

Online Supplementary

A Data protocol

A.1 Data preparation

We use data from CHES case records reported on 26 May 2020 with $N = 16,138$ entries. There were 15,750 with unique ID cases. The cases with identical IDs and with more than one entry were examined, and the following actions were taken:

1. When there were identical values for ICU admission date, but the entries had different hospital admission dates, the record with the earliest hospital admission date was kept.
2. When there were identical values for ICU admission date, but the entries had the same hospital admission dates, the record with the earliest swab date was kept.
3. When there were different values for ICU admission date and the ICU periods were not contiguous, all the entries were kept.
4. When there were different values for the ICU admission date and the ICU periods were contiguous, a unique

case was created based on the earliest hospital and ICU admission dates and the latest date of leaving ICU.

5. When any of the cases does not have an ICU admission date, the record with a non-missing value in this item was kept.
6. When any of the cases had an ICU admission date, but had different hospital admission dates, the record with the earliest date in the last item was kept.
7. When any of the cases had an ICU admission date, but had identical hospital admission dates, the record with the earliest swab date was kept.

Amongst the cases with different IDs and with only one entry, we found a small number of records with identical values in the hospital and ICU admission dates (date, hours and minutes), sex, age (year and month), trust-code and postcode. Given these coincidences - we interpreted these records as duplicates and we only kept the record with the earliest estimated date of onset symptoms or swab date.

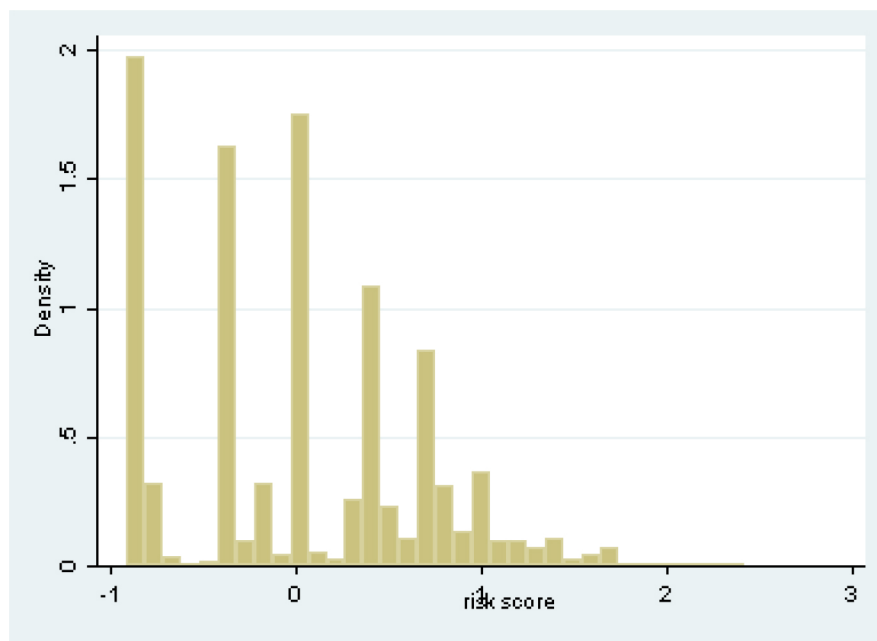
After the deduplication process, 493 records were removed from the initial dataset.

The deduplicated dataset was filtered out by records whose hospital admission date was before 16 March 2020 (beginning

Table A1: Cases removed which belonged to trusts with high levels of missingness per region

NHS region	Number of removed records
East of England	508
London	614
Midlands	488
North West	927
South East	27
Total	2,564

Figure A1: Histogram of the estimated standardised risk scores based on pre-existing health conditions/comorbidities



Source: own elaboration using CHES data.

Table A2: Results of the Rasch Partial Credit Model used to construct a score of pre-existing comorbid conditions and risk factors

Risk factor variable	Mean of variable	Available cases	Step parameter, b	Std. error of b
Chronic respiratory disease,excl. asthma	0.11	2,767	2.30	0.07
Asthma requiring medication	0.17	2,780	1.79	0.06
Chronic/congenital heart disease	0.14	2,744	2.06	0.06
Hypertension	0.42	2,863	0.32	0.05
Chronic renal disease	0.08	2,771	2.70	0.08
Chronic liver disease	0.02	2,618	4.04	0.13
Chronic neurological disease	0.05	2,728	3.24	0.09
Diabetes	0.31	2,748	0.91	0.05
Immunosuppression due to treatment	0.06	2,720	3.09	0.09
Immunosuppression due to disease	0.04	2,715	3.49	0.10
Obesity clinically apparent (maybe vs. not)	0.62	2,441	2.58	0.10
Obesity clinically apparent (yes vs. maybe)	0.62	2,441	-1.45	0.11

Source: own elaboration using CHES data. Step parameter b reflects the level of latent health risk required to have a 50% chance of reporting the risk factor.

of week 12) and after 24 May 2020 (end of week 21) ($n = 917$). Data before week 12 was excluded, given that the sample size was not enough for the current analysis ($n = 890$). Data after week 21 was omitted due to a lag of registering the complete information of each case. Therefore, we consider that the information of the most recent entries is unreliable.

We removed the records which had missing values or unknown sex, negative values in age or recorded age equal to 0, and unavailable information regarding whether a patient was admitted to ICU or not. This resulted in sample of 13,337 records.

From them, we removed 2,564 entries (Table A1). These records were reported by trusts whose missigness on predictor information is high.

From the remaining 10,773 registers, we selected dataset those who were admitted to ICU ($n=4,053$). Given that the procedure in ICU for under-18's is different, individuals who were younger than eighteen years old were also removed ($n=12$), leaving a sample of 4,041 cases.

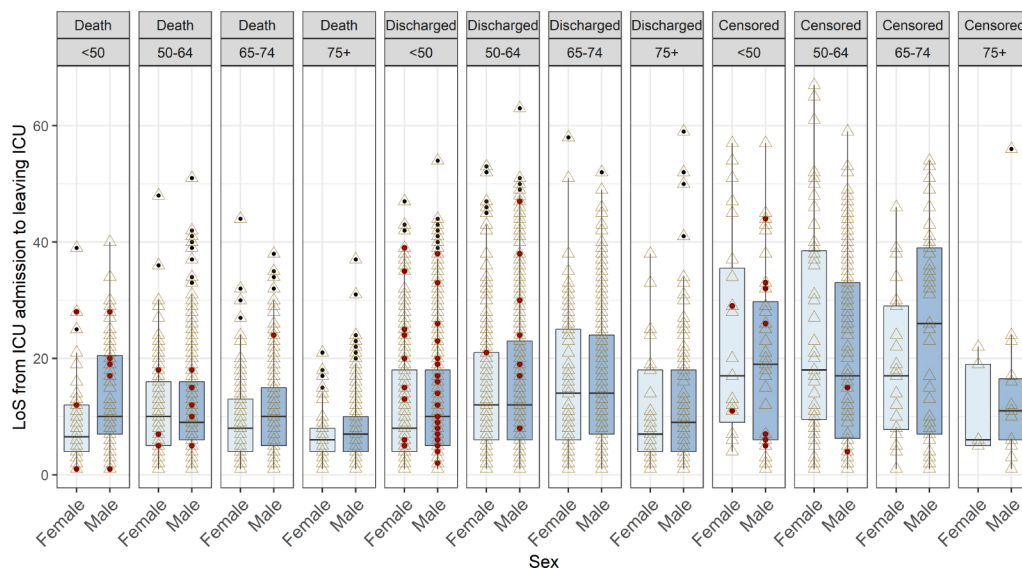
A.2 Assumptions on LoS variable

When the date of leaving ICU was the same as the final outcome date, it was assumed that the outcome took place while in ICU. The date of censoring was either the date of outcome "still on unit", or the date of last update of the record, or the date of swabbing for the SARS-CoV-2. All negative LoS were removed from the analysis.

A.3 Rasch risk factors scores

The results presented in Table A2 show the variables used and the key "step" parameters from the Rasch Partial Credit Model, which was implemented in Stata using `irt pcm` command. The fit of the model to the data was good. Principal Components Analysis of the model residuals showed no evidence of multi-dimensionality. The step parameters showed that hypertension, diabetes and asthma represented the lowest level of pre-existing health risk, and chronic liver

Figure B1: LoS from ICU admission to leaving ICU by final outcome, sex and age of patients



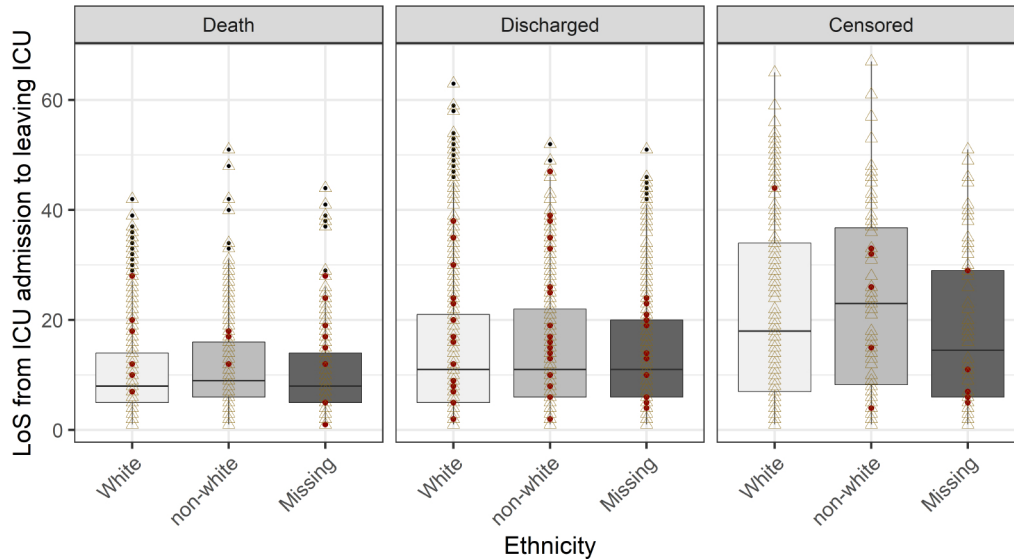
Note: red dots denote ECMO patients and yellow triangles refer to patients with Hospital-Acquired Infection (HAI). Last date of record update or swab date used for censoring. Source: own elaboration using CHES data and sample $n = 3,594$.

disease the highest. In Figure A1 we present a histogram of predicted risk scores. The distribution was right-skewed, with the majority of patients having low risk, but with a significant minority having much higher risk scores. For 545 cases it was impossible to assess whether missing information on the health-related risk factors was a genuine lack of thereof, or

information not recorded. Thus, we treated those cases as missing.

B Descriptive statistics of LoS from ICU admission to leaving ICU

Figure B2: LoS from ICU admission to leaving ICU by final outcome and ethnicity of patients



Note: red dots denote ECMO patients and yellow triangles refer to HAI (Hospital-Acquired Infection). Last date of record update or swab date used for censoring. Source: own elaboration using data and sample $n = 3,594$.

Table B1: Descriptive statistics of LoS from ICU admission to leaving ICU

Final outcome			Death or discharged					Censored					
Variable	Category	Total	Mean	SD	Median	Min	Max	Total	Mean	SD	Median	Min	Max
Sex	Female	833	12.6	10.0	10	1	58	87	24.6	17.0	20	1	67
	Male	1,923	13.2	10.4	10	1	63	206	22.3	15.8	21	1	59
Age groups	18-50	627	12.8	10.3	9	1	54	55	23.7	16.0	21	2	57
	50-64	1,161	14.1	10.7	11	1	63	161	23.6	16.4	22	1	67
	65-74	664	13.1	10.0	10	1	58	62	23.2	16.2	23	1	54
	>75	304	9.2	8.3	7	1	59	15	13.6	13.5	10	3	56
	White	1,619	12.8	10.3	10	1	63	160	22.9	16.1	20	1	65
Ethnicity	Non-white	655	13.6	10.7	10	1	52	80	24.5	16.6	24	1	67
	Missing	482	13.1	9.8	10	1	51	53	20.9	15.7	17	1	51
Rasch scores	<= -0.407	1,055	13.2	10.5	10	1	63	113	24.2	16.1	21	1	65
	(-0.407, 0.016]	597	13.9	10.9	10	1	59	62	26.1	16.9	27	1	67
	(0.016, 0.449]	412	13.1	9.7	10	1	53	48	23.7	16.5	24	1	57
	>0.449	692	12.1	9.7	9	1	52	70	17.7	14.7	13	1	59
ECMO patients	Yes	65	17.3	10.7	16	1	47	8	22.5	15.1	28	4	44
	Missing	2,691	12.9	10.3	10	1	63	285	23.0	16.3	20	1	67
HAI patients	Yes	514	19.3	11.6	18	1	59	66	27.0	15.7	28	1	59
	Missing	2,242	11.6	9.4	9	1	63	227	21.8	16.2	18	1	67
ICU admission	<25/03	359	14.5	11.4	11	1	63	28	30.0	21.9	36	2	67
	25/03 -07/04	1,370	13.8	10.7	10	1	59	82	30.7	17.2	34	1	59
	>07/04	1,027	11.5	9.1	9	1	47	183	18.5	12.8	17	1	47
Region	London	495	14.7	11.4	11	1	54	37	30.6	17.0	32	1	67
	South East	406	12.4	9.8	9	1	63	47	22.4	17.9	17.0	1	65
	South West	176	14.4	12.0	10	1	59	12	19.8	13.8	17.5	5	41
	East	210	13.4	10.8	11	1	52	32	20.5	16.0	18	1	51
	Midlands	492	13.7	9.7	12	1	53	57	21.7	16.0	16	1	54
	North East & Yorks	560	11.5	9.0	9	1	48	56	21.2	14.0	19	1	52
	North West	417	12.1	10.1	9	1	58	52	23.8	16.3	20	1	57
	Total	2,756	13.0	10.3	10	1	63	293	23.0	16.2	21	1	67

Source: own elaboration using CHES data. Note: last date of record update or SARS-Cov-2 swab date were assumed for censoring.

C Model results

Table C1: Weibull Accelerated Failure Time model with main effects

Variable	Coefficient	Robust SE	p
female	-0.03	0.03	0.345
admission period	(reference: after 07/04)		
before 25/03	0.17	0.06	0.008
25/03-07/04	0.09	0.04	0.035
age group	(reference: 50-64)		
<50	-0.16	0.06	0.005
65-74	-0.12	0.04	0.003
75+	-0.54	0.06	< 0.001
risk score	-0.07	0.03	0.026
ecmo	0.26	0.10	0.009
intercept	2.84	0.05	< 0.001
log(p)	0.23	0.02	< 0.001
n	3,049		
AIC	8,020.63		

Table C2: Predicted length-of-stay in days based on the Weibull model with main effects

Age	Measure	Admission period			Total
		<25/03	25/03-07/04	>07/04	
18-50	mean	16.1	15.0	14.1	14.7
	Median	13.0	12.0	11.3	11.8
	SD	1.2	1.3	1.5	1.5
	N	79	306	297	682
50-64	mean	18.8	17.2	15.8	16.8
	median	15.1	13.8	12.7	13.5
	SD	1.3	1.0	0.9	1.4
	N	141	646	535	1,322
65-74	mean	16.5	15.2	13.8	14.9
	median	13.3	12.2	11.1	11.9
	SD	0.9	0.7	0.7	1.2
	N	106	360	260	726
75+	mean	11.0	10.0	9.0	9.8
	median	8.8	8.1	7.3	7.9
	SD	0.5	0.5	0.4	0.8
	N	61	140	118	319
Total	mean	16.4	15.6	14.3	15.2
	median	13.2	12.5	11.5	12.2
	SD	2.8	2.3	2.2	2.4
	N	387	1,452	1,210	3,049

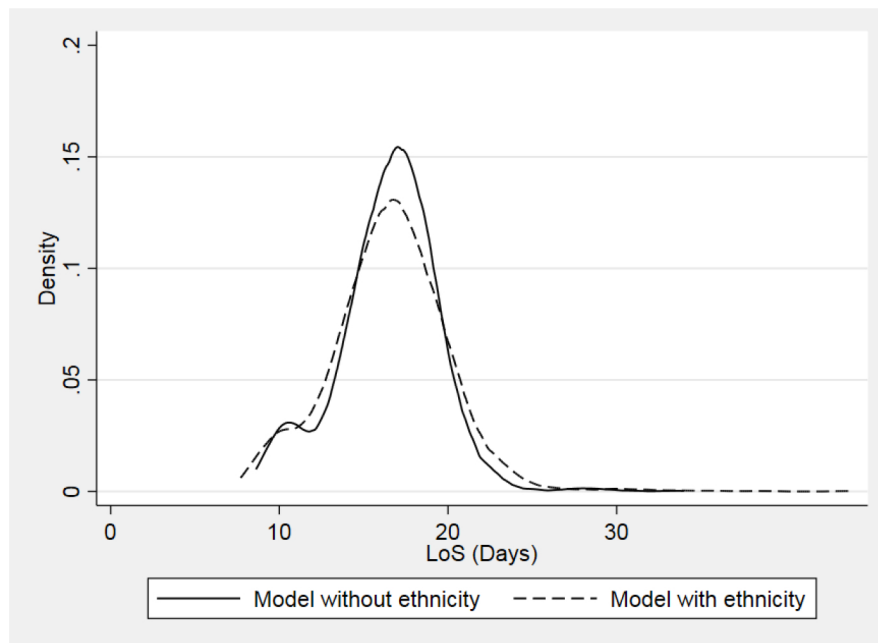


Table C3: Log-normal and Weibull Accelerated Failure Time model with main and interaction effects

variable	Weibull			Log-normal		
	Coef.	Rob. SE	p	Coef.	Rob. SE	p
ecmo	0.24	0.10	0.012	0.40	0.14	0.004
female	−0.03	0.03	0.367	−0.05	0.04	0.189
Period	(reference: after 07/04)					
1	0.23	0.10	0.019	0.27	0.10	0.007
2	0.08	0.06	0.155	0.08	0.07	0.215
age group	(reference: 50-64)					
18-50	−0.13	0.07	0.083	−0.20	0.08	0.012
65-74	−0.11	0.07	0.124	−0.11	0.08	0.142
75+	−0.62	0.09	< 0.001	−0.60	0.10	< 0.001
period*age group						
1*18-50	−0.08	0.12	0.508	−0.06	0.12	0.636
1*65-74	−0.09	0.13	0.467	−0.12	0.12	0.336
1*75+	−0.07	0.17	0.687	−0.04	0.17	0.820
2*18-50	−0.04	0.11	0.709	−0.02	0.11	0.814
2*65-74	0.00	0.08	0.991	0.02	0.09	0.801
2*75+	0.18	0.14	0.189	0.11	0.14	0.435
risk score	−0.06	0.04	0.12	−0.06	0.04	0.147
period*risk score						
1*risk score	−0.12	0.09	0.156	−0.08	0.09	0.374
2*risk score	0.01	0.05	0.807	0.02	0.05	0.647
intercept	2.84	0.05	< 0.001	2.42	0.06	< 0.001
log(p)	0.23	0.02	< 0.001			
log(σ)				−0.07	0.02	< 0.001
AIC	8,028.03			7,949.71		
n	3,049			3,049		

Note: Rob. SE denotes robust SE. Period 1 denotes period before 25 March; period 2 is between 25 March and 7 April. Source: own elaboration using CHES data.

Figure C1: Distribution of mean predicted LoS in ICU based on log-normal model with main effects, by ethnicity



Source: own elaboration using CHES data.

Table C4: Log-normal and Weibull Accelerated Failure Time model with main and age $\times$ sex interaction effects

Variable	Coefficient	Robust SE	p
ecmo	0.40	0.14	0.005
female	-0.02	0.05	0.654
age group	(reference: 50-64)		
18-50	-0.16	0.08	0.036
65-74	-0.10	0.07	0.171
75+	-0.61	0.12	< 0.001
female*age group			
female*18-50	-0.12	0.12	0.334
female*65-74	-0.03	0.08	0.756
female*75+	0.03	0.13	0.808
period	(reference: after 07/04)		
1	0.26	0.10	0.008
2	0.08	0.07	0.212
period*age group			
1*18-50	-0.05	0.12	0.694
1*65-74	-0.12	0.12	0.340
1*75+	-0.03	0.17	0.856
2*18-50	-0.03	0.11	0.801
2*65-74	0.02	0.09	0.806
2*75+	0.12	0.14	0.406
risk score	-0.06	0.04	0.137
period*risk score			
1*risk score	-0.08	0.09	0.362
2*risk score	0.03	0.06	0.607
intercept	2.41	0.06	< 0.001
log(σ)	-0.08	0.02	< 0.001
AIC	7,953.82		
n	3,049		

Note: Period 1 denotes period before 25 March; period 2 is between 25 March and 7 April. Source: own elaboration using CHESS data.

Table C5: Log-normal Accelerated Failure Time model with main effects including five ethnicity categories

Variable	Coeff.	Rob. SE	p	Coeff.	Rob. SE	p
female	-0.05	0.05	0.273	-0.05	0.05	0.264
period	(reference: after 07/04)					
before 25/03	0.27	0.07	< 0.001	0.27	0.07	< 0.001
25/03-07/04	0.13	0.05	0.017	0.13	0.05	0.016
age group	(reference: 50-64)					
<50	-0.23	0.08	0.003	-0.23	0.07	0.002
65-74	-0.16	0.06	0.005	-0.16	0.05	0.004
75+	-0.59	0.06	< 0.001	-0.58	0.06	< 0.001
risk score	-0.07	0.03	0.019	-0.08	0.03	0.011
ecmo	0.54	0.17	0.001	0.53	0.16	0.001
ethnicity	(reference: White)					
Asian				0.10	0.09	0.28
Black				0.12	0.10	0.248
Mixed				-0.01	0.19	0.976
Other				-0.05	0.08	0.509
intercept	2.41	0.06	< 0.001	2.39	0.06	< 0.001
log(σ)	-0.07	0.02	0.001	-0.07	0.02	< 0.001
AIC	6,572.26			6,574.48		
n	2,514			2,514		

Note: Rob. SE denotes robust SE. Source: own elaboration using CHESS data.