Effect of Universal TB Vaccination and Other Policy-Relevant Factors on the Probability of Patient Death from COVID-19

Amos Golan[*ab], Tinatin Mumladze[*] and Danielle Wilson[*]

[*]American University
[a]Santa Fe Institute
[b]Pembroke College, Oxford

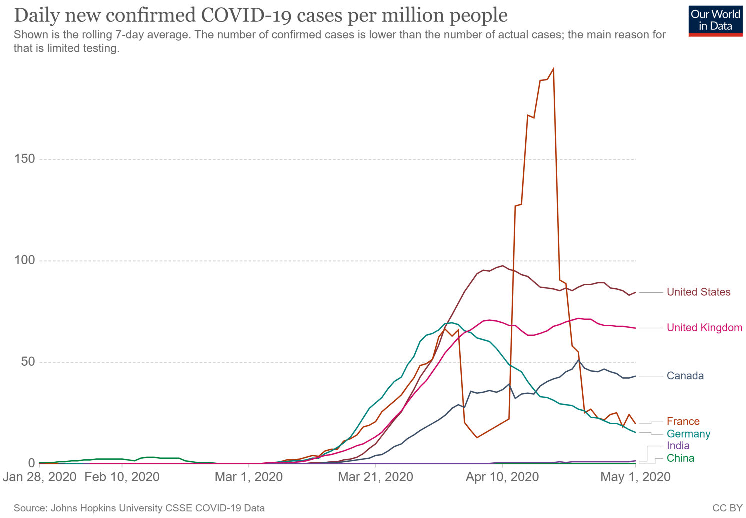
April 8, 2021

This research is based on previous collaborative work with:

Elissa Cohen,
William Mooney,
Troy McGuinness,
&
Jisung Moon

Context and Motivation

- Started the project in April - during the start of the first wave.
- Understanding of the disease was relatively limited at the time.
- Progression of the disease varied around the world.



Question

Have any pre-existing policies influenced susceptibility to COVID-19?
If so, what is the estimated impact?

Objective

Estimate patients' conditional survival/death probability

Data

Individual Level Data

- Age
- Gender
- Country
- Other demographic characteristics
- Pre-existing conditions

Country Level Data

- Prior vaccination policies
- Health indicators
- Economic indicator
- Welfare indicators

Data

Individual Level Data

- Age ✓
- Gender ✓
- Country ✓
- Other demographic characteristics
- Pre-existing conditions ✓

Country Level Data
2015 - 2018

- Prior vaccination policies ✓
- Health indicators ✓
- Economic indicator ✓
- Welfare indicators ✓

Data

Individual Level Data

- Age ✓
- Gender ✓
- Country ✓
- Other demographic characteristics
- Pre-existing conditions ✓

CORRESPONDENCE | VOLUME 20, ISSUE 5, P534, MAY 01, 2020

Open access epidemiological data from the COVID-19 outbreak

Bo Xu • Moritz U G Kraemer • on behalf of the Open COVID-19 Data Curation Group

Published: February 19, 2020 • DOI: https://doi.org/10.1016/S1473-3099(20)30119-5

COVID-19

Data and code repository for the Open COVID-19 Data Working Group: a global and multi-organizational initative that aims to enable rapid sharing of trusted and open public health data to advance the response to infectious diseases.

Acknowledgements

We first want to thank all those individuals and organizations across the world who have been willing and able to report data in as open and timely manner as possible. This work attempts to synthesize information from across a myriad set of data sources. Each entry in our database has an associatd source. A number of individuals have contributed to the specific data added here and their names and details are

Data

Limitations of Individual Level Data

- Initially comprised of primarily symptomatic individuals.
- As sample increased, amount of information for each case decreased.
- Sample size with age, gender, country, and *outcome* information is small.
- Covariates are moderately correlated.
- Ratio of death to infected relatively low compared to other diseases.

CORRESPONDENCE | VOLUME 20, ISSUE 5, P534, MAY 01, 2020

Open access epidemiological data from the COVID-19 outbreak

Bo Xu • Moritz U G Kraemer • on behalf of the Open COVID-19 Data Curation Group

Published: February 19, 2020 • DOI: https://doi.org/10.1016/S1473-3099(20)30119-5

COVID-19

Data and code repository for the Open COVID-19 Data Working Group: a global and multi-organizational initative that aims to enable rapid sharing of trusted and open public health data to advance the response to infectious diseases.

Acknowledgements

We first want to thank all those individuals and organizations across the world who have been willing and able to report data in as open and timely manner as possible. This work attempts to synthesize information from across a myriad set of data sources. Each entry in our database has an associatd source. A number of individuals have contributed to the specific data added here and their names and details are

Data

Limitations of Individual Level Data

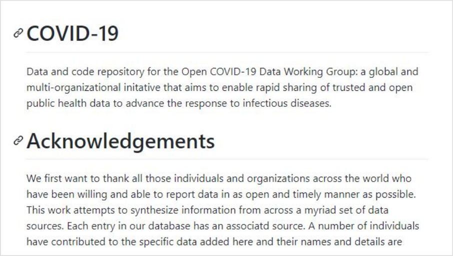
- Initially comprised of primarily symptomatic individuals.
- As sample increased, amount of information for each case decreased.
- Sample size with age, gender, country, and *outcome* information is small.
- Covariates are moderately correlated.
- Ratio of death to infected relatively low compared to other diseases.

CORRESPONDENCE | VOLUME 20, ISSUE 5, P534, MAY 01, 2020

Open access epidemiological data from the COVID-19 outbreak

Bo Xu • Moritz U G Kraemer • on behalf of the Open COVID-19 Data Curation Group

Published: February 19, 2020 • DOI: https://doi.org/10.1016/S1473-3099(20)30119-5



COVID-19

Data and code repository for the Open COVID-19 Data Working Group: a global and multi-organizational initative that aims to enable rapid sharing of trusted and open public health data to advance the response to infectious diseases.

Acknowledgements

We first want to thank all those individuals and organizations across the world who have been willing and able to report data in as open and timely manner as possible. This work attempts to synthesize information from across a myriad set of data sources. Each entry in our database has an associatd source. A number of individuals have contributed to the specific data added here and their names and details are

Modeling Approach

The information theoretic approach has following preferred properties (Golan, Judge and Perloff, 1996):

- Efficient
- Avoids parametric assumptions
- Suitable for small sample sizes
- Suitable for highly correlated covariates
- Permits the inclusion of non-sample information

Model

Model accommodating the noise from each observation is

$$death_{ij} = F\left(\mathbf{x'}_i \beta_j\right) + \epsilon_{ijs} \tag{1}$$

Primary model includes the following covariates:

- gender
- age
- bcg_nev_i: Individual in a country that has never had a BCG vaccination policy
- bcg_past_i: Individual in a country that previously had a BCG vaccination policy
- $immsl_i$: Percent of children in individual's country who received measles immunization.
- $ihepb_i$: Percent of children in individual's country who received hepatitis B immunization.
- $phexp_i$: Private health expenditure per capita at PPP for individual's country.
- $diehh_i$: Mortality rate attributed to household and ambient air pollution for indivudal's country.

Model Details

Results

Marginal Effects of Dependent Variable - Death from COVID-19

	(1)		(2)		(3)	
female	-0.059	**	-0.060	**	-0.052	
	(0.029)		(0.029)		(0.034)	
age	0.008	***	0.008	***	0.008	*
	(0.002)		(0.002)		(0.002)	
bcg_nev	0.474	***	0.446	***	0.372	
Never had BCG policy	(0.050)		(0.067)		(0.246)	
bcg_past	0.160	***	0.177	***	-0.003	
Previously had BCG policy	(0.059)		(0.062)		(0.309)	
immsl	-0.008	**	-0.008	*	-0.026	
Perct. measles immunization	(0.004)		(0.004)		(0.025)	
ihepb	0.014	***	0.013	***	0.012	
Perct. hepatitis B immunization	(0.004)		(0.004)		(0.008)	
phexp	-0.014	**	-0.013	**	0.012	
Private health expd. per capita	(0.006)		(0.006)		(0.028)	
diehh	0.002	***	0.002	***	0.000	
Pollution mortality rate	(0.001)		(0.001)		(0.004)	
centhc	.		-0.055		0.004	
Centralized health care			(0.088)		(0.121)	
posrate	.		.		-0.014	
Pos. test result rate					(0.029)	
_cons	-1.102	***	-1.059	***	0.623	
	(0.285)		(0.288)		(2.203)	
Number of obs =	485		485		358	
Entropy for probs. =	129.6		129.5		91.1	
Normalized entropy =	0.3855		0.3852		0.3673	
Ent. ratio stat. =	413.1		413.4		314	
Pseudo R2 =	0.6145		0.6148		0.6327	

*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.
Private health expenditure is in hundreds of international dollars at purchasing power parity (PPP).

The "centhc" variable is a dummy indicating whether a country has a centralized health care system. The "posrate" variable is share of tests returning a positive COVID-19 result. This variable is a measure of (i) how adequately countries are testing and (ii) how the virus is spreading across countries.

Marginal Effects of Dependent Variable - Death from COVID-19

	(1)		(2)		(3)	
female	-0.059	**	-0.060	**	-0.052	
	(0.029)		(0.029)		(0.034)	
age	0.008	***	0.008	***	0.008	*
	(0.002)		(0.002)		(0.002)	
bcg_nev	0.474	***	0.446	***	0.372	
Never had BCG policy	(0.050)		(0.067)		(0.246)	
bcg_past	0.160	***	0.177	***	-0.003	
Previously had BCG policy	(0.059)		(0.062)		(0.309)	
immsl	-0.008	**	-0.008	*	-0.026	
Perct. measles immunization	(0.004)		(0.004)		(0.025)	
ihepb	0.014	***	0.013	***	0.012	
Perct. hepatitis B immunization	(0.004)		(0.004)		(0.008)	
phexp	-0.014	**	-0.013	**	0.012	
Private health expd. per capita	(0.006)		(0.006)		(0.028)	
diehh	0.002	***	0.002	***	0.000	
Pollution mortality rate	(0.001)		(0.001)		(0.004)	
centhc	.		**-0.055**		0.004	
Centralized health care			(0.088)		(0.121)	
posrate	.		.		-0.014	
Pos. test result rate					(0.029)	
_cons	-1.102	***	-1.059	***	0.623	
	(0.285)		(0.288)		(2.203)	
Number of obs =	485		485		358	
Entropy for probs. =	129.6		129.5		91.1	
Normalized entropy =	0.3855		0.3852		0.3673	
Ent. ratio stat. =	413.1		413.4		314	
Pseudo R2 =	0.6145		0.6148		0.6327	

*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.
Private health expenditure is in hundreds of international dollars at purchasing power parity (PPP).

The "centhc" variable is a dummy indicating whether a country has a centralized health care system. The "posrate" variable is share of tests returning a positive COVID-19 result. This variable is a measure of (i) how adequately countries are testing and (ii) how the virus is spreading across countries.

All Regressions

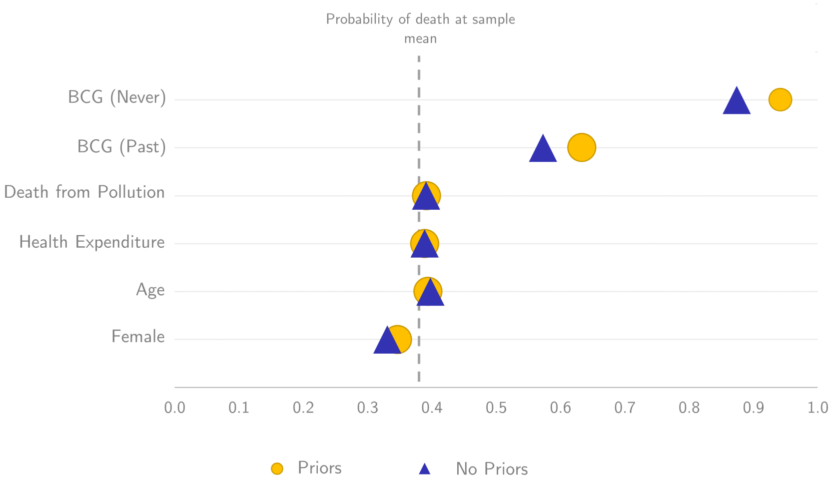
Priors

Sources priors from Karlberg, Hong and Lai (2004):

- Study on fatality rates from severe acute respiratory syndrome (SARS)
- Excluding medical personnel the patient probability of dying:

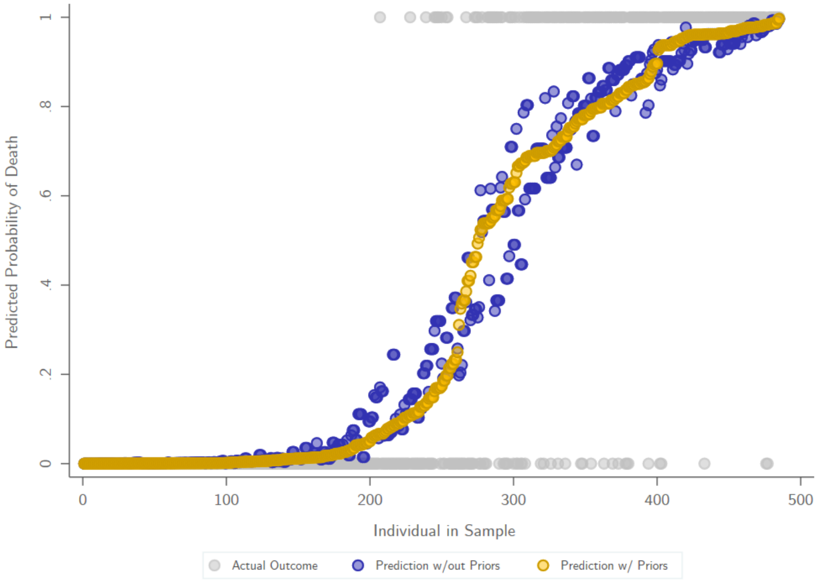
Age	Male	Female
0-44	7.7 %	3.7 %
45-74	32.6 %	24.5 %
75+	64.6 %	63.6 %

Marginal Effects of Dependent Variable - Death from COVID-19
Primary Model with Priors

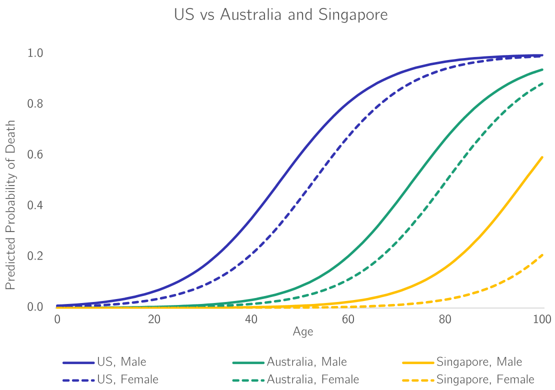


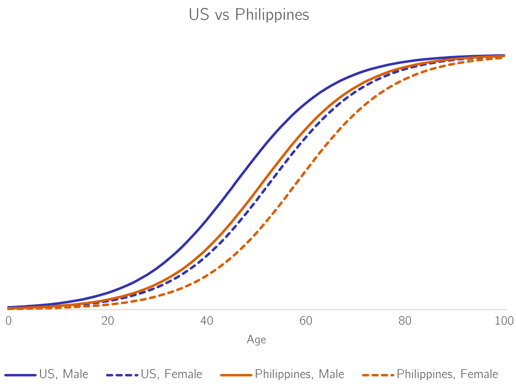
Table

Predicted Probabilities of Death from COVID-19



Case Studies: Predicted Probabilities of Death from COVID-19





	US/AUS	US/SGP	US/PHL
Predicted ratio of death rates	7.34	71.97	1.30
Real ratio of death rates	3.71	72.10	1.62

Summary and Concluding Thoughts

- Information-theoretic inferential approach was used to study the effects of pre-existing policies on susceptibility of Covid-19
- Three potential policies improve patients' conditional survival probabilities:
  - a universal BCG vaccination policy
  - reduction in air pollution
  - an increase in health-related expenditure
- The effects of each one of these policies is magnified with age

Further Research

Further extensions of the research should focus on:

- Re-evaluate the results as more data becomes available
- Control for demographic characteristics
- Study the differential effects of the policies on patients with pre-existing conditions.

Thank you

Questions?

Appendix

Model

When considering noise the relationship between outcome and covariates is:

$$death_{ij} = F\left(\mathbf{x'}_i \boldsymbol{\beta}_j\right) + \epsilon_{ijs} = p_{ij} + \sum_{h=1}^{3} v_{ijh} w_{ijh}$$

The corresponding problem is:

$$max_{p,w} H(p,w) = max_{p,w} \left\{ -\sum_{ij} p_{ij} ln(p_{ij}) - \sum_{ijh} w_{ijh} ln(w_{ijh}) \right\}$$

subject to

$$\sum_i death_{ij} x_{ik} = \sum_i x_{ij} p_{ij} + \sum_{ih} x_{ik} v_h w_{ijh},$$

$$\sum_j p_{ij} = 1, \ \sum_{h=1}^{3} w_{ijh} = 1$$

where h is corresponds to the element of the support space $v = [-1/\sqrt{2}, 0, 1/\sqrt{2}]$.

Return

Marginal Effects of Dependent Variable - Death from COVID-19

	(1)		(2)		(3)	
female	-0.059	**	-0.060	**	-0.052	
	(0.029)		(0.029)		(0.034)	
age	0.008	***	0.008	***	0.008	*
	(0.002)		(0.002)		(0.002)	
bcg_nev	0.474	***	0.446	***	0.372	
Never had BCG policy	(0.050)		(0.067)		(0.246)	
bcg_past	0.160	***	0.177	***	-0.003	
Previously had BCG policy	(0.059)		(0.062)		(0.309)	
immsl	-0.008	**	-0.008	*	-0.026	
Perct. measles immunization	(0.004)		(0.004)		(0.025)	
ihepb	0.014	***	0.013	***	0.012	
Perct. hepatitis B immunization	(0.004)		(0.004)		(0.008)	
phexp	-0.014	**	-0.013	**	0.012	
Private health expd. per capita	(0.006)		(0.006)		(0.028)	
diehh	0.002	***	0.002	***	0.000	
Pollution mortality rate	(0.001)		(0.001)		(0.004)	
centhc	.		-0.055		0.004	
Centralized health care			(0.088)		(0.121)	
posrate	.		.		-0.014	
Pos. test result rate					(0.029)	
_cons	-1.102	***	-1.059	***	0.623	
	(0.285)		(0.288)		(2.203)	
Number of obs =	485		485		358	
Entropy for probs. =	129.6		129.5		91.1	
Normalized entropy =	0.3855		0.3852		0.3673	
Ent. ratio stat. =	413.1		413.4		314	
Pseudo R2 =	0.6145		0.6148		0.6327	

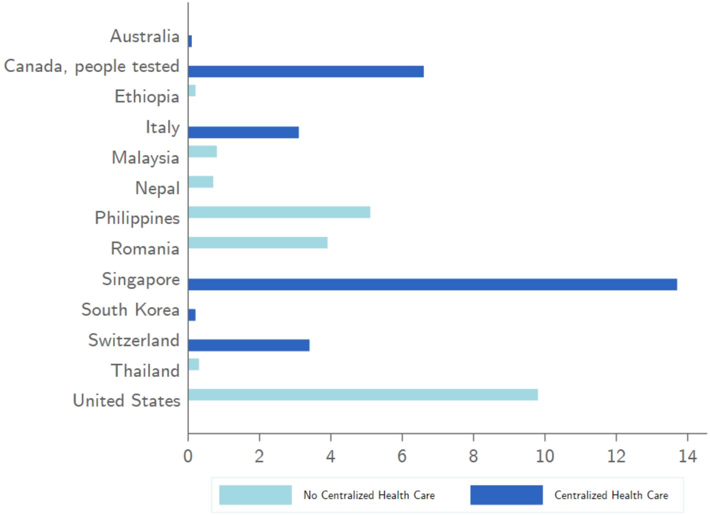
*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.
Private health expenditure is in hundreds of international dollars at purchasing power parity (PPP).

The "centhc" variable is a dummy indicating whether a country has a centralized health care system. The "posrate" variable is share of tests returning a positive COVID-19 result. This variable is a measure of (i) how adequately countries are testing and (ii) how the virus is spreading across countries.

Return

Positive Rates

Positive Rate by Countries in Sample



Return

Marginal Effects of Dependent Variable - Death from COVID-19

Return

	(1)		(1) w/ Priors	
female	-0.059	**	**-0.036**	
	(0.029)		(0.028)	
age	0.008	***	**0.005**	***
	(0.002)		(0.002)	
bcg_nev	0.474	***	**0.453**	***
Never had BCG policy	(0.050)		(0.047)	
bcg_past	0.160	***	**0.172**	***
Never had BCG policy	(0.059)		(0.057)	
immsl	-0.008	**	**-0.009**	
Perct. measles immunization	(0.004)		(0.005)	
ihepb	0.014	***	**0.015**	**
Perct. hepatitis B immunization	(0.004)		(0.006)	
phexp	-0.014	**	**-0.018**	*
Private health expd. per capita	(0.006)		(0.009)	
diehh	0.002	***	**0.003**	***
Pollution mortality rate	(0.001)		(0.001)	
_cons	-1.102	***	**-1.009**	***
	(0.285)		(0.382)	
Number of obs =	485		485	
Entropy for probs. =	129.6		154.1	
Normalized entropy =	0.3855		0.4583	
Ent. ratio stat. =	413.1		364.2	
Pseudo R2 =	0.6145		0.5417	

*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.
Private health expenditure is in hundreds of international dollars at purchasing power parity (PPP).

Model Predictions with Observed Events

- The total number of Covid-positive individuals is 131,197,518 (Apr 3, 2021)
- The share of elderly (65+) male and female individuals is assumed to be roughly 10 percent in the total number of observed Covid-positive cases

Predictions of 10 percent improvement in the policy variable relative to the mean

		Health Expenditure		Pollution Mortality Rate		Both	
		Male	Female	Male	Female	Male	Female
Reduction in:	Elderly deaths	242,534	198,199	359,171	290,179	592,525	268,068
	Total deaths	15.40%		22.70%		37.10%	