THE LANCET
Global Health

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Jassat W, Mudara C, Ozougwu L, et al. Difference in mortality among individuals admitted to hospital with COVID-19 during the first and second waves in South Africa: a cohort study. *Lancet Glob Health* 2021; published online July 9. http://dx.doi.org/10.1016/S2214-109X(21)00289-8.

Manuscript reference number: LANGLH-D-21-00575R2

Title: Difference in mortality among individuals admitted to hospital with COVID-19 during the first and second waves in South Africa: a cohort study

Author: Dr Waasila Jassat

Supplementary Materials

Supplementary Methods

SUPPLEMENTARY METHODS

Study population: DATCOV contains a field to capture reason for admissions but it is not well completed. DATCOV includes patients who had COVID-19 symptoms, were admitted for isolation, acquired nosocomial COVID-19 infection, or tested positive incidentally when admitted for other reasons. We are not able to determine whether incidental positive tests would have been more likely due to routine testing for all hospital admissions.

Data collection: DATCOV is the first electronic information system introduced to public hospitals in most provinces. There has not previously been a mandate or culture for clinicians or administrative staff to report on patient level information. When DATCOV was introduced to a few hospitals, a handful of medical doctors assumed responsibility for submitting data on the platform. During the first wave when the hospitals were busy, data capturers were required to enter data using the patient record. These data capturers may not have been able to extract all required information from the patient record, especially those related to comorbidities, complications and clinical treatment. Data completeness dropped due to the constraints on time during the surge of the wave and due to being completed by data capturers with less clinical knowledge. It is also true that medical records are poorly completed in the public sector as there are no billing implications and because there are inconsistent clinical governance efforts to address this. While DATCOV was developed as an online platform for direct data entry, when it was observed that the private hospital groups and the Western Cape public sector had available data from electronic information systems, DATCOV adapted to allow for data imports. It was difficult to match all fields consistently to DATCOV so a compromise was reached and only the minimum required fields were imported (patient identifiers, age, sex, race, pregnancy, HCW status, comorbidities, and details of the admission such as date of admission, ward of care, oxygen and ventilation required and the outcome and date of outcome). The remaining variables are therefore unknown for all patients from the private sector and Western Cape public sector, which is half the patients admitted.

Variables included: Race was self-defined by the individual as black, white, mixed (or Coloured) and Indian, and captured on DATCOV by the reporting hospital. Data completeness was poorer during the peaks of the wave and there were more individuals captured with unknown race in the second wave. This is because the hospitals were busier and data was being submitted by data capturers who were less likely to get access to all patient clinical records and to extract data completely from patient records.

Comorbidities included in the analysis were hypertension, diabetes, asthma, malignancy, and chronic pulmonary, cardiac and renal diseases, HIV and current and past TB. We included the comorbidities strongly suggested to be related to severe COVID-19 and did not include others such as Down Syndrome and other intellectual disabilities that have been subsequently reported to be associated with higher mortality. We included past TB within the previous 10 years, because it has been suggested that lung damage from previous TB could result in more severe pulmonary disease with COVID-19 (https://theunion.org/our-work/covid-19/covid-19-and-tb-frequently-asked-questions). Comorbidities recorded in the patients' clinical record were extracted by data capturers into DATCOV. The presence of comorbidities was not verified by the DATCOV team and no national information system exists that contains individuals' comorbidity information. To create the combined comorbidity variable, if a patient was recorded as having one or more of the listed comorbidities they were captured as "Yes", if they were recorded as having none or nil known they were recorded as "No". The presence or absence of any comorbidity was unknown in 22% of patients.

Obesity was not consistently recorded, and was either based on the calculation of body mass index or on the subjective opinion of the attending healthcare worker, which could result in misclassification. There were also 75% of records with missing Obesity data. We therefore excluded Obesity from the analysis.

Socio-economic variables are not collected on DATCOV as this data is not routinely collected in patients records and is not easily ascertainable or available through other routine data sources.

Reinfection: DATCOV does not consistently record data on previous COVID-19 infection as records are difficult to match between hospital admissions. Matching and linkage of patients has limitations due to lack of unique identifiers being consistently recorded, so we are also not able to systematically determine whether patients have had readmissions and repeat infections.

Sequencing is not currently routinely done in South Africa except for in a small number of research sites, and sequence data were not available for patients reported to DATCOV, or even for a subset of patients.

Data management: In terms of data quality assurance, on data entry most fields in DATCOV have drop-down menus or tick-boxes, and very little free text is captured. There are data validations included on DATCOV to ensure that certain data submission errors are avoided, e.g. dates not in the future, rules for minimum and maximum of weight, numerics only for certain fields, etc.

Data management: The DATCOV team do data import of private sector and Western Cape data daily, the import overwrites existing data where it has changed, e.g. outcome and date of outcome has now been added for a patient who was in hospital. Once all imports are completed, daily linelists and reports are produced and shared with the provinces and NDoH daily by 15h00 including weekends and public holidays. While the timelines for producing reports are demanding, the DATCOV team conduct a number of routine data quality assurance steps.

1. Database validation rules during import to check for errors in incomplete data, incorrect formatting of dates, data outliers such as age older than 100.
2. Routine audit of import file and database to ensure all fields correctly imported and updated.
3. Routine audit of summary data to look for outliers, e.g. long hospital stays over 3 weeks. The hospitals are contacted and required to update if these patients have recorded outcomes.
4. Routine matching with national SARS-CoV-2 laboratory case linelist, to match against DATCOV, assign laboratory case identifier, to ensure the admission did have a recorded positive diagnosis associated with the admission. Admissions that were not found on the case linelist are investigated by the hospital, then flagged as negative and removed from reports.
5. Routine matching with other sources of data where available, to obtain data that may be missing on DATCOV, e.g. Identity Number.

Data analysis: We used incidence of admissions to define wave period and since the second wave had a faster increase in admissions, the overall period for wave 2 was shorter than for wave 1. We do not however believe that this is a limitation of our approach. The CFR was calculated as number of deaths/ number of patients with outcomes (deaths plus discharges), which would have adjusted for the longer periods and greater admissions/deaths. In fact there were more admissions in wave 2, which was of shorter duration.

Supplementary Table 1. Factors associated with the second compared to the first COVID-19 wave among hospitalised individuals*, South Africa, 5 March 2020-27 March 2021. (complete case analysis showing univariate and multivariable analysis implemented on the unimputed dataset)

Characteristic	First wave n/N (%)	Second wave n/N (%)	Unadjusted OR	95% CI	Adjusted OR	95% CI	Adjusted OR	95% CI
					Unadjusted for weekly admissions		Adjusting for weekly admissions	
Age group								
<40 years	17158/70612 (24.3)	20624/100481 (20.5)	Reference		Reference		Reference	
40-64 years	36168/70612 (51.2)	50985/100481 (50.7)	**1.17**	1.10-1.25	**1.22**	1.12-1.33	**1.19**	1.09-1.30
≥65 years	17286/70612 (24.5)	28872/100481 (28.7)	**1.39**	1.26-1.53	**1.46**	1.30-1.65	**1.42**	1.26-1.61
Sex								
Female	39335/70494 (55.8)	55634/100221 (55.5)	Reference					
Male	31159/70494 (44.2)	44587/100221 (44.5)	1.01	0.98-1.05				
Race								
White	4260/51311 (8.3)	5459/62297 (8.8)	Reference		Reference		Reference	
Black	41455/51311 (80.8)	47899/62297 (76.9)	0.90	0.77-1.06	**0.41**	0.33-0.50	**0.40**	0.32-0.49
Mixed	2577/51311 (5.0)	4673/62297 (7.5)	**1.42**	1.17-1.71	1.14	0.87-1.50	1.13	0.87-1.48
Indian	3019/51311 (5.9)	4266/62297 (6.9)	1.10	0.91-1.33	**0.69**	0.52-0.90	**0.68**	0.51-0.89
Comorbid condition								
No co-morbidity	20818/46000 (45.3)	35220/64526 (54.6)	Reference		Reference		Reference	
Co-morbid condition	25182/46000 (54.7)	29306/64526 (45.4)	**0.69**	0.59-0.81	**0.69**	0.59-0.81	**0.69**	0.59-0.81
Health sector								
Private sector	36414/70612 (51.6)	47290/100481 (47.1)	Reference		Reference		Reference	
Public sector	34198/70612 (48.4)	53191/100481 (52.9)	**1.20**	1.02-1.40	**2.00**	1.45-2.74	**2.02**	1.46-2.78
Province								
Western Cape	11656/70612 (16.5)	22554/100481 (22.5)	Reference		Reference		Reference	
Eastern Cape	8753/70612 (12.4)	13440/100481 (13.4)	0.79	0.59-1.06	**3.06**	1.71-5.49	**3.17**	1.78-5.66
Free State	4463/70612 (6.3)	3360/100481 (3.3)	**0.39**	0.28-0.55	1.39	0.75-2.57	1.35	0.73-2.49
Gauteng	24610/70612 (34.9)	23364/100481 (23.3)	**0.49**	0.37-0.64	1.43	0.77-2.67	1.38	0.74-2.56
KwaZulu-Natal	12752/70612 (18.1)	23637/100481 (23.5)	0.96	0.72-1.28	**3.22**	1.78-5.83	**3.04**	1.68-5.49
Limpopo	1361/70612 (1.9)	5275/100481 (5.3)	**2.00**	1.42-2.83	**3.77**	1.85-7.69	**3.50**	1.72-7.13
Mpumalanga	2098/70612 (3.0)	3957/100481 (3.9)	0.97	0.66-1.44	**2.81**	1.20-6.56	**2.68**	1.13-6.37
North West	3939/70612 (5.6)	3772/100481 (3.8)	**0.49**	0.33-0.73	2.16	0.99-4.66	2.03	0.95-4.37

Northern Cape	980/70612 (1.4)	1122/100481 (1.1)	0.59	0.33-1.06	1.11	0.28-4.43	1.08	0.28-4.18
Weekly national admission number								
Low <3500	3317/70612 (4.7)	2997/100481 (3.0)	Reference				Reference	
Medium 3500-5749	18722/70612 (26.5)	13473/100481 (13.4)	**0.80**	0.65-0.97			0.82	0.64-1.05
High 5750-7999	13790/70612 (19.5)	14736/100481 (14.7)	1.18	0.94-1.48			0.96	0.71-1.29
Very high >8000	34783/70612 (49.3)	69275/100481 (68.9)	**2.20**	1.71-2.85			**1.68**	1.21-2.34

* Statistically significant estimates are highlighted in bold
OR=Odds Ratio; CI=Confidence Interval

Supplementary Table 2. Factors associated with in-hospital mortality among individuals hospitalised with COVID-19*, South Africa, 5 March 2020-27 March 2021. (univariate and multivariable analysis implemented on the unimputed dataset)

Characteristic	Case Fatality Risk n/N (%)	Unadjusted OR	95% CI	Adjusted OR	95% CI	Adjusted OR	95% CI
				Including wave periods		Adjusting for weekly admissions	
Age group (years)							
<40	3816/53210 (7.2)	Reference		Reference		Reference	
40-64	23646/108773 (21.7)	**3.75**	3.62-3.90	**3.43**	3.21-3.68	**3.41**	3.19-3.65
≥65	23575/57282 (41.2)	**9.38**	9.03-9.75	**8.64**	8.06-9.30	**8.60**	8.01-9.24
Sex							
Female	26466/121937 (21.7)	Reference		Reference		Reference	
Male	24490/96890 (25.3)	**1.32**	1.29-1.34	**1.32**	1.28-1.37	**1.32**	1.27-1.37
Race							
White	2689/12661 (21.2)	Reference		Reference		Reference	
Black	28602/114571 (25.0)	**0.77**	0.73-0.81	**1.18**	1.10-1.27	**1.17**	1.09-1.25
Mixed	2310/10053 (23.0)	**0.92**	0.85-0.99	**1.24**	1.13-1.37	**1.24**	1.12-1.36
Indian	2014/8591 (23.4)	**1.08**	1.00-1.17	**1.30**	1.12-1.44	**1.29**	1.17-1.42
Comorbid condition							
No co-morbidity	11705/70544 (16.6)	Reference		Reference		Reference	
Co-morbid condition	19295/70056 (27.5)	**2.06**	1.99-2.13	**1.55**	1.49-1.62	**1.55**	1.49-1.62
Health sector							
Private sector	19763/105409 (18.8)	Reference		Reference		Reference	
Public sector	31274/113856 (27.5)	**1.99**	1.70-2.33	**1.51**	1.29-1.75	**1.52**	1.30-1.77
Province							
Western Cape	9739/45341 (21.5)	Reference		Reference		Reference	
Eastern Cape	9334/28671 (32.6)	**2.32**	1.79-3.01	**1.83**	1.43-2.34	**1.83**	1.43-2.34
Free State	2627/11759 (22.3)	1.14	0.83-1.56	1.08	0.80-1.42	1.07	0.80-1.42
Gauteng	11938/59548 (20.1)	**0.76**	0.60-0.98	1.06	0.82-1.34	1.04	0.82-1.33
KwaZulu-Natal	10411/42712 (24.4)	**1.45**	1.13-1.86	**1.32**	1.04-1.68	**1.30**	1.02-1.66
Limpopo	2431/8110 (30.0)	**1.98**	1.44-2.72	1.22	0.88-1.68	1.20	0.87-1.65
Mpumalanga	2209/8287 (26.7)	**2.27**	1.62-3.20	1.33	0.91-1.90	1.31	0.91-1.90
North West	1651/11343 (14.6)	0.73	0.49-1.09	1.06	0.70-1.56	1.04	0.70-1.55

Northern Cape	697/3494 (20.0)	1.32	0.84-2.08	1.29	0.75-2.12	1.27	0.76-2.13
Wave period							
Pre-wave 1	1945/10833 (18.0)	**0.79**	0.75-0.84	**0.80**	0.73-0.88	**0.86**	0.76-0.97
Wave 1	14721/70612 (20.9)	Reference		Reference		Reference	
Pre-wave 2	3951/24024 (16.5)	**0.70**	0.68-0.73	**0.81**	0.76-0.89	**0.87**	0.77-0.97
Wave 2	27947/100481 (27.8)	**1.34**	1.31-1.37	**1.33**	1.26-1.38	**1.30**	1.24-1.36
Post-wave 2	2473/13315 (18.6)	**0.80**	0.76-0.84	0.97	0.84-1.11	1.06	0.94-1.20
Weekly national admission number							
Low <3500	8985/50840 (17.7)	Reference				Reference	
Medium 3500-5749	7541/35841 (21.0)	**1.25**	1.20-1.29			0.97	0.88-1.07
High 5750-7999	6562/28526 (23.0)	**1.44**	1.39-1.50			1.02	0.92-1.13
Very high >8000	27949/104058 (26.9)	**1.74**	1.69-1.79			**1.17**	1.06-1.29

* Statistically significant estimates are highlighted in bold
OR=Odds Ratio; CI=Confidence Interval

Supplementary Table 3. Multivariable analysis of factors associated with in-hospital mortality among individuals hospitalised with COVID-19, in public and private hospitals*, South Africa, 5 March 2020-27 March 2021. (analysis implemented on the imputed dataset)

Characteristic	Adjusted OR	95% CI	Adjusted OR	95% CI
	Private sector		Public sector	
Age group				
<40 years	Reference		Reference	
40-64 years	**3.74**	3.49-4.01	**2.99**	2.85-3.13
≥65 years	**10.50**	9.74-11.31	**6.84**	6.51-7.18
Sex				
Female	Reference		Reference	
Male	**1.32**	1.28-1.36	**1.28**	1.24-1.32
Race				
White	Reference		Reference	
Black	**1.25**	1.15-1.36	**1.17**	1.06-1.30
Mixed	**1.25**	1.13-1.39	**1.15**	1.01-1.31
Indian	**1.32**	1.21-1.43	**1.42**	1.21-1.67
Comorbid condition				
No co-morbidity	Reference		Reference	
Co-morbid condition	**1.59**	1.52-1.66	**1.76**	1.69-1.84
Province				
Western Cape	Reference		Reference	
Eastern Cape	**1.74**	1.37-2.21	**2.43**	1.75-3.39
Free State	**1.35**	1.06-1.73	1.33	0.87-2.02
Gauteng	1.11	0.95-1.30	0.97	0.65-1.44
KwaZulu-Natal	**1.26**	1.05-1.50	**1.70**	1.20-2.39
Limpopo	**1.53**	1.12-2.10	**1.97**	1.33-2.92
Mpumalanga	1.16	0.86-1.56	**2.83**	1.84-4.35
North West	0.91	0.70-1.19	0.95	0.55-1.66
Northern Cape	1.20	0.81-1.76	**1.81**	1.03-3.20
Wave period				
Pre-wave 1	**1.16**	1.02-1.32	**0.88**	0.79-0.97
Wave 1	Reference		Reference	

Post-wave 1	**0.85**	0.76-0.95	**0.84**	0.77-0.92
Wave 2	**1.52**	1.46-1.58	**1.15**	1.11-1.20
Post-wave 2	1.07	0.97-1.19	0.95	0.87-1.04
Weekly national admissions number				
Low <3500	Reference		Reference	
Medium 3500-5749	1.04	0.94-1.14	0.96	0.89-1.04
High 5750-7999	**1.14**	1.03-1.26	1.05	0.96-1.14
Very high >8000	**1.35**	1.22-1.48	**1.18**	1.09-1.27

* Statistically significant estimates are highlighted in bold
OR=Odds Ratio; CI=Confidence Interval