BMJ Open

Educational status and COVID-19 related outcomes in India: hospital-based cross-sectional study

Arvind K Sharma,[1] Rajeev Gupta [ORCID] ,[2,3] Vaseem Naheed Baig,[1] Veer Teja Singh,[1] Surabhi Chakraborty,[1] Jagdish P Sunda,[1] Prahalad Dhakar,[4] Shiv Prakash Sharma,[1] Raja Babu Panwar,[2] Vishwa Mohan Katoch[5]

[1]Department of Community Medicine, RUHS College of Medical Sciences, Jaipur, India
[2]Academic Research Development Unit, Rajasthan University of Health Sciences, Jaipur, Rajasthan, India
[3]Department of Medicine, Eternal Heart Care Centre and Research Institute, Jaipur, Rajasthan, India
[4]Department of Medicine, RUHS College of Medical Sciences, Jaipur, Rajasthan, India
[5]ICMR NASI Chair in Public Health Research, Rajasthan University of Health Sciences, Jaipur, Rajasthan, India

Correspondence to
Dr Rajeev Gupta;
rajeevgg@gmail.com

ABSTRACT

Objective Association of educational status, as marker of socioeconomic status, with COVID-19 outcomes has not been well studied. We performed a hospital-based cross-sectional study to determine its association with outcomes.

Methods Successive patients of COVID-19 presenting at government hospital were recruited. Demographic and clinical details were obtained at admission, and in-hospital outcomes were assessed. Cohort was classified according to self-reported educational status into group 1: illiterate or ≤primary; group 2: higher secondary; and group 3: some college. To compare intergroup outcomes, we performed logistic regression.

Results 4645 patients (men 3386, women 1259) with confirmed COVID-19 were recruited. Mean age was 46±18 years, most lived in large households and 30.5% had low educational status. Smoking or tobacco use was in 29.5%, comorbidities in 28.6% and low oxygen concentration (SpO_2 <95%) at admission in 30%. Average length of hospital stay was 6.8±3.7 days, supplemental oxygen was provided in 18.4%, high flow oxygen or non-invasive ventilation 7.1% and mechanical ventilation 3.6%, 340 patients (7.3%) died. Group 1 patients had more tobacco use, hypoxia at admission, lymphocytopaenia, and liver and kidney dysfunction. In group 1 versus groups 2 and 3, requirement of oxygen (21.6% vs 16.7% and 17.0%), non-invasive ventilation (8.0% vs 5.9% and 7.1%), invasive ventilation (4.6% vs 3.5% and 3.1%) and deaths (10.0% vs 6.8% and 5.5%) were significantly greater (p<0.05). OR for deaths were higher in group 1 (1.91, 95% CI 1.46 to 2.51) and group 2 (1.24, 95% CI 0.93 to 1.66) compared with group 3. Adjustment for demographic and comorbidities led to some attenuation in groups 1 (1.44, 95% CI 1.07 to 1.93) and 2 (1.38, 95% CI 1.02 to 1.85); this persisted with adjustments for clinical parameters and oxygen support in groups 1 (1.38, 95% CI 0.99 to 1.93) and 2 (1.52, 95% CI 1.01 to 2.11).

Conclusion Low educational status patients with COVID-19 in India have significantly greater adverse in-hospital outcomes and mortality.

Trial registration number REF/2020/06/034036.

Strengths and limitations of this study

► Studies in high-income countries have reported that low socioeconomic status is a risk factor for adverse outcomes in COVID-19. Similar studies are not available in lower-middle and low-income countries.

► This study shows that low educational status patients with COVID-19 in India have significantly higher in-hospital mortality compared with the better educated.

► Low educational status patients have more severe disease at presentation with greater requirement of oxygen and ventilation.

► Important limitations are lack of area-based measures, neighbourhood details, biochemical and inflammatory markers of severity of illness and absence of long-term follow-up.

INTRODUCTION

COVID-19 pandemic continues to devastate human lives and livelihoods, especially in low and lower-middle income countries.[1]

After the initial spread to the high-income countries in Europe and North America, the epidemic is now rapidly escalating in middle-income and low-income countries of South America, South Asia, South East Asia and Africa.[2] Epidemiological studies from China, Europe, UK and USA have shown greater disease burden in socioeconomically deprived neighbourhoods and minority ethnic groups.[3] A review that included more than 18.7 million patients from 50 studies in UK and USA reported that individuals from black and Asian ethnicities had 1.5–2.0 time greater risk of COVID-19 infection compared with white individuals and individuals of Asian ethnicity were at greater risk for intensive care unit admission and death.[4] Multiple reasons have been postulated for these socioeconomic disparities and include factors such as poverty, racism and other structural factors, lower availability, access, affordability and utilisation of healthcare and low value care.[5 6] Greater load of infection and longer exposure to the virus due to crowded environments, limited housing, large household

sizes, low quality jobs, unsafe commute and undernutrition are also important.[6 7]

Educational status is an important marker of socioeconomic status, and hundreds of studies in fields of communicable and non-communicable diseases have reported association of low educational status with adverse health-related events.[8–10] It is also an independent risk factor for morbidity and mortality from infectious diseases .[8 11] Association of socioeconomic status with COVID-19 related outcomes has not been well studied. A rapid review identified 42 studies that evaluated social determinants of COVID-19 incidence, clinical presentation, health service use and outcomes[3] and reported significant associations of race, ethnicity and social deprivation with increased COVID-19 incidence and hospitalisation. The review also reported that there was limited evidence regarding other key determinants including occupation, education, housing status and food security and suggested larger epidemiological studies to obtain high-quality evidence. A number of more recent studies have highlighted importance of socioeconomic inequalities in COVID-19 related morbidity and mortality,[12–14] and a review that included 34 studies has reported substantial racial, ethnic and socioeconomic variation in incidence of COVID-19 in USA with greater incidence among poorer communities.[15]

India has one of the largest burdens of COVID-19 cases and deaths.[16] A macrolevel study reported that Indian states with greater human development index and other socioeconomic indices had higher per capita COVID-19 incidence and deaths.[17] Although anecdotal evidence and modelling data exist,[1 18] there are no significant data on association of individual-level socioeconomic status with disease incidence and outcomes. Therefore, to examine association of self-reported educational status,[9 10] as a marker of socioeconomic status, in confirmed COVID-19 cases successively admitted to a dedicated COVID-19 government hospital in India, we performed a prospective registry-based study.

METHODS
We conducted a hospital-based prospective observational study on patients with laboratory-confirmed COVID-19 admitted to a 1200-bed dedicated COVID-19 government hospital (Rajasthan University of Health Sciences Hospital, Jaipur) from April to mid-September 2020. Initial data on patients have been reported earlier.[19 20] It is registered with Clinical Trials Registry of India at www.ctri.nic.in.

Patient data
Successive patients aged 18 years or more, presenting to the hospital for admission with suspicion of COVID-19 infection were enrolled in the study. Only those who tested positive for COVID-19 on nasopharyngeal and oropharyngeal reverse transcriptase PCR (RT-PCR) test have been included. All RT-PCR positive patients admitted from 1 April to 15 September have been included.

Patients recruited into the study in mid-September were followed up to discharge or death and outcome events were recorded.

A questionnaire was developed, and details of sociodemographic, clinical, laboratory, treatments and outcomes variables were recorded using patients' history and medical records.[19] Demographic details were obtained at the time of admission. These included name, age, sex, residence address and educational status. Status of highest educational level achieved was self-reported similar to most of the previous studies.[9 10] Other sociodemographic variables were not available for majority of patients and are not reported. Although it is possible to obtain individual details from unique identification number (Aadhaar number) or other identifiers of all the COVID-19 cases, we did not use these data. All the COVID-19 RT-PCR reports along with the government identifier are uploaded on the official website of Indian Council of Medical Research at wwwicmrgovin. Details of physical examination at the time of admission were obtained from patient case files. These included history of duration of symptoms at admission, pulse, blood pressure (BP), respiratory rate and surface oxygen concentration (SpO_2). Details of investigations at admission were obtained from the case files and biochemistry, microbiology and pathology departments as reported earlier.[20] We do not have data on serial investigations. We obtained data on duration of hospital stay from medical record department. For patients discharged alive from the hospital, we obtained data on patients who required oxygen support (nasal prongs, facial mask or high-flow nasal cannula), non-invasive ventilation (continuous positive airways pressure (CPaP) or bilevel positive airways pressure (BiPaP) support) or invasive ventilation after endotracheal intubation. Binary outcomes were obtained for all patients and included either recovery, referral to non-government hospitals on request of family or death. All these data have also been sent to the Department of Health, Government of Rajasthan, India, but are not currently accessible.

Patient and public involvement
Patients and the public were not involved in the design, or conduct, or reporting of this research. The preprint (*medRxiv preprints*. https://doi.org/10.1101/2021.05.17.21257364) has been shared with the administrative authorities of Government of Rajasthan.

Statistical analyses
The data were computerised, and data processing was performed using commercially available statistical software, SPSS V.20.0. Educational status was self-reported, and patients were classified into three groups: group 1: illiterate or ≤primary education, group 2: >primary to higher secondary school education and group 3: any graduate or postgraduate college education. Numerical data are expressed as mean±1 SD and categorical data as per cent. Significance of intergroup differences were

calculated using either χ^2 test or analysis of variance (ANOVA) as appropriate. χ^2 test residuals were determined for categorical variables in various groups and significant were age, sex, household size and some clinical parameters. Tests of normality for continuous variables was performed in the statistical programme, and all followed a normal Gaussian distribution. The variables where significant residuals identified were adjusted using logistic regression. We also compared mean and proportionate differences in groups 1 and 2 as compared with group 3 using unpaired t-test or χ^2 test as appropriate. To evaluate association of educational status with clinical outcomes, we performed stepwise logistic regression. Univariate and multivariate ORs and 95% CIs were calculated for group 1 and group 2 compared with group 3 for outcomes of in-hospital death, invasive ventilation and non-invasive ventilation. We initially calculated the univariate ORs and subsequently performed a stepwise logistic regression with sequential adjustment with: (1) age and sex, (2) household size, (3) cardiovascular risk factors and comorbidities, (4) clinical features and investigations at presentation and finally with (5) oxygenation during hospital stay, and determined multivariate ORs. P value of <0.05 is considered significant.

RESULTS

Patients were enrolled from March 2020 to mid-September 2020. A total of 7349 patients were hospitalised with confirmed or suspected COVID-19 during this period, 5103 patients (69.0%) tested positive for the disease on RT-PCR test and for the present study 4645 individuals (91.0% of confirmed cases), men 3386 (72.9%) and women 1259 (27.1%), in whom detailed clinical data were available have been included (table 1). The mean age of the cohort was 45.9±18 years, 54% were less than 50 years and about half lived in large family households. Prevalence of low educational status was high and greater in women while tobacco use was more in men (online supplemental file). Comorbidities were present in 28.6% with hypertension and diabetes being the most common. Details of symptoms, laboratory investigations and clinical status at admission is shown in table 1. Data on haematological investigations were available in 4456 (95.9%) and for biochemical tests in 867 (18.7%) patients. All patients received standard treatment according to guidelines available from Indian Council of Medical Research and the state government.[21] Management included oral or intravenous hydration, paracetamol and oral or intravenous antibiotics if required. A number of patients also received hydroxychloroquine, ivermectin, azithromycin, doxycycline, lopinavir-ritonavir, favipiravir, etc. The average length of stay in hospital was 6.8±3.7 days and was significantly greater in men (6.9±3.8 days) than in women (6.5±3.6 days) (p=0.004). Oxygen requirement was significantly greater in women, but other outcomes such as requirement of high flow oxygen, non-invasive or invasive ventilation were not significantly different

Table 1 Clinical characteristics of the study cohort at admission to hospital and outcomes

Variables	Total (n=4645)
Men	3386 (72.9)
Women	1259 (27.1)
Age (mean, years)	45.9±18.0
Age groups	
<30	1125 (24.2)
30–49	1397 (30.1)
50–69	1650 (35.5)
70+	473 (10.2)
Family members/house	
1–4	2395 (51.2)
5–9	2000 (42.8)
≥10	281 (6.0)
Educational status	
Illiterate or up to primary education	1424 (30.5)
Secondary school and/or higher secondary education	1538 (32.9)
Some college	1667 (35.7)
Tobacco or smoking (ever)	1369 (29.5)
Medical comorbidities	1335 (28.6)
Hypertension	831 (17.8)
Pulmonary disease	193 (4.1)
Type 2 diabetes	777 (16.6)
Thyroid disease	38 (0.8)
Heart disease	75 (1.6)
Neurological disease	15 (0.3)
Current or past tuberculosis	106 (2.3)
Duration of symptoms at admission (days)	
Clinical findings	
Pulse rate/min	83.9±11.4
Systolic BP mm Hg	125.4±12.2
Diastolic BP mm Hg	82.8±8.1
Respiratory rate/min	19.0±3.7
SpO$_2$ at admission	
≥95%	2144 (70.0)
90%–94%	561 (18.3)
<90%	357 (11.7)
Laboratory Investigations (biochemistry n=867; haematology n=4456)	
Creatinine, mg/dL	0.95+0.50
SGOT, IU	44.9±96.5
SGPT, IU	43.4±56.2
Sodium, mEq/L	136.1±12.5
Potassium, mEq/L	5.4±1.1
Haemoglobin (gm/dL)	12.7±2.3
White cells (10^9 cells/L)	7527±3830

Continued

Table 1	Continued	
Variables		**Total (n=4645)**
Lymphocytes (10^9 cells/L)		1589±1325
Lymphocyte/neutrophil ratio		0.36±0.32
Outcome measures		
Mean duration of hospital stay (days)		6.8±3.7
Oxygen requirement		861 (18.4)
High flow O_2/non-invasive ventilation		334 (7.1)
Mechanical ventilation		169 (3.6)
Recovered		4217 (90.2)
Referred		119 (2.5)
Deaths		340 (7.3)

Numbers ± indicate 1 SD; Numbers in parentheses are percent.
BP, blood pressure; SGOT, serum glutamic oxalate transferase;
SGPT, serum glutamic pyruvate transferase; SpO_2, saturation of
peripheral oxygen.

(online supplemental table 1). Number of in-hospital deaths were significantly greater in men (n=282, 8.3%) as compared with women (n=58, 4.6%) (p<0.001).

The cohort was divided into the three groups based on educational status. Important demographic and clinical characteristics and in-hospital outcomes are shown in table 2. Low educational status (groups 1 and 2) was more common in women, while more men had college education. Family size was larger among the less literate group. Tobacco use and smoking was also greater in group 1. Prevalence of comorbidities, especially hypertension and diabetes, was significantly greater among the more literate, similar to previous studies in India.[22] No significant differences were observed in complaints or clinical findings (data not shown). Data on duration of illness prior to admission were not available. Low SpO_2 (<90% as well as <95%), lymphopaenia, higher transaminases and higher creatinine values at admission were observed among the less literate. The length of hospital stay was not significantly different in the three groups.

Univariate ORs (categorical variables) and mean differences (continuous variables) in less literate groups 1 and 2 compared with the more literate group 3 are shown in table 3. Patients in less literate groups were younger, more women and lived in larger households (>10 persons/house). Presence of tobacco use was greater, while cardiovascular risk factors were lower. Various clinical outcomes are shown in figure 1 and compared with group 3, in group 1 there was greater oxygen requirement (unadjusted OR 1.34, 95% CI 1.12 to 1.61), non-invasive ventilation (1.14, 95% CI 0.87 to 1.49) and invasive ventilation (1.54, 95% CI 1.06 to 2.23) (table 3). Compared with group 3 (deaths n=92, 5.5%), deaths were significantly greater in group 1 (n=143, 10.0%, unadjusted OR 1.91, 95% CI 1.46 to 1,51) as well as in group 2 (n=104, 6.8%, unadjusted OR 1.24, 95% CI 0.93 to 1.66) (p<0.001).

We performed a stepwise logistic regression analysis to identify influence of various sociodemographic, risk factor, clinical and treatment variables on outcomes. Compared with the most literate group 3, unadjusted OR for deaths were higher in less literate groups 1 and 2 (table 4). Following adjustments for age, sex, household size, risk factors and comorbidities, the ORs attenuated but remained significant in both group 1 (1.44, 95% CI 1.07 to 1.93) and group 2 (1.38, 95% CI 1.02 to 1.85). However, after addition of clinical features at admission and laboratory investigations, the risks attenuated to marginally significant in group 1 (1.39, 95% CI 0.99 to 1.93) and significant in group 2 (1.53, 95% CI 1.10 to 2.11) and remain the same after further adjustments for oxygenation (table 3). OR for other outcomes assessed in the cohort (need for invasive ventilation and non-invasive ventilation) are shown in table 4 and demonstrate a marginal significance.

DISCUSSION

This study shows that illiterate and less literate patients with COVID-19 have significantly greater in-hospital mortality compared with the better educated. The higher risk of death among the less literate persists after adjustment for various sociodemographic factors (age, sex and household size), lifestyle factors and comorbidities but attenuates after adjustment for clinical features at presentation, investigations and oxygen treatment. This suggests that more adverse features at presentation (hypoxia, deranged liver and kidney functions) could be responsible for higher deaths among the less educated patients with COVID-19 in India.

Clinical and epidemiological studies from most developed countries in Europe and North America have consistently reported higher communicable disease-related mortality among the less literate and lower socioeconomic individuals.[11] In the COVID-19 pandemic, studies from most developed countries have reported greater COVID-19 related mortality and adverse outcomes among the ethnic minorities.[3–5] However, association of mortality among low socioeconomic or less educational status individuals are inconclusive.[3 4 12–14] In England, OpenSAFELY platform evaluated ethnic differences in COVID-19 related hospitalisation, intensive care unit admission and death in 17 million adults from the National Health Service.[23] As compared with the British white patients, deaths were higher in South Asians in the first wave (OR 1.08, 95% CI 1.07 to 1.09), and the second wave of COVID-19 epidemic (OR 1.87, 95% CI 1.68 to 2.07) as well as in the overall cohort (OR 1.26, 95% CI 1.15 to 1.37). Deaths were the highest in the most deprived groups.[23] A study from Brazil reported that those with low education attainment were more likely to die from COVID-19 (OR 1.13, 95% CI 1.07 to 1.19).[24] Increased deaths among the poor and low educational status patients has also been reported in recent studies from USA,[25] South Korea[26] and African countries.[27] An epidemiological study in

Table 2 Clinical characteristics and outcomes according to educational status (group 1=<primary education; group 2=>primary to higher secondary education; group 3=some college)

Variables	Group 1 (n=1424)	Group 2 (n=1538)	Group 3 (n=1667)	χ2 test or ANOVA (p value)
Age groups (years)				<0.001
<30	353 (25.0)	414 (27.1)	348 (21.0)	
30–49	416 (29.4)	459 (30.0)	510 (30.8)	
50–69	509 (36.0)	501 (32.8)	620 (37.4)	
70+	136 (9.6)	154 (10.1)	179 (10.8)	
Age mean (years)	45.8±17.9	44.6±18.4	47.1±17.6	<0.001
Men	980 (29.0)	1061 (31.4)	1339 (39.6)	<0.001
Women	444 (35.5)	477 (38.2)	328 (26.3)	
Members/house				<0.001
1–4	710 (49.9)	769 (50.0)	893 (53.6)	
5–9	624 (43.8)	652 (42.4)	703 (42.2)	
≥10	90 (6.3)	117 (7.6)	71 (4.3)	
Tobacco or smoking	496 (34.6)	485 (31.5)	375 (22.5)	<0.001
Medical comorbidities	391 (27.5)	411 (26.7)	531 (31.9)	0.002
Hypertension	248 (17.4)	218 (14.2)	365 (21.9)	0.000
Pulmonary disease	44 (3.1)	59 (3.8)	89 (5.3)	0.006
Type 2 diabetes	220 (15.4)	232 (15.1)	325 (19.5)	0.001
Thyroid disease	18 (1.3)	13 (0.8)	7 (0.4)	0.034
Coronary heart disease	18 (1.3)	20 (1.3)	36 (2.2)	0.074
Clinical findings at admission				
Systolic BP mm Hg (mean±SD)	125.4±12.6	124.7±11.6	125.9±12.4	0.021
Respiratory rate (mean±SD)	19.1±3.7	18.9±3.5	19.1±3.9	0.225
SpO_2 <90%	173 (12.1)	165 (10.7)	168 (10.1)	0.765
SpO_2 90%–94%	273 (19.2)	272 (17.7)	270 (16.2)	0.312
Investigations (mean+SD)				
Haemoglobin, g/dL	12.8±2.2	12.6±2.4	12.7±2.2	0.056
White cells, 10^9 cells/L	7559±3917	7611±3759	7419±3832	0.340
Lymphocyte, 10^9 cells/L	1574±1269	1561±1187	1631±1489	0.282
Lymphocyte/neutrophil ratio	0.35±0.35	0.36±0.35	0.36±0.27	0.624
SGPT, units	46.9±72.6	43.0±50.1	38.5±29.2	<0.001
SGOT, units	50.8±143.5	43.6±59.02	38.3±26.9	<0.001
Sodium, mEq/L	135.1±15.3	136.4±10.9	136.6±11.1	0.002
Creatinine, mg/dL	0.96±0.57	0.90±0.51	0.90±0.32	<0.001
Clinical outcomes				
Oxygen requirement	308 (21.6)	257 (16.7)	284 (17.0)	<0.001
Non-invasive ventilation	114 (8.0)	91 (5.9)	118 (7.1)	0.582
Invasive ventilation	66 (4.6)	54 (3.5)	51 (3.1)	<0.001
In-hospital outcomes				
Recovered	1247 (87.6)	1400 (91.0)	1526 (91.5)	<0.001
Referred	34 (2.4)	34 (2.3)	49 (2.9)	0.582
Deaths	143 (10.0)	104 (6.8)	92 (5.5)	<0.001

$χ^2$ test used for categorical variables and analysis of variance for continuous variables.

Numbers ± indicate 1 SD; numbers in parentheses are per cent; ORs and 95% CI calculated for categorical variables; mean difference and 95% CI calculated for numerical variables.

ANOVA, analysis of variance; BP, blood pressure; SGOT, serum glutamic oxalate transferase; SGPT, serum glutamic pyruvate transferase; SpO_2, saturation of peripheral oxygen.

Table 3 ORs (categorical variables) or mean difference (continuous variables) and 95% CIs among groups 1 and 2 patients compared with group 3 (college education)

Variables	OR/mean difference (95% CI) group 2 versus 3	P value	OR/mean difference (95% CI) group 1 versus 3	P value
Age groups (years)				
<30	1.39 (1.18 to 1.64)	0.0001	1.25 (1.05 to 1.47)	0.0083
30–49	0.97 (0.83 to 1.12)	0.6229	0.93 (0.80 to 1.09)	0.3980
50–69	0.81 (0.70 to 0.94)	0.0065	0.94 (0.81 to 1.08)	0.4211
70+	0.85 (0.67 to 1.06)	0.5178	0.87 (0.69 to 1.11)	0.2729
Age mean (years)	1.84 (0.69 to 2.99)	0.0001	−5.92 (−7.10 to −4.69)	0.0423
Men	0.54 (0.46 to 0.64)	<0.0001	0.54 (0.45 to 0.64)	<0.0001
Women	1.83 (1.56 to 2.15)	<0.0001	1.85 (1.57 to 2.18)	<0.0001
Members/house				
1–4	0.87 (0.75 to 0.99)	0.0416	0.86 (0.75 to 0.99)	0.0402
5–9	1.00 (0.87 to 1.16)	0.9089	1.07 (0.93 to 1.23)	0.3704
≥10	1.85 (1.37 to 2.51)	0.0001	1.52 (1.10 to 2.08)	0.0127
Tobacco or smoking	1.58 (1.35 to 1.85)	<0.0001	1.79 (1.52 to 2.09)	<0.0001
Medical comorbidities	0.78 (0.67 to 0.91)	0.0012	0.81 (0.69 to 0.95)	0.0077
Hypertension	0.59 (0.49 to 0.71)	<0.0001	0.75 (0.63 to 0.90)	0.0018
Pulmonary disease	0.71 (0.50 to 0.99)	0.0344	0.56 (0.39 to 0.82)	0.0026
Type 2 diabetes	0.73 (0.61 to 0.88)	0.001	0.75 (0.62 to 0.91)	0.0029
Thyroid disease	2.02 (0.80 to 5.08)	0.1403	3.04 (1.26 to 07.29)	0.0055
Coronary heart disease	0.59 (0.34 to 1.03)	0.0535	0.58 (0.33 to 1.03)	0.0597
Clinical findings				
Systolic BP mm Hg (mean±SD)	1.21 (0.37 to 2.03)	0.0248	0.51 (−0.38 to 1.38)	0.2674
Respiratory rate (mean+SD)	0.20 (−0.05 to 0.45)	0.1278	0.00 (−0.26 to 0.27)	1.0000
SpO_2 <90%	1.07 (0.85 to 1.34)	0.5781	1.23 (0.98 to 1.55)	0.0768
SpO_2 90%–94%	1.11 (0.92 to 1.33)	0.2579	1.22 (1.02 to 1.45)	0.0290
Investigations (mean+SD)				
Haemoglobin, g/dL	0.10 (−0.05 to 0.26)	0.2185	−0.1 (−0.25 to 0.05)	0.2079
White cells, 10^9 cells/L	−192 (−455 to 71)	0.1528	−140 (−414 to 134)	0.3163
Lymphocyte, 10^9 cells/L	70 (−23 to 163)	0.1433	57 (−41 to 155)	0.2566
Lymphocyte/neutrophil ratio	0.00 (−0.02 to 0.02)	1.00	0.01 (−0.01 to 0.03)	0.3705
SGPT, units	−4.5 (−7.3 to −1.7)	0.0017	−8.4 (−12.2 to −4.6)	<0.0001
SGOT, units	−5.3 (−8.4 to −2.2)	0.0009	−12.5 (−19.5 to −5.5)	<0.0001
Sodium, mEq/L	0.29 (−0.47 to 1.05)	0.7193	1.61 (0.67 to 2.53)	0.0017
Creatinine, mg/dL	0.00 (−0.02 to 0.02)	1.00	−0.06 (−0.09 to −0.02)	0.0002
Clinical outcomes				
Oxygen requirement	0.97 (0.81 to 1.17)	0.8207	1.34 (1.12 to 1.61)	0.0012
Non-invasive ventilation	0.82 (0.62 to 1.09)	0.1694	1.14 (0.87 to 1.49)	0.3442
Invasive ventilation	1.15 (0.78 to 1.70)	0.5261	1.54 (1.06 to 2.23)	0.0295
In-hospital outcomes				
Recovered	0.94 (0.73 to 1.19)	0.6166	0.65 (0.51 to 0.82)	0.0004
Referred	0.75 (0.47 to 1.16)	0.2874	0.81 (0.52 to 1.26)	0.3901
Deaths	1.24 (0.93 to 1.66]	0.1252	1.91 (1.46 to 2.51]	<0.0001

Univariate logistic regression used for categorical variables and unpaired t-test for continuous variables.

Numbers ± indicate 1 SD; numbers in parentheses are per cent; mean difference and 95% CI calculated for numerical variables and ORs and 95% CI for categorical variables.

BP, blood pressure; SGOT, serum glutamic oxalate transaminase; SGPT, serum glutamic pyruvate transaminase.
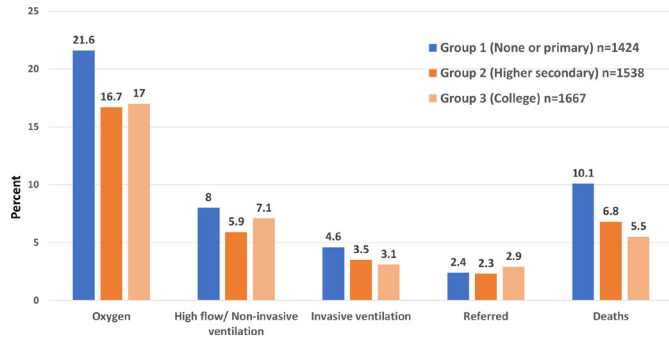
Figure 1 Clinical outcomes in various educational status groups.

Santiago, Chile, reported a strong association between socioeconomic status and mortality, measured either by COVID-19 attributed deaths or excess deaths with greater case fatality rates in the young patients with COVID-19 in deprived localities.[28] A large meta-analysis that combined population and hospital based data in the USA involving 4.3 million patients from 68 studies reported that disease incidence was more in African-American and Hispanic-American individuals, while risk of hospitalisation was greater in Asian Americans. Mortality rates in Hispanics and Asian Americans correlated positively with residence in more deprived locations.[29] In this study, influence of individual-level socioeconomic factors was not reported. Our study is one of the first reports from India that has evaluated socioeconomic difference in COVID-19 related mortality and shows a 1.4 to 1.9-fold greater mortality among low educational status men and women and is similar to the recent international studies. Our study also shows that greater mortality among low educational status individuals could be due to delayed presentation and more severe disease (lower oxygen, more impaired liver and renal functions) and greater need of oxygen and non-invasive and invasive ventilation in these patients (table 2). We did not obtain exact information regarding use of various non-evidence based empirical therapies (hydroxychloroquine, ivermectin, lopinavir-ritonavir, favipiravir, etc)[30] or proven evidence-based therapies such as corticosteroids, remdesivir and tocilizumab,[31] and this is a study limitation.

A variety of approaches to conceptualisation and measurement of socioeconomic status have been used. Four measures are consistently associated with greater risk: low education, low income, lower employment status and neighbourhood socioeconomic factors.[32] Use of self-reported educational status as marker of level of highest education achieved is similar to previous studies.[8–10] Low education or socioeconomic status is well known as a leading modifiable risk factor for overall as well as infectious disease mortality and is an important social determinant of health.[33] Our previous studies in India and other low and lower middle income countries have reported strong correlation of self-reported educational status with measures of income, household wealth, occupation, etc.[34 35] There are multiple social, clinical and system level

Table 4 Stepwise multivariate logistic regression analyses and OR (95% CIs) for adverse outcomes in educational status group 1 (<primary education) and group 2 (>primary to higher secondary education) compared with group 3 (some college)

	Educational status groups (reference 3)	Unadjusted ORs	Age and sex adjusted	Plus household size	Plus risk factors, comorbidities	Plus clinical factors, investigations	Plus oxygenation
Deaths	Group 1	1.91 (1.46 to 2.51)	1.33 (0.99 to 1.83)	1.37 (1.01 to 1.83)	1.44 (1.07 to 1.93)	1.39 (0.99 to 1.93)	1.38 (0.99 to 1.93)
	Group 2	1.24 (0.93 to 1.66)	1.31 (0.91 to 1.82)	1.32 (0.98 to 1.78)	1.38 (1.02 to 1.85)	1.53 (1.10 to 2.11)	1.52 (1.01 to 2.11)
Invasive ventilation	Group 1	1.54 (1.06 to 2.23)	1.19 (0.80 to 1.81)	1.21 (0.81 to 1.79)	1.29 (0.86 to 1.92)	1.34 (0.86 to 2.11)	1.39 (0.88 to 2.19)
	Group 2	1.15 (0.78 to 1.70)	1.06 (0.71 to 1.59)	1.07 (0.71 to 1.60)	1.11 (0.74 to 1.67)	1.31 (0.84 to 2.04)	1.33 (0.85 to 2.07)
Non-invasive ventilation	Group 1	1.14 (0.87 to 1.49)	0.95 (0.71 to 1.32)	0.96 (0.72 to 1.27)	1.03 (0.77 to 1.36)	0.79 (0.56 to 1.12)	0.78 (0.54 to 1.13)
	Group 2	0.82 (0.62 to 1.09)	1.01 (0.76 to 1.33)	1.00 (0.76 to 1.33)	1.02 (0.77 to 1.35)	0.88 (0.63 to 1.22)	0.91 (0.64 to 1.29)

Data are in ORs and 95% CIs.

contributors that lead to greater disease risk among the poor and include structural barriers to good health, particularly among the less literate and poor, increased risk of exposure (unhygienic working conditions and crowded housing), unequal access to testing and high-quality care, higher rates of associated medical conditions and less access to vaccination.[7 36] In the present study, we observed some of these barriers among our patients (crowded housing, greater tobacco use and delayed presentation with more severe disease). COVID-19 in India could act as a catalyst to improve overall healthcare systems with opportunities for policymakers, advocacy groups and researchers for evaluation of various interventions.[37] It is hoped that COVID-19 would lead to global focus on creation of health equity by influencing and coaxing politicians towards the right direction.[38]

The study has strengths as well as limitations. This is the largest case series from India; we used data from a government hospital that is more representative of general population; there are substantial number of less literate patients reflecting local educational status. This has led to data granularity and robust evaluation of outcomes. We used self-reported educational status to determine the highest level of literacy achieved and this is a study limitation; however, most of the previous studies have used similar methods.[8–10] Other limitations include lack of other sociodemographic factors (housing, neighbourhoods, occupation, income, working conditions, etc), clinical parameters (detailed history, pulmonary findings, radiological evaluation, chest computerised tomographic scans and blood biomarkers: C reactive protein, interleukins, d-Dimer, ferritin, lactic dehydrogenase, etc) and type of therapy the patients received. We also did not evaluate cardiovascular biomarkers (troponins and N-terminal probrain natriuretic peptide) that are important in prognostication.[29] These are due to lack of guidelines regarding routine measurement of many of these variables[21] and low healthcare funding in the country.[1] There could be multiple causes of deaths in COVID-19 (acute respiratory distress syndrome, myocardial infarction, acute heart failure, pulmonary embolism, secondary chest infection, sepsis, acute renal failure, etc),[29] and we did not have data on specific causes of death. About 2.5% persons were transferred from our hospital to other centres, and although we have obtained information on death in these patients using telephonic interview with families, details of specific outcomes are not available. Finally, data from a single hospital with about 4500 patients and 340 deaths may not be applicable to the whole country, which has one of the largest burden of COVID-19 in the world.[16] In view of the massive second wave of COVID-19 in India,[39] we should strive for larger multicentric studies for identifying reasons for greater mortality among the low socioeconomic status patients with this disease in the country.

In conclusion, our study shows a significantly greater mortality from COVID-19 in less educated (lower socioeconomic status) individuals in India. Khalatbari-Soltani *et al*[40] have suggested that low educational status is associated with increased prevalence of smoking and poor nutrition leading to more severe disease, prevalence of comorbidities is high in these individuals and low health literacy results in increased disease incidence and severity due to poor understanding pf public health preventive measures and delayed healthcare-seeking behaviours. Our study shows that the less educated patients with COVID-19 have more severe disease at presentation to hospital with need for greater oxygen and ventilatory support. Strategies to increase early diagnosis and access to care for these patients are important and should include public health measures for early detection of disease and early referral to treatment centres for appropriate therapeutic measures.

ORCID iD
Rajeev Gupta http://orcid.org/0000-0002-8356-3137

REFERENCES
1 Cash R, Patel V. Has COVID-19 subverted global health? *Lancet* 2020;395:1687–8.

2 Dawood FS, Ricks P, Njie GJ, *et al*. Observations of the global epidemiology of COVID-19 from the prepandemic period using web-based surveillance: a cross-sectional analysis. *Lancet Infect Dis* 2020;20:1255–62.

3 Upshaw TL, Brown C, Smith R, *et al*. Social determinants of COVID-19 incidence and outcomes: a rapid review. *PLoS One* 2021;16:e0248336.

4 Sze S, Pan D, Nevill CR, *et al*. Ethnicity and clinical outcomes in COVID-19: a systematic review and meta-analysis. *EClinicalMedicine* 2020;29-30:100630.

5 Treweek S, Forouhi NG, Narayan KMV, *et al*. COVID-19 and ethnicity: who will research results apply to? *Lancet* 2020;395:1955–7.

6 Egede LE, Walker RJ. Structural Racism, Social Risk Factors, and Covid-19 - A Dangerous Convergence for Black Americans. *N Engl J Med* 2020;383:e77.

7 Lavizzo-Mourey RJ, Besser RE, Williams DR. Understanding and Mitigating Health Inequities - Past, Current, and Future Directions. *N Engl J Med* 2021;384:1681–4.

8 Leon DA, Poverty WG. *Poverty, inequality and health: an international perspective*. Oxford: Oxford University Press, 2001.

9 Marmot M, Wilkinson R. *Social determinants of health*. Oxford: Oxford University Press, 2005.

10 Gupta R, Joseph P, Rosengren A, *et al*. Location and level of care, education, availability of medicines and cardiovascular mortality. In: Fuster V, Narula J, Vaishnava P, eds. *Hurst's The Heart*. 15th Ed. New York: McGraw Hill, 2022.

11 Bollyky TJ. *Plagues and the paradox of progress*. Cambridge, MA: The MIT Press, 2018.

12 Gao Y-D, Ding M, Dong X, *et al*. Risk factors for severe and critically ill COVID-19 patients: a review. *Allergy* 2021;76:428–55.

13 Liao TF, De Maio F. Association of social and economic inequality with coronavirus disease 2019 incidence and mortality across US counties. *JAMA Netw Open* 2021;4:e2034578.

14 Clouston SAP, Natale G, Link BG. Socioeconomic inequalities in the spread of coronavirus-19 in the United States: a examination of the emergence of social inequalities. *Soc Sci Med* 2021;268:113554.

15 Mackey K, Ayers CK, Kondo KK, *et al*. Racial and Ethnic Disparities in COVID-19-Related Infections, Hospitalizations, and Deaths : A Systematic Review. *Ann Intern Med* 2021;174:362–73.

16 Ritchie H, Ortiz-Ospina E, Beltekian D. India: coronavirus pandemic country profile. Available: https://ourworldindata.org/coronavirus/country/india [Accessed 5 May 2021].

17 Gaur K, Khedar RS, Mangal K, *et al*. Macrolevel association of COVID-19 with non-communicable disease risk factors in India. *Diabetes Metab Syndr* 2021;15:343–50.

18 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.

19 Sharma AK, Ahmed A, Baig VN, et al.. Characteristics and outcomes of hospitalized young adults with mild to moderate COVID-19 at a university hospital in India. *J Assoc Physicians India* 2020;68:62–5.

20 Sharma S, Sharma AK, Dalela G, et al.. Association of SARS CoV-2 cycle threshold (CT) with clinical outcomes: a hospital-based study. *J Assoc Physicians India* 2021;69:86–90.

21 Government of India,, Ministry of Health and Family Welfare. Clinical management protocol: COVID-19. Available: http://www.rajswasthya.nic.in/PDF/COVID%20-19/FOR%20HOSPITALS/27.06.2020.pdf [Accessed 30 Apr 2021].

22 Gupta R, Gaur K, Ram CVS. Emerging trends in hypertension epidemiology in India. *J Hum Hypertens* 2019;33:575–87.

23 Mathur R, Rentsch CT, Morton CE, *et al*. Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. *Lancet* 2021;397:1711–24.

24 Li SL, Pereira RHM, Prete CA, *et al*. Higher risk of death from COVID-19 in low-income and non-white populations of São Paulo, Brazil. *BMJ Glob Health* 2021;6:e004959.

25 Azar KMJ, Shen Z, Romanelli RJ, *et al*. Disparities in outcomes among COVID-19 patients in a large health care system in California. *Health Aff* 2020;39:1253–62.

26 Oh TK, Choi J-W, Song I-A. Socioeconomic disparity and the risk of contracting COVID-19 in South Korea: an NHIS-COVID-19 database cohort study. *BMC Public Health* 2021;21:e144.

27 Salyer SJ, Maeda J, Sembuche S, *et al*. The first and second waves of the COVID-19 pandemic in Africa: a cross-sectional study. *Lancet* 2021;397:1265–75.

28 Mena GE, Martinez PP, Mahmud AS, *et al*. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science* 2021;372:eabg5298.

29 Magesh S, John D, Li WT, *et al*. Disparities in COVID-19 outcomes by race, ethnicity, and socioeconomic status: a Systematic-Review and meta-analysis. *JAMA Netw Open* 2021;4:e2134147.

30 Siemieniuk RA, Bartoszko JJ, Ge L, *et al*. Drug treatments for covid-19: living systematic review and network meta-analysis. *BMJ* 2020;370:m2980.

31 RECOVERY: randomized evaluation of COVID-19 therapy. news. Oxford. Nuffield department of population health, 2021. Available: https://www.recoverytrial.net/news [Accessed 6 May 2021].

32 Braveman P, Egerter S, Williams DR. The social determinants of health: coming of age. *Annu Rev Public Health* 2011;32:381–98.

33 The Lancet Public Health . Education: a neglected social determinant of health. *Lancet Public Health* 2020;5:e361.

34 Gupta R, Gupta VP, Ahluwalia NS. Educational status, coronary heart disease, and coronary risk factor prevalence in a rural population of India. *BMJ* 1994;309:1332–6.

35 Gupta R, Kaur M, Islam S, *et al*. Association of household wealth index, educational status, and social capital with hypertension awareness, treatment, and control in South Asia. *Am J Hypertens* 2017;30:hpw169–381.

36 Smedley BD, Syme SL. *Promoting health: intervention strategies from social and behavioral research*. Washington: National Academy Press. Institute of Medicine, 2000.

37 Gupta R. Health systems in Post-Covid-19 era: strengthening primary care and district hospital. *RUHS J Health Sciences* 2020;5:61–5.

38 Williams DR, Cooper LA. COVID-19 and Health Equity-A New Kind of "Herd Immunity". *JAMA* 2020;323:2478–80.

39 Gupta R, Gaur K, Katoch VM. SARS-CoV-2 variants of concern and changing trends of COVID-19 in India and Rajasthan: epidemiological perspective. *RUHS J Health Sciences* 2021;6:141–5.

40 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 Epidem Comm Health* 2020;74:620–3.