the first wave, there was a significant inversely proportional trend between the DI and the risk of testing positive. As deprivation increased, the risk of being positive when tested decreased. During the second wave, individuals with a higher level of socio-economic deprivation had a higher statistically significant probability of testing positive. This may be explained by the fact that during the early stages of the epidemic outbreak, with affected geographical areas still circumscribed, most of the cases in the Apulia region were due to returning residents [34]. Indeed, the gradual implementation of control measures in Italy sparked substantial movements of people travelling from northern regions at the epicentre of the epidemic toward other regions, such as Apulia. Then, the swift extension of lockdown to the entire country [22] mitigated the impact of these COVID-19 seeding events and the epidemic in Apulia was successfully contained. For these reasons, during the first wave in Apulia, the epidemic probably spread mainly among individuals who had the financial means to travel or who were economic migrants in Lombardy or northern regions, areas of intense economic activities. During the first wave, individuals with a higher level of socio-economic deprivation did not have a statistically significant higher probability of being hospitalised or dying if positive. These results were in line with those reported by Furtunato et al. [35]. While during the second wave, individuals with a higher level of socio-economic deprivation had a higher statistically significant higher probability of being hospitalised or dying if positive.

Concerning the second wave, the two DI showed a clear association between the probability of testing positive and the level of deprivation. However, the geographic DI showed a stronger trend. The results of this study showed that most people with a higher individual DI live in the most deprived census areas. Since people with a higher individual DI were more often in jobs that were less amenable to remote working and so they benefited less from lockdown restrictions than those able to work from home; they had a higher likelihood of acquiring the infection [36]. Moreover, they generally live in higher-density environments where family members could be infected secondarily [37]. This may have meant that in more deprived areas the virus circulated more widely than in less deprived areas. As individual DI increased, so did the probability of hospitalisation and death. The same trend was not observed for geographical DI, which showed a significantly increased probability of being hospitalised or dying if positive only for the highest level of deprivation. In accordance with our results, Mateo-Urdiales and colleagues, using the same geographical DI used in

this study, but at the municipality level, found a higher incidence of cases in the most deprived municipalities compared with the least deprived ones and no differences in case-hospitalisation and case-fatality according to deprivation were observed in any period under study. The same result has also been reported by other studies using geographical measures of deprivation conducted in Spain [38, 39]. While many factors could explain this finding, an alternative could be that hospitalisation and death cases were similar across areas with different levels of deprivation in a well-developed universal healthcare system, such as Italy and Spain.

The use of individual DI enabled to understand that actually the most disadvantaged people had a higher risk of hospitalisation and death, regardless of the area in which they lived. To avoid exacerbating existing social inequalities and marginalisation, it is essential to be able to monitor them. According to our results, using geographical DI as a proxy for individual DI may lead to inaccurate assessments. However, the geographic DI can have a relevance in providing insight into the distribution of infection within different neighbourhoods, accounting for their heterogeneity. This may have important implications for public health action planning.

Understanding the relationship between deprivation and COVID-19 outcomes is multifaceted and complex. This is why the secondary objective of this manuscript was to test the hypotheses about how individual and geographical DI interact to predict COVID-19 outcomes. According to the results of the multilevel logistic and GLM models, there was no association between COVID-19 outcomes and the interaction between PCA individual DI and geographical DI. Although this may seem at odds with the robust evidence in the literature demonstrating an interaction between socioeconomic status in one's neighbourhood, individual deprivation and health, this evidence is mainly based on health outcomes of non-communicable chronic diseases [40–43].

The COVID-19 pandemic has been described as a syndemic pandemic [44]. Originating in anthropology, a syndemic describes a set of closely intertwined and mutual enhancing health problems that significantly affect the overall health status of a population within the context of a perpetuating configuration of noxious social conditions [45]. Deprivation—which is an area measure of poverty, low income, and a reflection of the wider social determinants of health (such as housing, working conditions, unemployment, health-care access, etc.)—results in multiple, interacting, and additive adverse risk factors for COVID-19 outcomes. These can be summarised by way of four inter-related pathways: unequal exposure, unequal transmission, unequal vulnerability, and unequal susceptibility [46]. Living in a more deprived neighbourhood

may increase the likelihood of exposure and transmission. Although it has been proven that individuals living in more deprived neighbourhoods have worse profiles on many components of subjective health, more risk factors and higher morbidity and mortality rates than their counterparts living in less deprived neighbourhoods, the level of neighbourhood deprivation does not influence with a higher extent the health of those with low individual DI compared to those with a higher individual DI [43]. This could be the reason why not all individuals living in the most deprived neighbourhoods have a higher vulnerability and susceptibility to COVID-19 and would explain why the analyses in this study give different results when using geographical DI than when using individual DI.

This study had some limitations. Firstly, policies for the execution of the diagnostic test changed during the period under study. During the first wave, the diagnostic capacity was limited, and the number of positives could be under-reported, whereas for the second wave data could represent laboratory-confirmed COVID-19 cases who sought care. Secondly, the death status only reflects deaths occurring in individuals diagnosed with COVID-19, but this number may be underestimated due to non-diagnosis. Third, the Charlson comorbidities index was utilized as a comorbidity index, though it may not be the most suitable or up-to-date choice for studies related to SARS-CoV-2. Unfortunately, we could not independently calculate a comorbidity index due to a lack of available data. Finally, the DI is a composite measure of deprivation, and it is difficult to know which of its constituent factors are driving associations. There are other potential variables, such as income and homeownership, that could enhance the deprivation index, but we were unable to assess them due to information unavailability.

Conclusion

Historically, pandemics have been experienced unequally with higher rates of infection and mortality among the most disadvantaged communities [47]. Evidence from a variety of countries suggests that these inequalities are being mirrored in the COVID-19 pandemic [36, 48] and this study also adds to this evidence. The causal forces at work at the interface of disease epidemics and social inequality are complex: they are operant at the level of individuals, neighbourhoods, and local communities. The results of this study remind us to be cautious about using geographical DI as a proxy for the level of individual social disadvantage because of the inevitable potential ecological bias that can result from attributing a collective measure to an individual and may lead to inaccurate assessments.

The geographical DI is often used due to a lack of individual data. However, on the determinants of

health and health inequalities, monitoring has to have a central focus. Health inequalities monitoring provides evidence on who is being left behind and informs equity-oriented policies, programmes and practices [49]. The emergence of novel data sources and monitoring approaches for public health surveillance is triggered by technological advances in data retrieval and data analysis. Unfortunately, the information richness resulting from the data science revolution is not exploited to better understand health inequities and inform the development and implementation of services and policies that tackle inequities in health [50, 51].

Reducing these inequalities—and those that may result from future pandemics—requires long-term action to improve equity in health and wealth. Future research and data collection should focus on improving surveillance systems, for example by integrating individual measures of inequalities into national health information systems. Understanding the social nature of emerging infectious disease pandemics will ultimately help reduce the burden of disease.

Abbreviations

DI	Deprivation Index
PCA	Principal component analysis
SEP	Socioeconomic position
SDH	Social determinant of health
SSN	Servizio Sanitario Nazionale
R and AP	Regions and two autonomous provinces
ICU	Intensive care unit
GLM	Generalized linear logistic regression models
OR	Odds Ratio
CI	Confidence Interval

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12942-024-00367-6>.

Additional file 1: Figure S1. The geographical boundaries of Italy and its regions with the Apulia region in red. **Figure S2.** Summary of the three steps performed to build the multilevel logistic regression. **Figure S3.** Epidemic curve of positive COVID-19 PCR test results (7 days moving average). **Figure S4.** Scree plot of PCA performed on census variables. **Figure S5.** Interaction between geographical and individual DI in the prospective prediction of being positive if tested. Estimated values reflect statistical adjustment for sex, age, and Charlson comorbidity index. **Table S1.** Coefficients forming components retained in the PCA.

Acknowledgements

We would like to acknowledge Damiano Morando for his statistical support.

Author contributions

SM, GP, and LT conceived the study and developed the study protocol. AC and AN performed data merging and data cleaning. SM and GP performed data analysis and results interpretation in consultation with LT, DM, and LB. SM, GP and LT drafted the manuscript in collaboration with DM, LB, and PLL. LB, DM, PLL, and LT provided expert insights and contributed to manuscript drafting. All authors read, revised, and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

The data that support the findings of this study are available from Strategic Regional Health and Social Agency of Puglia (AReSS Puglia) but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of AReSS Puglia.

Declarations**Ethics approval and consent to participate**

The study protocol was approved on March 26, 2021, by the Joint Ethical Committee for research of the Scuola Normale Superiore (Pisa, Italy) and the Sant'anna School for Advanced Studies (Pisa, Italy). Number of approval: 6/2021.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Scuola Normale Superiore, Pisa, Italy. ²Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy. ³Health Science Interdisciplinary Research Centre, Sant'Anna School of Advanced Studies, Pisa, Italy. ⁴Strategic Regional Health and Social Agency of Puglia (AReSS Puglia), Bari, Italy. ⁵Department of Biological and Environmental Sciences and Technology, University of Salento, Lecce, Italy. ⁶Hygiene Unit, Policlinico Foggia Hospital, Department of Medical and Surgical Sciences, University of Foggia, Foggia, Italy.

Received: 28 September 2023 Accepted: 13 March 2024

Published online: 04 April 2024

References

- Cash R, Patel V. Has COVID-19 subverted global health? *Lancet*. 2020;395(10238):1687–8.
- Worobey M, Pekar J, Larsen BB, Nelson MI, Hill V, Joy JB, et al. The emergence of SARS-CoV-2 in Europe and North America. *Science*. 2020;370(6516):564–70.
- WHO. World Health Organization. Coronavirus disease (COVID-19) Situation Report—148 Data as received by WHO from national authorities by 10:00 CEST, 16 June 2020. *Fam Pract Guidel*. 2020;(June):1–18. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>. Accessed 20 Mar 2023.
- World Health Organization. WHO Coronavirus (COVID-19) Dashboard. <https://covid19.who.int/>. Accessed 17 Mar 2023.
- Istituto Superiore di Sanità. Dati della Sorveglianza integrata COVID-19 in Italia. 2023. <https://www.epicentro.iss.it/coronavirus/sars-cov-2-dashbord>. Accessed 20 Mar 2023.
- Zheng Z, Peng F, Xu B, Zhao J, Liu H, Peng H, Li Q, Jiang C, Zhou Y, Liu SYC. Risk factors of critical & mortal COVID-19 cases: a systematic literature review and meta-analysis. *J Infect*. 2020;81: e16.
- Hamer M, Kivimäki M, Gale CR, Batty GD. Lifestyle risk factors, inflammatory mechanisms, and COVID-19 hospitalization: a community-based cohort study of 387,109 adults in UK. *Brain, Behav Immun*. 2020;87:184–7.
- Wang J, Geng L. Effects of socioeconomic status on physical and psychological health: lifestyle as a mediator. *Int J Environ Res Public Health*. 2019;16(2):281.
- Drozdz M, Pujades-Rodriguez M, Lillie PJ, Straw S, Morgan AW, Kearney MT, et al. Non-communicable disease, sociodemographic factors, and risk of death from infection: a UK Biobank observational cohort study. *Lancet Infect Dis*. 2021;21(8):1184–91. [https://doi.org/10.1016/S1473-3099\(20\)30978-6](https://doi.org/10.1016/S1473-3099(20)30978-6).
- Caul S. Deaths involving COVID-19 by local area and socioeconomic deprivation. *Off Natl Stat*. 2020;20(May):1–23.
- Bukhman G, Mocumbi AO, Atun R, Becker AE, Bhutta Z, Binagwaho A, et al. The lancet NCDI poverty commission: bridging a gap in universal health coverage for the poorest billion. *Lancet*. 2020;396(10256):991–1044.
- Horton R. Offline: COVID-19 is not a pandemic. *Lancet*. 2020;396(10255):874. [https://doi.org/10.1016/S0140-6736\(20\)32000-6](https://doi.org/10.1016/S0140-6736(20)32000-6).
- Marmot M, Allen J, Goldblatt P, Herd E, Morrison J. Build Back Fairer: The COVID-19 Marmot Review. The Health Foundation; 2020. <https://www.health.org.uk/publications/build-back-fairer-the-covid-19-marmot-review>.
- Jung J, Manley J, Shrestha V. Coronavirus infections and deaths by poverty status: the effects of social distancing. *J Econ Behav Organ*. 2021;182:311–30.
- Caranci N, Biggeri A, Grisotto L, Pacelli B, Spadea T, Costa G. L'indice di deprivazione italiano a livello di sezione di censimento: definizione, descrizione e associazione con la mortalità The Italian deprivation index at census block level: definition, with general mortality. *Epidemiol Prev*. 2010;34(4):167–76.
- Rosano A, Pacelli B, Zengarini N, Costa G, Cislighi C, Caranci N. Aggiornamento e revisione dell'indice di deprivazione italiano 2011 a livello di sezione di censimento Update and review of the 2011 Italian deprivation index calculated at the census section level. *Epidemiol Prev*. 2020;44:162–70.
- Mateo-Urdiales A, Fabiani M, Rosano A, Vescio MF, Del Manso M, Bella A, et al. Socioeconomic patterns and COVID-19 outcomes before, during and after the lockdown in Italy (2020). *Heal Place*. 2021;71(March):102642. <https://doi.org/10.1016/j.healthplace.2021.102642>.
- Politi J, Martín-Sánchez M, Mercuriali L, Borrás-Bermejo B, Lopez-Contreras J, Vilella A, et al. Epidemiological characteristics and outcomes of COVID-19 cases: mortality inequalities by socioeconomic status, Barcelona, Spain, 24 February to 4 May 2020. *Eurosurveillance*. 2021;26(20):1–12. <https://doi.org/10.2807/1560-7917.ES.2021.26.20.2001138>.
- France G, Taroni F, Donatini A. The Italian health-care system. *Health Econ*. 2005;14(SUPPL. 1):187–202.
- Ferré F, de Belvis AG, Valerio L, Longhi S, Lazzari A, Fattore G, et al. Italy: health system review. *Health Syst Transit*. 2014;16(4):1–166.
- Osservatorio Nazionale sulla Salute nelle Regioni Italiane. Rapporto Osservasalute 2019 Stato di salute e qualità dell'assistenza nelle regioni italiane. 2019. <https://osservatoriosullasalute.it/wp-content/uploads/2020/06/Rapporto-Osservasalute-2019.pdf>
- Decree of the Prime Minister. Ulteriori disposizioni attuative del decreto-legge 23 febbraio 2020, n. 6, recante misure urgenti in materia di contenimento e gestione dell'emergenza epidemiologica da COVID-19, applicabili sull'intero territorio nazionale. (20A01605). *GU Ser Gen n 64 del 11 marzo 2020*. <http://www.trovanorme.salute.gov.it/norme/dettaglioAtto?id=73643>. Accessed 20 Mar 2023.
- Il Presidente del Consiglio dei Ministri. Decreto del presidente del consiglio dei ministri 26 aprile 2020. Ulteriori disposizioni attuative del decreto-legge 23 febbraio 2020, n. 6, recante misure urgenti in materia di contenimento e gestione dell'emergenza epidemiologica da COVID-19. *Gazz Uff della REPUBBLICA Ital*. 2020. <https://www.gazzettaufficiale.it/eli/id/2020/04/27/20A02352/sg>. Accessed 20 Mar 2023.
- Il Presidente del Consiglio dei Ministri. Decreto del presidente del consiglio dei ministri 11 giugno 2020. Ulteriori disposizioni attuative del decreto-legge 23 febbraio 2020, n. 6, recante misure urgenti in materia di contenimento e gestione dell'emergenza epidemiologica da COVID-19. *Gazz Uff Della REPUBBLICA Ital*. 2020. <https://www.gazzettaufficiale.it/eli/id/2020/06/11/20A03194/sg>. Accessed 20 Mar 2023.
- Il Presidente della Repubblica. Misure urgenti connesse con la proroga della dichiarazione dello stato di emergenza epidemiologica da COVID-19 e per la continuità operativa del sistema di allerta COVID, nonché per l'attuazione della direttiva (UE) 2020/739 del 3 giugno 2020. *Gazz della REPUBBLICA Ital*. 2020. <https://www.trovanorme.salute.gov.it/norme/dettaglioAtto?id=76574>. Accessed 20 Mar 2023.
- Il Presidente del Consiglio dei Ministri. Decreto del presidente del consiglio dei ministri 11 giugno 2020 Ulteriori disposizioni attuative del decreto-legge 25 marzo 2020, n. 19, recante misure urgenti per

- fronteggiare l'emergenza epidemiologica da COVID-19, e del decreto-legge 16 maggio 2020, n. Gazz Uff della Repubb Ital. 2020. <https://www.gazzettaufficiale.it/eli/id/2020/06/11/20A03194/sg>. Accessed 20 Mar 2023.
27. Modi C, Ferraro S, Stein G. Estimating COVID-19 mortality in Italy early in the COVID-19 pandemic. *Nat Commun*. 2021;12:2729.
 28. Istituto Nazionale di Statistica (ISTAT). Il Censimento permanente della popolazione in Puglia—Prima diffusione dei dati definitivi 2018 e 2019. 2021; https://www.istat.it/it/files//2021/02/Censimento-permanente-della-popolazione_Puglia.pdf. Accessed 20 Mar 2023.
 29. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373–83.
 30. Sommet N, Morselli D. Keep calm and learn multilevel logistic modeling: a simplified three-step procedure using stata, R, Mplus, and SPSS. *Int Rev Soc Psychol*. 2017;30(1):203–18.
 31. Di Girolamo C, Bartolini L, Caranci N, Moro ML. Socioeconomic inequalities in overall and covid-19 mortality during the first outbreak peak in emilia-romagna region (Northern Italy). *Epidemiol Prev*. 2020;44(5–6):288–96.
 32. Rostila M, Cederström A, Wallace M, Brandén M, Malmberg B. Disparities in coronavirus disease 2019 mortality by Country of Birth in Stockholm, Sweden: a total-population—based cohort study. *Am J Epidemiol*. 2021;190(3):1510–8.
 33. Pagani G, Conti F, Giacomelli A, Oreni L, Beltrami M, Pezzati L, et al. Differences in the prevalence of SARS-CoV-2 infection and access to care between Italians and non-Italians in a social-housing neighbourhood of Milan Italy. *Environ Res Public Heal*. 2021;18:10621.
 34. Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R, et al. Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures. *Proc Natl Acad Sci U S A*. 2020;117(19):10484–91.
 35. Fortunato F, Lillini R, Martinelli D, Iannelli G, Ascagnolo L, Casanova G, et al. Association of socio-economic deprivation with COVID-19 incidence and fatality during the first wave of the pandemic in Italy: lessons learned from a local register-based study. *Int J Health Geogr*. 2023;22(1):1–12. <https://doi.org/10.1186/s12942-023-00332-9>.
 36. McGowan VJ, Bambra C. COVID-19 mortality and deprivation: pandemic, syndemic, and endemic health inequalities. *Lancet Public Heal*. 2022;7(11):e966–75. [https://doi.org/10.1016/S2468-2667\(22\)00223-7](https://doi.org/10.1016/S2468-2667(22)00223-7).
 37. Gibson M, Petticrew M, Bambra C, Sowden AJ, Wright KE, Whitehead M. Housing and health inequalities: a synthesis of systematic reviews of interventions aimed at different pathways linking housing and health. *Heal Place*. 2011;17(1):175–84. <https://doi.org/10.1016/j.healthplace.2010.09.011>.
 38. García CN. Socioeconomic, demographic and healthcare determinants of the COVID-19 pandemic: an ecological study of Spain. *BMC Public Health*. 2021;21(1):1–8.
 39. Zaldo-Aubanell Q, Campillo I, López F, Bach A, Serra I, Olivet-Vila J, Saez M, et al. Community risk factors in the covid-19 incidence and mortality in Catalonia (Spain). A population-based study. *Int J Environ Res Public Health*. 2021;18(7):3768.
 40. Arcaya MC, Tucker-Seeley RD, Kim R, Schnake-Mahl A, So M, Subramanian SV. Research on neighborhood effects on health in the United States: a systematic review of study characteristics. *Soc Sci Med*. 2016;168:16–29. <https://doi.org/10.1016/j.socscimed.2016.08.047>.
 41. Diez Roux AV, Jacobs DR, Kiefe CI. Neighborhood characteristics and components of the insulin resistance syndrome in young adults: the coronary artery risk development in young adults (CARDIA) study. *Diabetes Care*. 2002;25(11):1976–82.
 42. Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *J Epidemiol Community Health*. 2001;55(2):111–22.
 43. Boylan JM, Robert SA. Neighborhood SES is particularly important to the cardiovascular health of low SES individuals. *Soc Sci Med*. 2017;188:60–8.
 44. Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. *J Epidemiol Community Health*. 2020;74(11):964–8.
 45. Singer M, Bulled N, Ostrach B, Mendenhall E. Syndemics and the biosocial conception of health. *Lancet*. 2017;389(10072):941–50. [https://doi.org/10.1016/S0140-6736\(17\)30003-X](https://doi.org/10.1016/S0140-6736(17)30003-X).
 46. Bambra C. Pandemic inequalities: emerging infectious diseases and health equity. *Int J Equity Health*. 2022;21(1):4–7. <https://doi.org/10.1186/s12939-021-01611-2>.
 47. Summers JA, Stanley J, Baker MG, Wilson N. Risk factors for death from pandemic influenza in 1918–1919: a case-control study. *Influenza Other Respi Viruses*. 2014;8(3):329–38.
 48. Magesh S, John D, Li WT, Li Y, Mattingly-App A, Jain S, et al. Disparities in COVID-19 outcomes by race, ethnicity, and socioeconomic status: a systematic-review and meta-analysis. *JAMA Netw Open*. 2021;4(11):1–16.
 49. Marmot M, Goldblatt P. Importance of monitoring health inequalities. *BMJ*. 2013;347(November):1–2. <https://doi.org/10.1136/bmj.f6576>.
 50. Connelly R, Playford CJ, Gayle V, Dibben C. The role of administrative data in the big data revolution in social science research. *Soc Sci Res*. 2016;59:1–12. <https://doi.org/10.1016/j.ssresearch.2016.04.015>.
 51. Card DE, Chetty R, Feldstein MS, Saez E. Expanding Access to Administrative Data for Research in the United States. *SSRN Electron J*. 2012;

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.