medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

The first three months of the COVID-19 epidemic: Epidemiological evidence for two separate strains of SARS-CoV-2 viruses spreading and implications for prevention strategies

Two epidemics of COVID-19

KNUT M. WITTKOWSKI[1*]

[1] ASDERA LLC, New York, NY

* Corresponding author:
E-mail: knut@asdera.com (KMW)

Abstract

About one month after the COVID-19 epidemic peaked in Mainland China and SARS-CoV-2 migrated to Europe and then the U.S., the epidemiological data begin to provide important insights into the risks associated with the disease and the effectiveness of intervention strategies such as travel restrictions and social distancing. Respiratory diseases, including the 2003 SARS epidemic, remain only about two months in any given population, although peak incidence and lethality can vary. The epidemiological data suggest that at least two strains of the 2020 SARS-CoV-2 virus have evolved during its migration from Mainland China to Europe. South Korea, Iran, Italy, and Italy's neighbors were hit by the more dangerous "SKII" variant. While the epidemic in continental Asia is about to end, and in Europe about to level off, the more recent epidemic in the younger US population is still increasing, albeit not exponentially anymore. The peak level will likely depend on which of the strains has entered the U.S. first. The same models that help us to understand the epidemic also help us to choose prevention strategies. Containment of high-risk people, like the elderly, and reducing disease severity, either by vaccination or by early treatment of complications, is the best strategy against a respiratory virus disease. Social distancing or "lockdowns" can be effective during the month following the peak incidence in infections, when the exponential increase of cases ends. Earlier containment of low-risk people merely prolongs the time the virus needs to circulate until the incidence is high enough to initiate "herd immunity". Later containment is not helpful, unless to prevent a rebound if containment started too early.

About the Author

Dr. Wittkowski received his PhD in computer science from the University of Stuttgart and his ScD (Habilitation) in Medical Biometry from the Eberhard-Karls-University Tübingen, both Germany. He worked for 15 years with Klaus Dietz, a leading epidemiologist who coined the term "reproduction number", on the Epidemiology of HIV before heading for 20 years the Department of Biostatistics, Epidemiology, and Research Design at The Rockefeller University, New York. Dr. Wittkowski is currently the CEO of ASDERA LLC, a company discovering novel treatments for complex diseases from data of genome-wide association studies.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Table of Content

Introduction **3**

Materials and Methods **4**

Data 4

Methods and models 4

Statistical and Bioinformatics Methods 4

Epidemiological Models 4

Results **5**

Incidence by Country. Norther Hemisphere 5

Time-course by country/region 7

A Window of Opportunity for Containment (Social Distancing) 14

Discussion **16**

Strengths and shortcomings 16

Evidence for (at least) two different strains of SARS-CoV-2 17

Changes in infectivity and lethality between China and Europe 18

Predictions for COVID-19 in North America 19

A historical perspective 19

Implications for prevention 20

Conclusions **22**

References **22**

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Introduction

The first cases of a new coronavirus strain, termed SARS-CoV-2 (Severe Acute Respiratory Syndrome CoronaVirus) by the International Committee on Taxonomy of Viruses (ICTV) (Cascella 2020), were reported on 31-12-2019 in Wuhan, the capital of the Hubei province of China.(Jernigan 2020) As of 2020-03-31, 10:00 CET, 777,798 symptomatic cases and 37,272 deaths have been reported from virtually every country in the northern hemisphere (see Section Data), The disease was termed COVID-19 by the WHO on 2020-02-11, and categorized as a pandemic on 2020-03-12, yet the details of the spread and their implications for prevention have not been discussed in sufficient detail.

Between 02-14 and 03-16, the Dow Jones fell 31% from 29,440 to 20,188, raising fears for the economy, in general, and retirement savings, in particular. Several administrations have imposed severe restrictions aimed at containment of the virus. For instance,

- On 03-08, the Italian government imposed a quarantine on 16 million people in the north of Italy, which was followed up on 03-11 with a nationwide closure of all restaurants and bars along with most stores.(WSJ, 2020-03-11)
- on 03-11, the U.S. administration banned travel from 26 European countries (Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, and Switzerland). https://www.cnn.com/world/live-news/coronavirus-outbreak-03-12-20-intl-hnk/index.html.
- On 03-15, New York's Mayor de Blasio reversed his previous position that NY schools should remain open to avoid health care workers from "staying home and watching their children" and announced NY public schools to be closed, following many other school systems.
- From 03-17, all New York, New Jersey, and Connecticut restaurants were had to close. From 03-22, hairdressers and barbers were also closed.
- From 03-19 California was "shut down".(Executive order N-33-20)
- On 03-22, the lockdown in the Italian region of Lombardy was tightened to ban sports and other physical activity, as well as the use of vending machines.
- Also, on 03-22, the National Guard was activated in New York, California, and Washington State, five senators self-quarantined. NY governor Cuomo mandated that all nonessential businesses close or work from home.

By the end on 03-20, the Dow Jones was down at 19,173 (35%) from 02-14. On 03-26, the U.S. Senate approved a $2T stimulus package.

For most of the first three months of the epidemic, much of the response was driven by "fear, stigma, or discrimination"(Ren 2020), including naming SARS-CoV-2 the "China virus"(Rogers 2020), despite the fact that seasonal respiratory zoonotic pathogens typically originate in China, where life-animal markets provide chances for animal viruses to transmit to humans.(Malik 2020)

After three months, enough data are available to discuss important epidemiological characteristics of COVID-19 and the potential impact of interventions. In particular, we have now seen the number of new cases (and deaths) to decline in China and South Korea and to at least stabilize in some European countries. Changes in number of deaths follow the changes in number cases (albeit at a lower level) by about two weeks. Hence, we can discuss both the infectiousness and the lethality of the virus, two important characteristics to assess public health impact of the disease.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

One of the key findings published herein is evidence that at least two strains of SARS-CoV-2 with different infectiousness and lethality have evolved, and by following the likely path for each of these strains we can obtain novel insights into the nature of the epidemic and, thus, the effectiveness of prevention strategies. Another finding is that the timepoint when a public health intervention starts in the course of the epidemic (especially the "turning point" where the increase in new cases begins to decline) is crucial for the impact of the intervention.

Materials and Methods

Data

All data were downloaded on 2020-03-31 from the European Centre for Disease Prevention and Control (ECDC) Web site at https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide, where data are collected daily between 6:00 and 10:00 CET. Updates were collected from the Johns Hopkins online tracker available at https://systems.jhu.edu/research/public-health/ncov/. New York City data was downloaded from https://www1.nyc.gov/assets/doh/downloads/pdf/imm/covid-19-daily-data-summary.pdf .

Population data were accessed from https://www.worldometers.info/world-population/population-by-country/ on 2020-03-12. Data on ages by country were accessed from https://data.worldbank.org/indicator/SP.POP.65UP.TO.ZS.

Methods and models

Statistical and Bioinformatics Methods

The data were processed with MS Excel. To avoid biases from inappropriate model assumptions only basic descriptive statistics were employed. In some cases, data from only the day before or after (or both) was averaged (up to a three day moving average) to reduce the effects of apparent reporting artifacts (Darwin's *natura non facit saltum*,[(Berry 1985)]) without creating undue biases. The two "smoothers" applied were:

- averaging x_0 with a previous x_{-1} (or, rarely, following x_{+1}) data and
- applying a moving average of $(x_{-1}, x_0, x_{+1}) \rightarrow (2\,x_{-1} + x_0,\ x_{-1} + x_0 + x_{+1},\ x_0 + 2\,x_{+1})/3$

No other changes were applied to the data.

Like China in mid-February, the German Robert-Koch-Institute (RKI) changed the reporting system in mid-March, which resulted in a near 6-fold increase of the data reported on 03-20 over the data reported on 03-19. Such changes in reporting systems add to the difficulties in interpreting the data.

Epidemiological Models

If a disease causes immunity after an infectious period of a few days only, like respiratory diseases, an epidemic extinguishes itself as the proportion of immune people increases. Under the SIR model,[(Kermack 1991)] for a reproduction number[(Dietz 1993)] (secondary infections by direct contact in a susceptible population) of R_0=1.5–2.5 over 7 days (recovery rate: γ=1/7=.14), the noticeable part of the epidemic lasts about 90–45 days ($R_0/\gamma=\beta$=.21–.36) in a homogeneous population of 10M. The period is shorter for smaller more homogenous and longer for larger, more heterogenous populations. For a given infectious period $1/\gamma$ (here, e.g., 7 days. SARS and COVID-19 incubation period plus 2 days[(Lauer 2020)]), R_0 also determines how long it will take for early cases to become visible after a single import (150–60 days), the peak prevalence of infections (5–22%), and how many people will become immune (55–90%). To allow for comparisons between models,

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

an arbitrary proportion of symptomatic cases among those becoming infected (.05%) is used and 2% of cases are assumed to die.

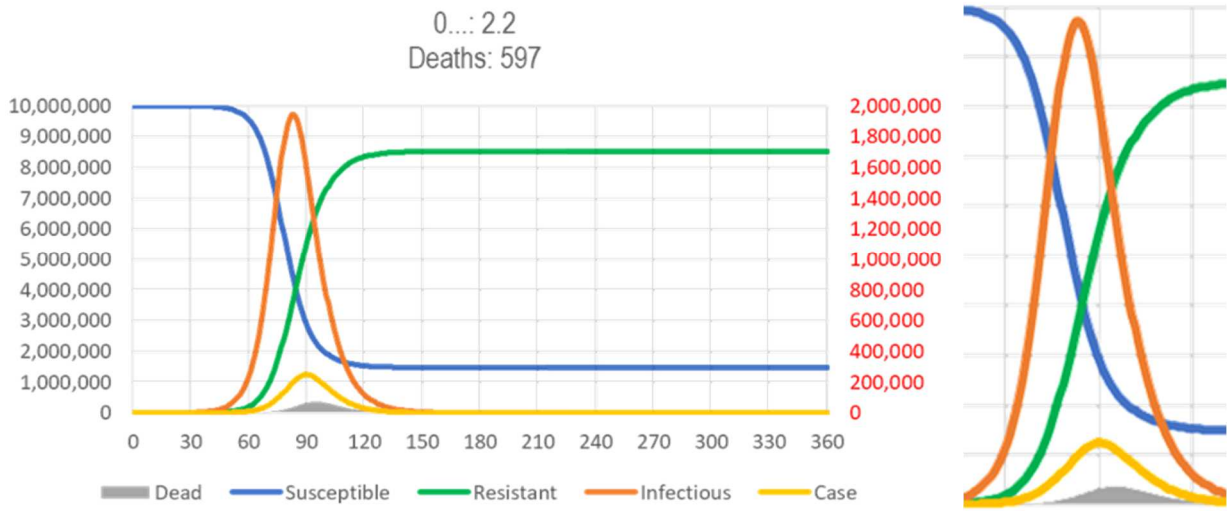


Fig 1: SIR Model of SARS. Number of susceptible (blue), infectious (red), resistant (green), case (orange) and dead (gray) people after a population of 10,000,000 susceptible people is exposed to 20 subjects infected carrying a novel virus. Assumptions: R_0 = 2.2, infectious period = 7 days,(available from https://app.box.com/s/pa446z1csxcvfksgi13oohjm3bjg86ql)

The model reflects that the key milestones of the epidemic, the turning point/half maximal point in infections (red, day 68), the turning point/half maximal point in the number of cases (day 75, orange), the peak in infections (day 83, red), the peak in cases (day 70, orange) and the peak in deaths (day 77, gray) follow each other, following the previous milestone after about a week.

Results

Incidence by Country. Norther Hemisphere

Table 1 shows the raw daily incidence by population sizes for countries with epidemiological relevance in the northern hemisphere. Countries within proximity are grouped by their peak incidence (red background).

The Hubei province in China (with the capital Wuhan), South Korea, Iran, Italy (especially the Lombardy region), and Spain have the highest peak incidence, followed by the countries neighboring Italy.

It should be noted, however, that there is no uniform definition of “cases”. In some countries a case needs to have symptoms, in other countries, it suffices to have antibodies (be immune).

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Table 1: Incidence by Country. Dates: Feb 13 (peak intensity in Mainland China, mostly Hubei and neighboring provinces), Feb 19 to Mar 22. Countries with low population size (PSz) or low number of cases (Total) are hidden. Red background indicates countries/dates with high incidence. Countries/regions are sorted by proximity among each other and distance from Mainland China.

Country	PSz[M]	Total	Total/M	02-05	02-21	02-22	02-23	02-24	02-25	02-26	02-27	02-28	02-29	03-01	03-02	03-03	03-04	03-05	03-06	03-07	03-08	03-09	03-10	03-11	03-12	03-13	03-14	03-15	03-16	03-17	03-18	03-19	03-20	03-21	03-22	03-23	03-24	03-25	03-26	03-27	03-28	03-29	03-30	03-31
CN(Hubei)	58.5	82,241	1405.83	66	15	14	11	3.7	8.8	7	7.5	5.6	7.3	9.8	3.5	2.2	2	2	2.9	1.7	0.8	0.8	0.3	0.5	0.4	0.4	0.3	0.4	0.4	0.7	0.4	0.8	1.7	0.9	1.1	2.4	1.2	1.3	1.7	1.6	2	1.9	1.7	1.4
Vietnam	97	229	2.36	0																0	0	0.1	0	0	0	0.1	0.1	0	0	0	0.1	0.1	0.1	0	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.3	0		
Cambodia	17	109	6.41																		0.1			0.1	0.1	0.1	0.1	0.3	0.7			0.7	0.7	1.1	1.1	0.1	0.2	0.3	0.2	0.2	0.1	0.1	0.1	
South Korea	51	9,786	191.88	0	1.5	3.7	5	3.2	2.5	5	8.8	12	13	14	13	12	10	9.4	9.4	8.7	7.2	5.6	3.7	3.7	2.2	2.2	2.1	1.5	1.5	1.6	1.8	2.6	2.5	2.5	1.9	1.3	1.5	2	2	2.3	2.3	2.1	2	2
Singapore	6	865	144.08	1	0.2	0.2	0.5		0.2	0.2	0.3	0.5	0.3	0.7	0.7	0.3	0.3	0.3	1.3	1.4	1.6	2	1.7	1	2	1.5	2.2	2.3	2	2.8	3.8	6.6	6.6	6.7	6.5	6.9	7.3	8.2	6	10	13	12	6.8	3.4
Malaysia	32	2,626	82.06	0.1								0	0.1		0.1	0.1	0.2	0.4	0.5	0.5	0.3	0.3	0.4	0.4	0.5	0.5	1.2	2.8	3.7	4.6	3.8	3.6	3.7	3.9	4.5	5.1	5.7	4	5.3	6.7	4.4	4.6	4.8	4.9
Japan	126	1,953	15.50	0	0.1	0.1	0.2	0.1		0.2	0.2	0.2	0.2	0.1	0.1		0.1	0.4	0.3	0.4	0.4	0.3	0.2	0.4	0.4	0.4	0.5	0.3	0.3	0.1	0.2	0.2	0.6	0.5	0.4	0.3	0.4	0.4	0.6	0.8	1.1	1.5	1.2	0.9
Taiwan	24	316	13.16	0		0.1		0.1	0.1	0	0	0.1	0.2	0		0	0		0.1	0				0.1		0	0.2	0.3		0.3	0.4	0.9	0.8	0.8	0.7	0.8	1	0.9	0.8	0.7	0.7	0.8	0.8	0.4
Thailand	70	1,651	23.59	0.1					0	0			0		0			0	0	0	0			0.1	0.2		0.2	0.2	0.2	0.5	0.5	0.3	0.3	2.6	2.4	2.3	1.5	1.5	1.6	1.3	0.8	0.8	2.9	2.9
Indonesia	274	1,414	5.16												0				0	0	0	0	0	0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.3	0.4	0.4	0.3	0.3	0.4	0.4	0.5	0.4	0.4	0.5	
Philippines	110	2,084	18.95																	0	0	0	0.2		0.1	0	0.1	0.3	0.3	0.2	0.2	0.2	0.2	0.7	0.7	0.2	0.5	0.8	0.8	0.7	1.3	1.9	4	5.1
India	1380	1,251	0.91													0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
Pakistan	221	1,625	7.35								0			0			0			0		0	0		0	0		0	0.4	0.4	0.3	0.3	0.6	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	
Afghanistan	39	141	3.62						0											0	0		0	0			0.1	0.2	0.1	0				0.1	0.1	0.1	0.2	0.4	0.4	0.2	0.2	0.3	0.4	0.5	
Sri Lanka	21	120	5.71																						0	0	0.1	0.2	0.4	0.5	0.4	0.4	0.3	0.6	0.6	0.4	0.5	0.2	0.1	0.1	0.2	0.2	0.1	0.1	
Iran	84	41.495	493.99		0	0.2	0.1	0.2	0.2	0.4	0.5	1.3	1.7	2.4	4.6	6.2	8.5	8.5	11	11	11	11	8.8	8.8	11	13	15	16	14	13	14	14	13	13	13	12	17	21	26	28	35	36	36	37
Iraq	40	630	15.75						0	0.1		0	0	0.2	0.2	0.1	0.1	0.1	0.2		0.4	0.2			0.2	0.1	0.3	0.5	0.5	0.4	0.4	0.3	0.3	0.5	0.5	0.6	0.9	1.1	0.8	1.2	1.6	1.1	1.4	1.7
Saudi Arabia	35	1,453	41.51												0	0		0.1	0.1	0	0	0.1	0.1	0.1	0.7	0.5	0.7	0.5	0.5	0.4	0.4	1	1.6	1.8	2.6	3.4	3.7	3.7	3.6	3.2	2.8	2.8	3.6	3.6
United Arab	10	611	61.10			0.2	0.2					0.6		0.2		0.3	0.3	0.1	0.1	0.8	0.8	0.7	0.7	1.5	0.6	0.6			0.7	0.7	1	1.4	1.8	0.7	0.7	1.5	3.2	4.8	5.7	5.2	4.8	6.3	7.2	7.2
Oman	5	179	35.80						0.4	0.4		0.4				0.2		1.6	0.2			0.2	0.2			0.2	0.2		0.4	0.4	1	1	1	1.3	1.3	0.6	2.2	3.6	3	3.2	3.2	4.2	3	2.4
Kuwait	4	266	66.50					0.8	0.5	1.5	3.8	4.3	0.5		0.3	1.3	1.3	0.3	0.3	0.4	0.4	0.8	0.3	1	1.2	2.6	4	1	2	2.8	2.2	2.1	2	3.5	3.5	3	0.3	0.5	1.8	2.8	3.9	3.8	3.8	2.8
Qatar	3	693	231.00											0.3	0.7	0.8	0.8		0.5	0.5	0.3	1	1	28	27	26	19	5.7	18	12	4.9	3.3	2.7	3.3	3.7	4.3	4.3	4.8	5.2	4.1	5.9	7.7	15	20
Bahrain	2	515	257.50					0.5	0.5	11	5		2.5	1.5	1.5	1.5	1	1.5		2	3.5	8	7.8	7.8	13	13	24	0.5	1.5	5	7	9	7	8.3	9.7	14	10	14	18	12	12	8.5	8.5	7.5
Jordan	10	268	26.80													0.1												0.7	0.7	1	1	1.1	1.1	1.4	1.4	1.5	1.3	1.5	2.4	2.8	3.3	2.3	1.9	1.4
Israel	9	4,473	497.00			0.1		0.1				0.1	0.4		0.3		0.3	0.3	0.2	0.2	0.7	1.6	1.7	1.7	1.3	1.6	2.2	6.9	8	6.9	6.8	6.6	16	16	19	21	41	52	45	38	53	53	47	47
Lebanon	7	446	63.71			0.1					0.1	0.1		0.1	0.9	0.4			0.4	0.9		1.4	1.3		2.9	0.7	1.9	1.9	0.9	1.5	1.5	1.9	2.2	4.6	7	2.6	2.7	4.9	4.8	4.7	3.3	3	2.4	2.4
Palestine	5	115	23.00																1.4	1.8	0.6		0.2	1.8	0.2	0.5	0.5	0.6		0.2	0.4	0.6	0.6	0.5	0.5	0.9	0.5	0.1	2.4	2.4	1.4	1.5	1.5	1.8	
Russia	146	1,836	12.58													0	0			0	0			0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.3	0.4	0.4	0.5	0.5	0.4	1.1	1.2	1.3	1.6	1.8	2.1	
Azerbaijan	2	273	136.50										0.5		1			0.8	0.8	1.5			0.5	0.5	1		1.5	1.5		2.3	2.3	4	4	2.3	2.3	5.2	5.7	6.2	8.8	8.8	15	15	23	23
Georgia	4	98	24.50								0.3		0.3	0.3					0.8	0.8	0.8	0.3	0.5	2	0.3	0.3	1.3		0.3	0.3	0.4	0.8	0.8	0.8	1.4	1.5	1.6	1.8	1.5	1.3	0.8	0.8	1.6	1.6
Belarus	9	105	11.67									0.1				0.2	0.2	0.2					0.2	0.2	0.3	0.5	0.5	0.3	0.3	0.7	0.7	0.7	1.2	1.1	1.1	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.6	0.6
Italy	60	101,739	1695.65			0.2	1	0.9	1.6	1.6	1.3	4.2	4	4	9.4	2.4	11	9.8	13	13	21	25	30	24	33	42	42	53	53	67	63	73	83	100	104	94	84	87	92	96	100	100	87	68
Switzerland	9	15,412	1712.44							0.1		0.8	0.4	0.7	0.7	0.7	0.8	2.2	3.3	14	6.1	7.6	4.7	13	17	24	30	26	47	47	50	59	81	103	118	118	101	101	103	126	127	129	126	126
Austria	9	9,618	1068.67							0.2		0.3	0.2	0.3	0.4	0.4	0.7	0.6	1.3	3.7	2.8	0.3	3.2	5.7	7.1	13	16	17	21	25	29	44	49	54	50	68	86	88	97	97	74	66	58	89
Slovenia	2	763	381.50															0.5	2.5	1.5	1.5	2	3.8	3.8	13	20	23	20	19	17	11	11	11	16	16	16	16	19	22	25	28	26	23	17
France	65	44,550	685.38							0	0	0.3	0.3	0.7	0.5	0.7	0.5	1.1	2.1	2.9	3.2	4.1	5	5.7	7.6	9.2	12	13	14	19	18	22	26	25	28	36	40	45	50	55	59	61	59	58
Spain	47	85,195	1812.66						0	0.1	0.1	0.3	0.2	0.7	0.4	0.7	0.8	1	1.3	2.4	1.2	8.2	8.2	9.3	15	15	26	32	37	37	42	54	73	75	81	87	96	140	169	183	170	160	151	136
Portugal	10	6,439	643.87												0.1	0.1	0.2	0.1	0.4	0.4	0.8	0.9	0.9	1	1	1.9	3.4	5.7	7.6	8.6	12	18	19	24	28	35	41	41	49	58	72	87	71	56
Germany	84	61,913	737.06							0	0	0.3	0.1	0.6	0.2	0.3	0.5	0.8	1.6	2.9	2	1.1	2.3	2.3	6.4	6.4	8.3	8.7	12	14	13	13	71	48	39	39	45	47	56	64	64	56	57	55
United Kingd	68	22,141	325.60					0.1				0	0	0.1	0.2	0.1	0.2	0.5	0.4	0.7	0.6	1	0.7	0.8	1.2	2	1.7	6.4	3.7	2.2	8	8	9.9	9.9	13	13	18	18	21	42	37	37	37	
Poland	38	2,055	54.08														0		0.1	0.1	0	0.1	0.2	0.1	0.2	0.5	0.7	0.7	1	1	1.4	1.4	1.8	2.2	2.6	2.6	3	4	4.1	4.3	5.1	5.6	6.1	5.1
Romania	19	1,952	102.74								0.1		0.1				0.1		0.1	0.1	0.3	0.1	0.1	0.4	1.1	1	1.3	1.3	1.4	2.2	2.1	2.1	1.1	1.9	2.6	4.8	6.9	9	7.2	9.3	11	12	10
Bulgaria	7	359	51.29																		0.3	0.3		0.2	0.2	1.7	1.7	1.4	1.4	1.6	2.1	2.1	2.3	3.4	4.5	3.1	2.4	2.7	3	3.5	4.2	5	2.1	1.9
Netherlands	17	11,750	691.18									0.1	0.1	0.3	0.4	0.3	0.6	0.6	2.6	3	3.6	4.2	3.3	3.6	7.1	8.1	8.9	9.8	10	16	17	20	24	33	34	35	37	43	49	60	69	68	65	52
Belgium	12	11,899	991.58												0.1	0.5	0.4	0.8	2.3	4.9	5	2.6	3.3	2.3	3.9	7.1	13	14	14	15	15	20	26	39	47	41	35	48	69	91	121	121	115	115
Czech Repub	11	3,002	272.91												0.3	0.2		0.3	0.4	0.6	0.6	0.5	0.7	2.1	2.8	2	3.1	5.8	6.5	6.7	6.8	12	12	14	14	11	11	14	24	31	31	30	15	15
Slovakia	2	346	173.13																	0.5	1	0.5	0.5	1	1.5	5.5	4.5	7	8.5	12	6.5	6.5	6.5	12	10	9.2	4.8	4.8	6	20	20	10	10	5.1
Greece	10	1,212	121.20								0.1	0.2	0.1	0.3				0.3	2.2	1.3	2.1	0.7	1.1	0.6	0.9	3.4	5.7	6	5.4	4.8	3.4	3.7	4.1	3.3	3.3	9.4	7.1	6.3	6.3	7.1	7.4	9.5	8.2	6.9
Hungary	10	447	44.70														0.1	0.1	0.1	0.1	0.3	0.1	0.1	0.2	0.2	0.6	0.6	0.7	0.7	0.6	0.6	1.8	1.8	2.3	2.3	3.1	3.2	3.3	3.7	3.7	4.3	5.6	3.5	1.3
Denmark	6	2,577	429.50								0.2		0.2	0.2	0.2	0.2	0.5	0.3	1.7	0.5	1.3	1.2	13	25	42	27	21	3.8	8.5	11	13	15	12	12	12	11	15	18	23	25	27	29	29	30
Finland	6	1,313	218.83								0.2		0.2		0.3	0.3	0.2	0.4	0.4	1.2	0.9	0.9	1.7	1.6	1.6	8	8	9.2	6.6	6.1	5.5	6.8	6.8	8.3	12	16	15	14	15	13	11	16	16	16
Sweden	10	4,009	400.93								0.1	0.5	0.5	0.1	0.1	0.1	0.9	1.1	2.6	5	5	4.2	4.5	7.8	14	16	16	15	11	8.9	4.6	13	13	16	16	14	18	21	24	28	32	32	29	29
Norway	5	4,226	845.20								0.2	0.6	0.4	1.8	0.8	1.2	1.6	4.6	6	5.4	6.8	4.4	4.6	25	29	32	29	29	29	27	25	24	29	34	38	42	46	55	55	67	67	52	43	34
Iceland	2	1,086	543.00										0.5		1	1.5	5	5	4.5	4	5	1	5	5	5	8	8	11	15	15	17	22	27	36	36	29	29	37	37	38	38	34	33	32
Ireland	5	2,910	582.00											0.2			0.2	0.8	1.4	1	0.2	0.4	1.4	1.4	1.6	5.4	4.2	7.6	8	11	14	15	34	28	22	24	43	44	45	51	60	59	50	50
Croatia	4	790	197.50							0.3	0.3	0.3	0.5		0.5	0.3	0.3		0.3	0.3	0.3			0.3	0.8	2.3	1.5	1.5	2.8	2	3.3	3	5.8	13	13	11	15	19	12	17	22	18	17	17
Serbia	9	782	86.89																	0.1			0.2	0.2	1.1	1.1	1.2	1.2	0.6	0.6	1.7	2.4	2.3	2.3	1.6	4.3	3.8	4	6	8	15	15	9.1	4.6
North Maced	2	285	142.50								0.5									0.5	0.5		1	1	0.5	0.5	1	1		3	6	5.5	7	7	11	11	9.3	11	12	11	11	11	11	11
Lithuania	3	516	171.83									0.3											0.3	0.3		0.5	0.5	1.3	1.3	1	2.7	2.7	5	7	12	13	12	16	16	14	14	21	21	11
Latvia	2	376	188.00												0.3	0.3				0.3	0.3	0.5	1.3	1.3	2	2	1.5	3	3	2.5	8.8	8.8	10	10	7	7	17	14	11	12	18	20	16	12
Egypt	102	609	5.97												0			0	0	0.1	0.2	0.2	0.1	0	0.1	0.1	0.1	0.1	0.1	0.2	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.4	0.5	0.4	0.3	0.4	0.4	0.3
Algeria	44	511	11.61							0					0		0	0.2	0.1		0	0				0.1	0	0.3	0.3	0.2	0.2	0.2	0.2	0.2	0.1	1.1	1.1	0.9	0.9	0.9	1.2	1.2	1.2	1.2
Morocco	37	556	15.03												0	0		0	0					0	0	0	0	0.3	0.3	0.2	0.2	0.3	0.4	0.4	0.4	0.4	0.5	1	1.4	1.6	1.7	1.8	1.8	2.1
Tunisia	12	312	26.00												0	0						0	0	0.3	0.2	0.5	0.3		0.2	0.2	0.3	0.4	0.8	1	1	1	1.2	2.1	2.5	2.5	3	2.9	2.8	2.8
United States	331	164,620	497.34		0	0.1			0.1		0	0	0	0	0.1		0.1	0.1	0.2	0.3	0.3	0.4	0.6	0.8	0.9	1.1	1.5	2.3	2.5	2.7	5.3	9	15	16	22	26	31	34	37	51	56	59	60	62
Mexico	129	1,094	8.48										0	0	0						0	0		0	0	0	0.1	0.1	0.1	0.2	0.1	0.2	0.4	0.3	0.4	0.5	0.4	0.3	0.5	0.9	1	1.1	1	0.9
Canada	38	7,424	195.37	0	0				0.1		0	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.1	0.4	0.4	0.5	0.7	1	1.7	2.2	2.6	3.6	3.7	3.8	5.1	5.1	5.2	5.7	8.2	38	17	18	18	23	31
Panama	4	1,075	268.75																				0.3	1.5	1.8	3.3	2.3	1.8	3.3	3.3	4.8	5.7	6.6	14	14	14	17	19	29	29	28	26	24	22
Dominican R	11	901	81.91												0.1					0.1			0.3				0.5				0.5	0.5	4.1	4.1	8.2	3.9	6.1	7.3	8.7	11	11	8.3	8.3
Costa Rica	5	330	66.00																	0.2		0.8	0.8	0.8	1.8	0.2	0.6	0.9	0.9	1.2	1.8	3.8	4.1	3.2	2.3	3.4	4.5	4.5	4.5	6	6.4	6.4	3.8	3.2
Ecuador	18	1,966	109.22											0.1	0.3	0.1	0.1	0.1	0.2			0.1	0.1	0.1		0.2	0.2	0.3	0.5	1.2	2.9	3.2	5.4	6.7	8.1	13	10	7.3	8.3	10	12	8.7	6.3	3.8
Nigeria	206	5,438	26.40									0										0	0							0	0	0	0	0	0	0	0	0	0.1	0.1	0.1	0.1	0.1	0.1
South Africa	59	1,348	22.85																0		0	0	0.1		0.1	0.1	0.1	0.2	0.2	0.2	0.4	0.5	0.7	0.7	0.7	1.1	1.8	2.5	2.9	3.7	2.2	2.2	1.6	0.8
Cameroon	27	142	5.26																0	0	0	0						0	0	0.1	0.1	0.1	0.1	0.2	0.2	0.5	0.5	0.6	0.3	0.3	0.2	0.2	0.8	0.8
Senegal	17	162	9.53												0	0	0.1	0.1	0.1							0.1	0.8	0.1	0.3	0.1	0.2	0.3	0.3	0.3	0.5	0.6	0.7	0.6	0.6	0.6	0.6	0.7	0.8	1
Togo	8	34	4.25																0.1	0.1													0.5	0.5	0.4	0.4	0.4	0.3	0.3	0.1	0.3	0.3	0.4	0.4
Australia	25	4,607	184.28	0	0.1	0.2	0				0		0.1	0	0.1	0.2	0.3	0.4	0.3	0.2	0.4	0.4	0.5	0.6	0.6	1.2	1.6	2	2.4	2.7	3.2	4.4	5.8	11	11	19	14	14	15	15	13	13	11	19
New Zealand	5	597	119.40								0.2					0.1	0.1	0.2	0.2	0.2							0.2	0.2	0.2	1.2	1.2	1.6	2.2	2.7	4.2	-4	8	11	13	15	14	14	14	19
Brazil	212	4,579	21.60							0				0				0	0	0		0.1		0	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.6	0.9	1.2	1.5	1.7	1.6	1.5	1.5	1.9	2.3	2.3	1.7	1.5
Colombia	51	798	15.65																	0		0	0	0.1	0.1	0.1	0.1	0.4	0.2	0.2	0.4	0.4	0.5	0.7	0.9	0.9	0.9	1.4	1.8	0.6	0.9	1.2	1.8	1.9
Argentina	45	966	21.47													0	0	0	0	0.1	0	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.4	0.7	0.9	1	1.1	1.2	1.7	2.3	1.9	1.9	1.7	1.5	3.2
Peru	33	950	28.79																	0	0.1	0.1	0.1	0.1	0.2	0.2	0.5	0.2	0.8

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Time-course by country/region

Among the Hubei population of 58.5M, the incidence rose from the first case reported in late 2019 to about 60 new cases per million people per day by 02-05 and then steadily declined (Fig 1) from ~4000 on 02-05 to below 50 cases per day since 03-08.

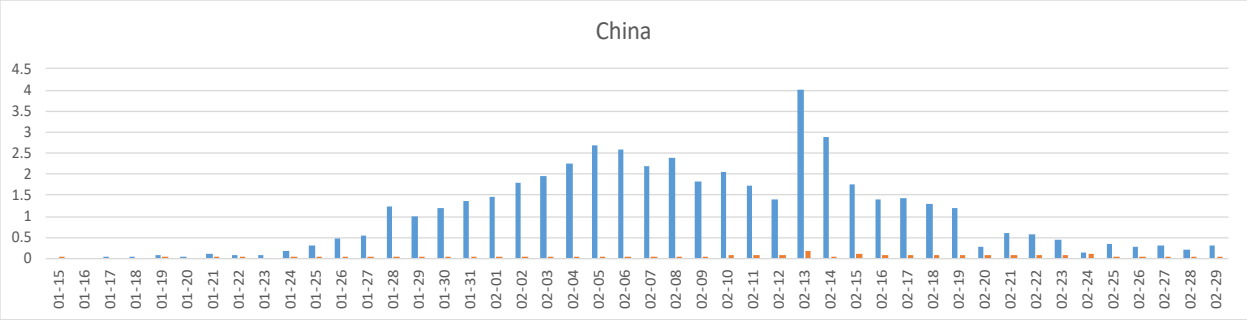


Fig 2: COVID-19 cases in Mainland China. Blue: cases/M/d, red: deaths/M/d. Around Feb 13, the case definition was expanded, resulting in additional cases from previous days being added. Hence, the 02-13 cases have been truncated. Most cases were seen in the Hubei province of 58.5M people (see Table 1 for population sizes).

By Mid-January 2020, the first cases of COVID-19 were seen in other Asian countries, but incidence remained below about 3/M/d outside of continental China, although Malaysia/Brunei, Singapore, and the Philippines have been rising to about 5/M/d. Japan, instead, with the largest proportion of people >65 years of age in the world (28%) remains below 1.5/M/d. Still, the upward curvature (Fig 3) might indicate an emerging epidemic with no indication of a turning point, yet.

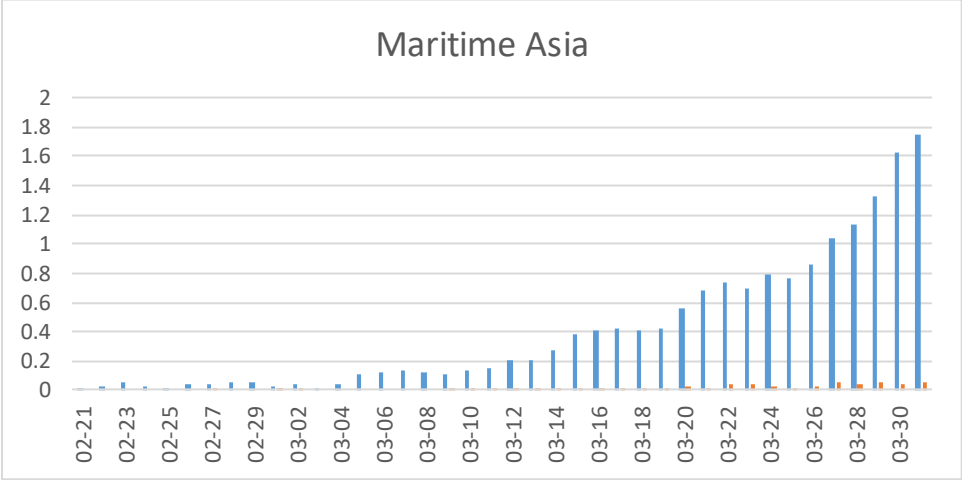
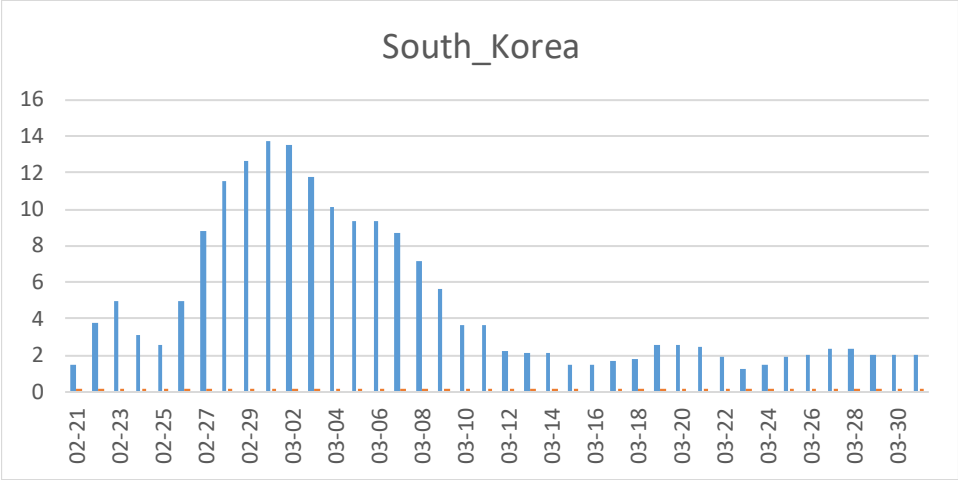


Fig 3: COVID-19 cases in Maritime Asia. See Fig 1 for legend.

In continental South Korea (population 51M, cumulative incidence 192/M), however, the incidence soon rose to a peak of about 14/M/d between 02-29 and 03-02, before declining to less than 150 cases per day (2.5/M/d) since 03-12 (Fig 2a).

In Iran (cumulative incidence 385/M), incidence rose about a week after South Korea. The top incidence before 03-23 (~15.5 cases/M/d) was about the same. The increase thereafter to 35–40M/d may indicate a "rebound" into a population not sufficiently immunized by the previous wave(s). Lethality in Iran was notably higher and followed the increase in cases with the expectel delay of several days (Fig 2b).

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .



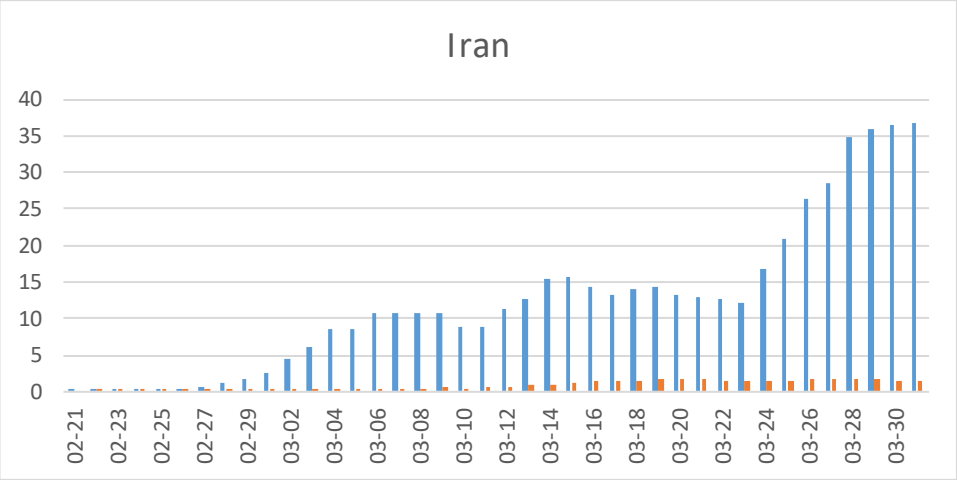


Fig 4: COVID-19 cases in South Korea and Iran. See Fig 1 for legend.

From 03-19 to 03-20, several European countries have seen a more than two-fold increase in the number of cases reported (Germany: 570%, San Marino: 340%, Ireland: 260%, Switzerland: 240%, Austria: 202%). As *natura non facit saltum* (Darwin: nature doesn't jump),[(Berry 1985)] such abrupt increases must be, at least in part, the result of reporting or other artifacts. In Germany, for instance, the reporting system was changed between 03-16 and 03-19, so that the number likely includes cases previously reported only through a parallel system. France, Italy, and Spain also reported an unusual increase by 27–35 percent. Since 3-26, all these countries reported the incidence to decline.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

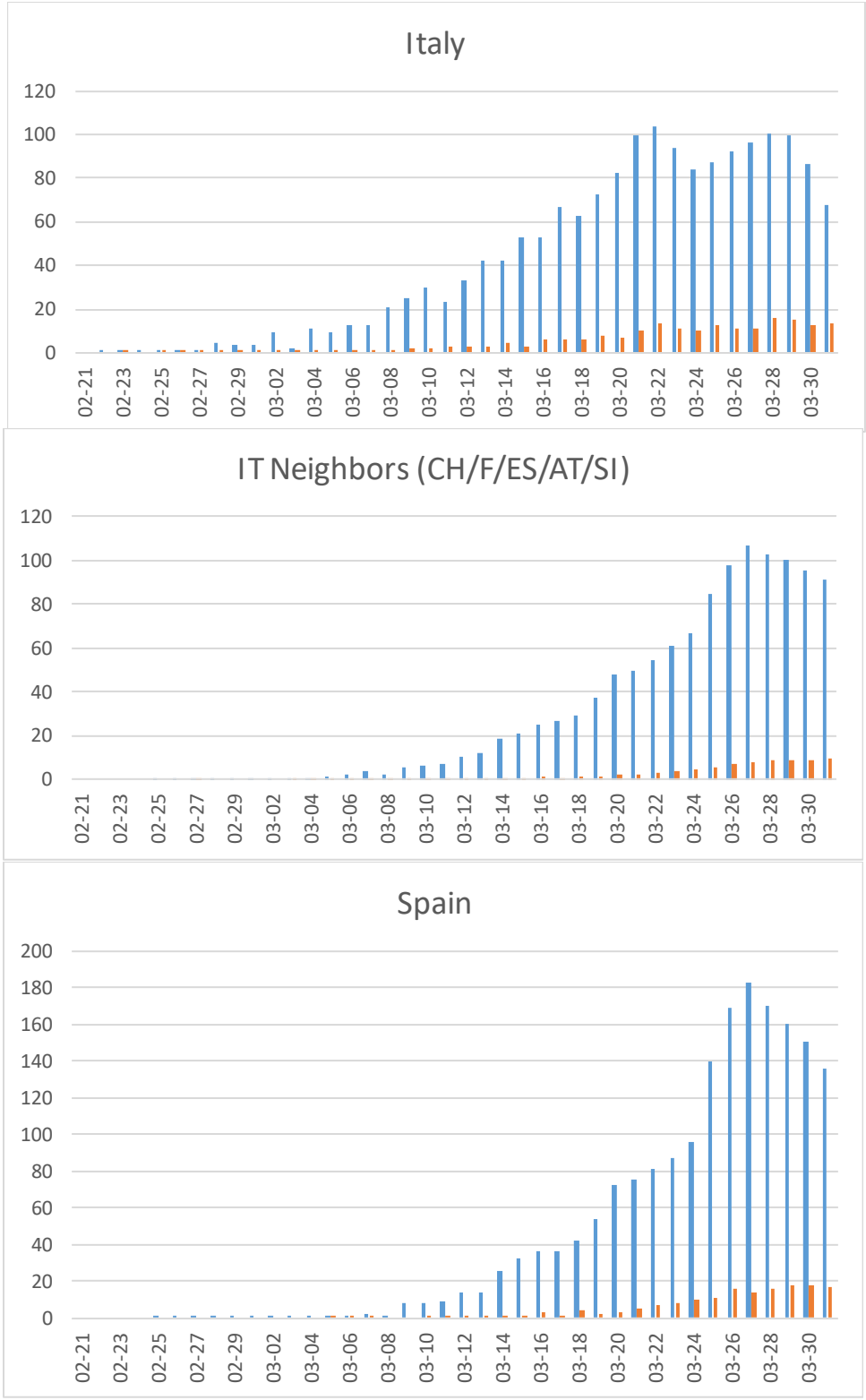


Fig 5: COVID-19 cases in Italy and it's neighboring countries. Italy (top), European countries neighboring Italy (IT+, middle. Spain also shown separately**.** See Fig 1 for legend.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Among European countries with a population of more than 2M, Italy's neighbors Spain and Switzerland now have a higher cumulative incidence per capita than Italy (Table 1, Fig 3), all higher than that in the Hubei province (their all with a population of 50–60 M is also similar).

The overall epidemic in Europe (Fig 4a) is a population weighted average of the high incidence regions (Fig 4a, weighted average of Fig 3) and the remaining low incidence countries (Fig 4c).

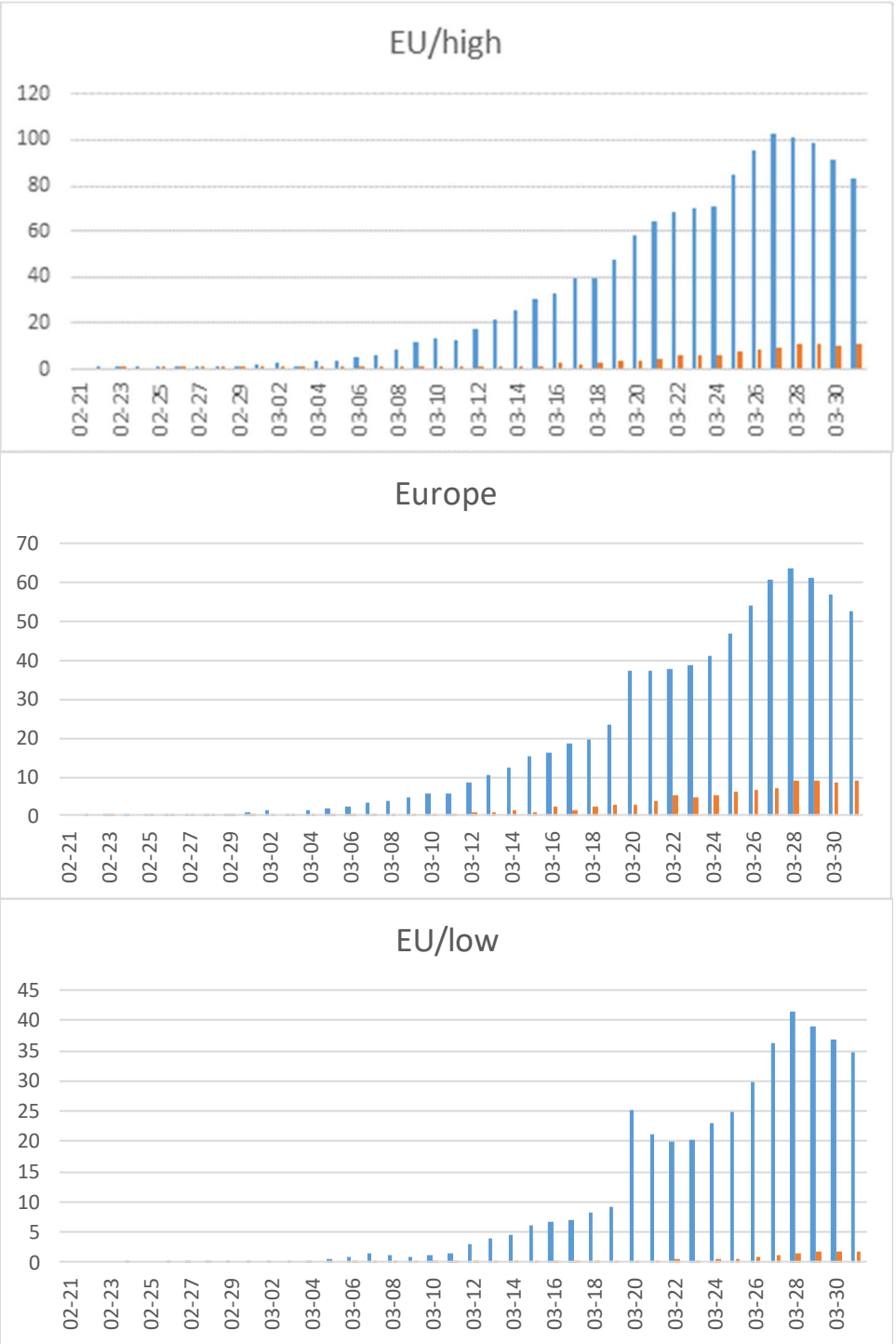
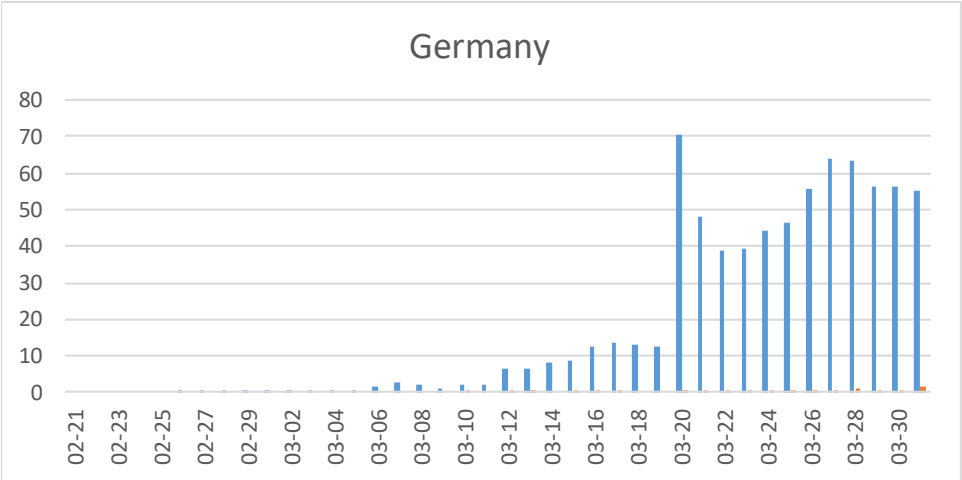
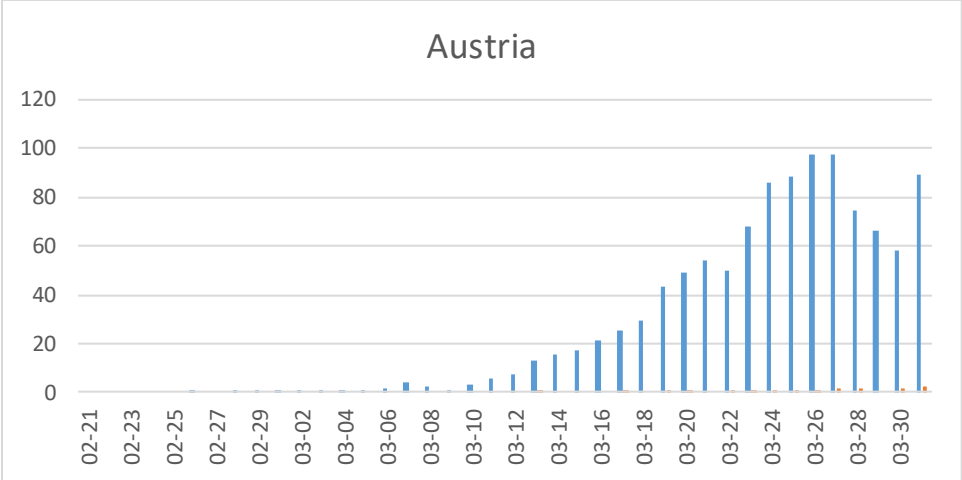


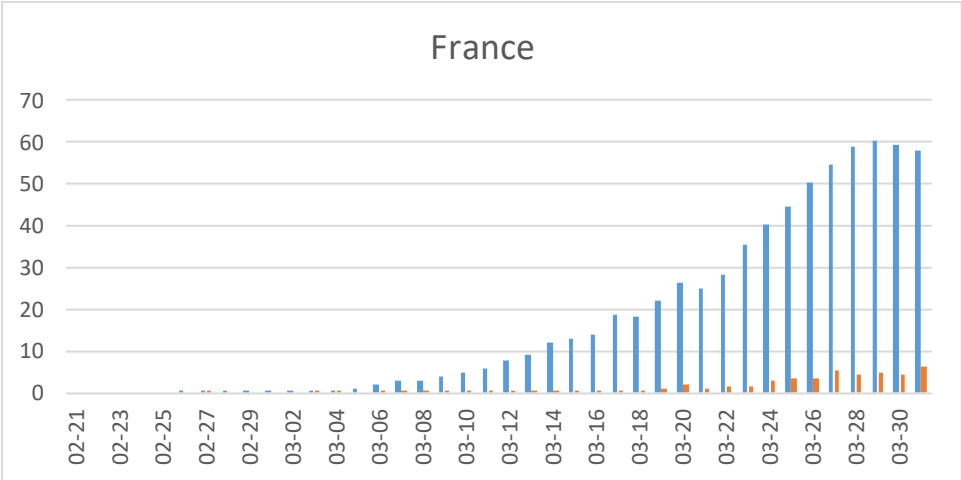
Fig 6: COVID-19 cases in Europe. Early onset/high lethality (IT and neighbors, top), total (middle), and late onset/low letality (other European countries, bottom)**.** See Fig 1 for legend.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

The incidence in Italy has stayed below the 100/M/d peak for 9 days, after rising for about 4 weeks from 1/M (02-26..~03-22, neighbors: 03-01..~03:26), while Hubei's and South Korea's rose over 2 weeks only (01-19..02-05, Fig 2, and 02-21..03-06, Fig 4). Other European countries peaked in 03-27/28. Germany is notable for reporting 61,913 cases, but only 583 deaths (0.9%, by 03-31), less than even Austria and much less than France, and, in particular, Italy and Spain.







medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Fig 7: COVID-19 cases in selected European countries. See Fig 1 for legend. Data in Germany is based on cases reported electronically to the Robert-Koch-Institute (RKI) and transmitted to the ECDC, but the RKI also provides two sources of data on its Web site that are difficult to reconcile with these data.

Of particular interest is Sweden, where no "social distancing" policy was implemented. Still, there was no difference in the shape of the epidemic or the height of its peak to the other Scandinavian countries:

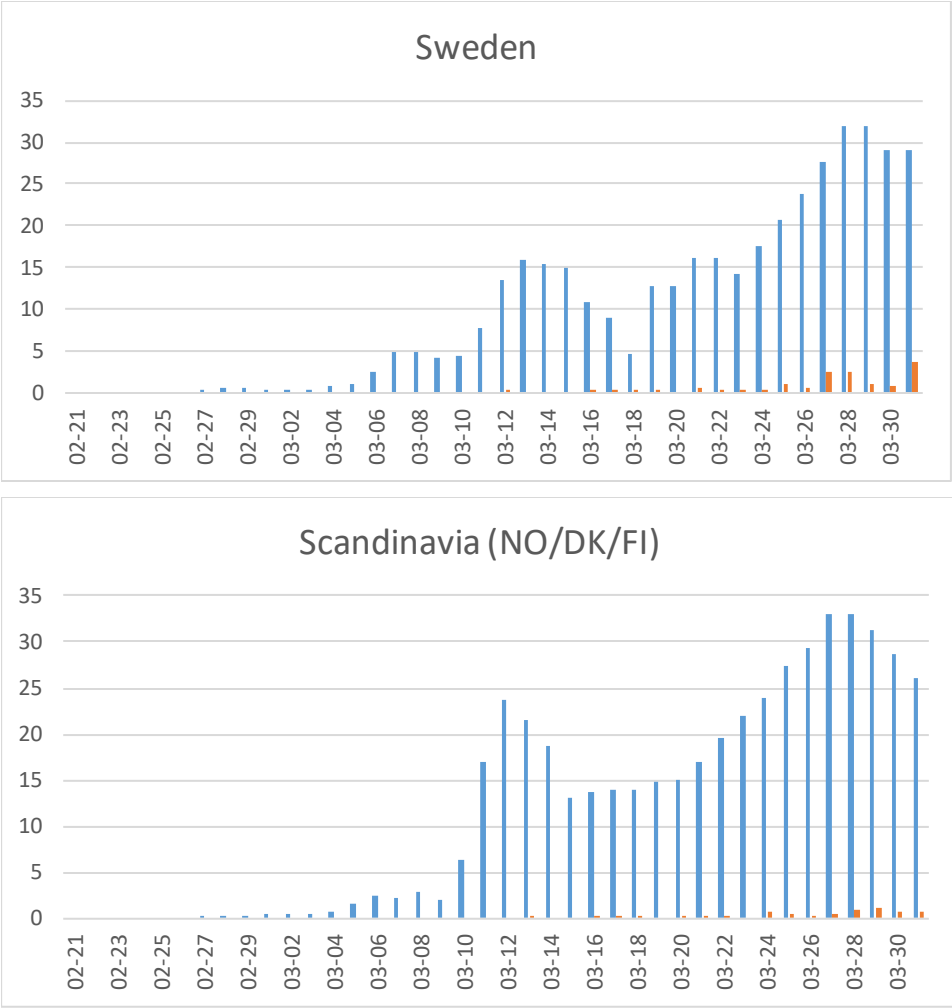


Fig 8: COVID-19 cases in the Scandinavia. See Fig 1 for legend.

The epidemic in North America started later, especially in the US (except for a few isolated cases likely imported directly from Asia). The increase in incidence is consistent with the dynamics of an earlier epidemic and the cumulative incidence in the US (497/M) is still lower than in Europe (747/M). The US may have reached the "turning point", where the rate of new cases begins to decline (when 50% of the peak incidence is reached) at about 03-26, compared to Europe's 03-22. Reporting in Canada is not consistent enough to draw conclusions, although levels may remain below those in the U.S. Even in New York City, which is particularly hard hit, the number of new cases seem to be leveling off, albeit the number of deaths is still increasing, indicating an extremely high lethality approaching 25%, even higher than Italy's.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

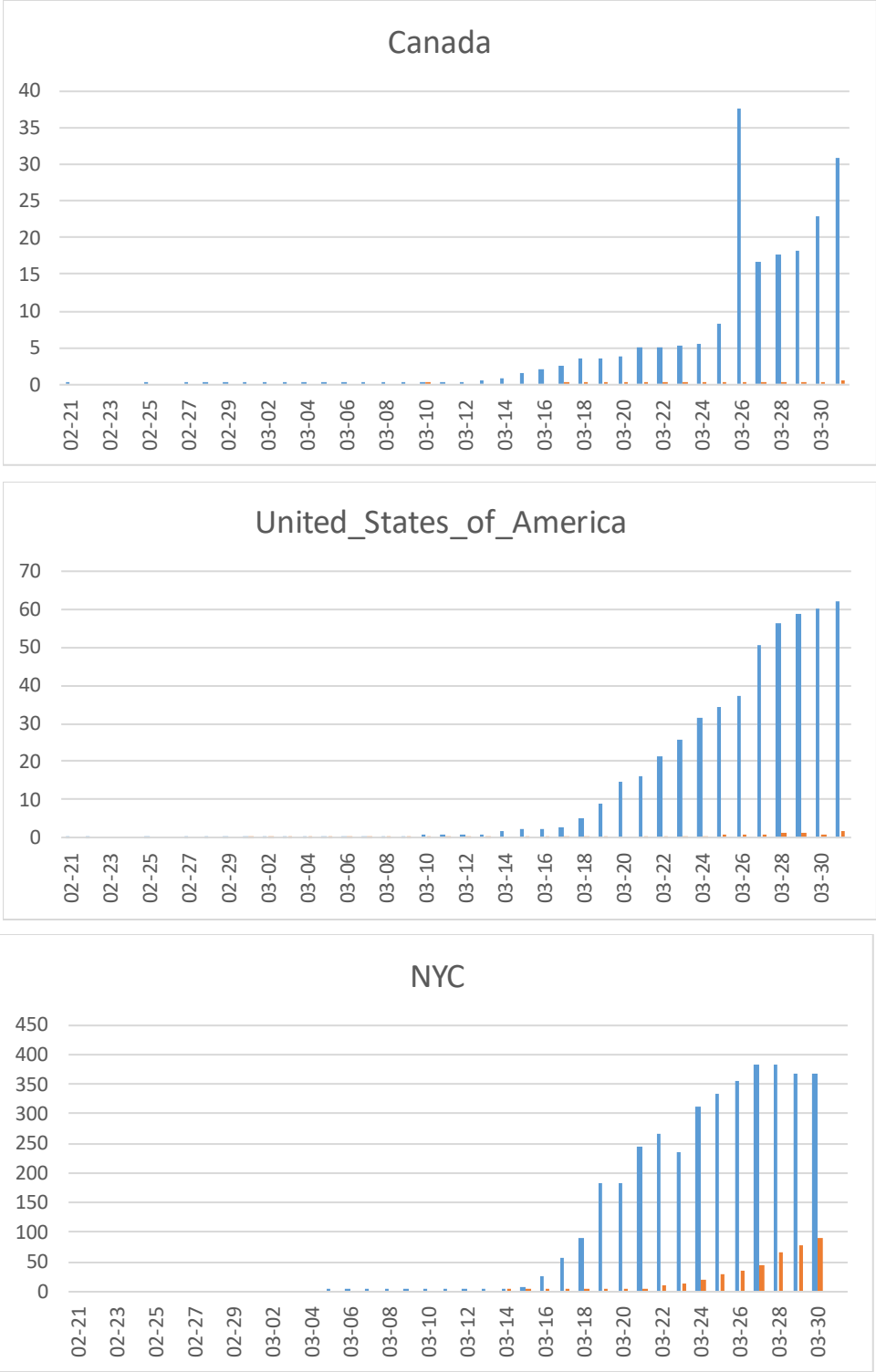


Fig 9: COVID-19 cases in the North America. See Fig 1 for legend. The high peak in the Canadian data on 03-26 cannot be reconciled with the dynamics of a respiratory disease spreading.

Incidence by Country. Southern Hemisphere

Most parts of the southern hemisphere (with the possible exception of Chile) have seen only few cases, but Oceania (Australia and New Zealand) shows evidence for an epidemic with a peak

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

incidence on 03-26/27. As in Canada and, earlier, in Germany and other European countries, isolated spikes tend to reflect delayed reporting, rather than changes in incident trends. The number of deaths (20/5204 = 0.4%) is extremely low (Fig 10).

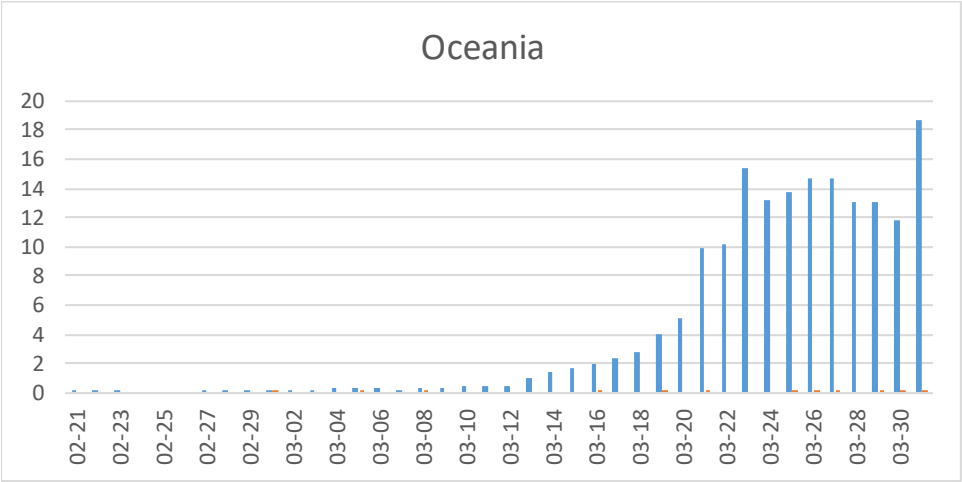


Fig 10: COVID-19 cases in the Oceania (Australia and New Zealand). See Fig 1 for legend..

Modeling Results for Effectiveness of Containment (Social Distancing)

The effect of reducing the reproduction number by reducing the number of contacts ("containment", "Social Distancing") depends on when it starts in the course of the epidemic. Fig 11 shows the effect of a one- month intervention cutting R_0 in half starting at the point of the peak prevalence of infectious subjects. Compared to Fig 1, the duration of the epidemic is shortened, albeit at the price of reducing the R/S ratio, so that a subsequent epidemic with the same or a similar virus (cross-immunity) could start earlier.

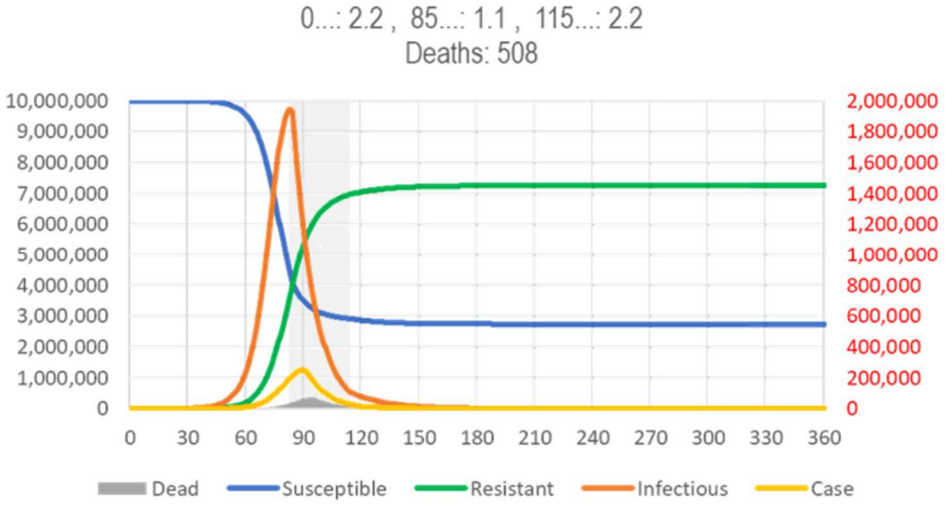


Fig 11: SIR Model of SARS, Window of Opportunity for Fast Eradication of the Epidemic. (see Fig 1 for legend). The gray area indicates the period where containment can give a "coup de grace" to a respiratory disease epidemic. The more narrow bell curve with a post-peak intervention indicates the reduction in number of infections and, thus, deaths.(spreadsheet for model calculations available from https://app.box.com/s/pa446z1csxcvfksgi13oohjm3bjg86ql)

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Fig 12 shows a one-month intervention starting about two weeks earlier, at the turning point where the curve of the new cases changes from increasing faster to increasing more slowly. This intervention reduces the number of deaths, but the epidemic is extinguished two months later and the R/S ratio ("herd immunity") is further decreased.

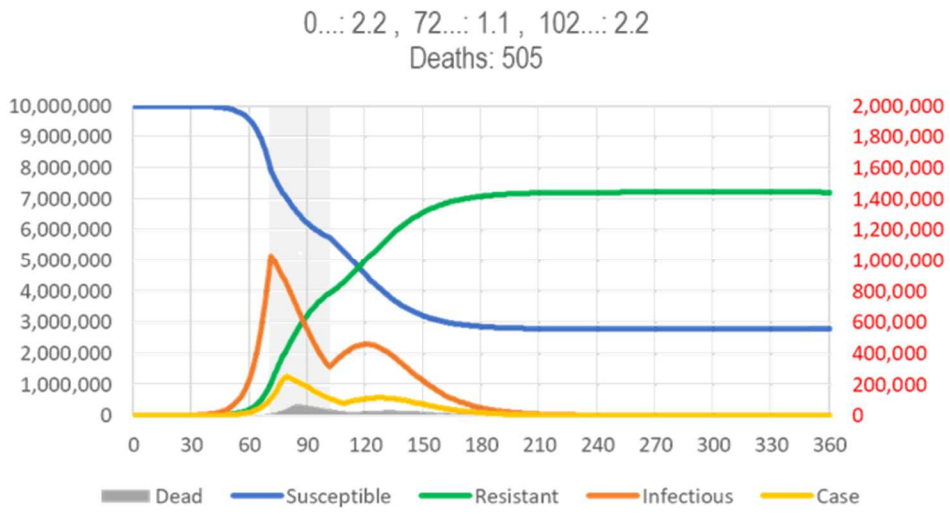


Fig 12: SIR Model of SARS, Window of Opportunity for Maximal Reduction of Total Deaths. (see Fig 1 for legend). The gray area indicates the period where containment can have the most impact on total number of deaths. However, the epidemic is not eradicated.(spreadsheet for model calculations available from https://app.box.com/s/pa446z1csxcvfksgi13oohjm3bjg86ql)

Fig 13 shows the effect of an intervention that starts even earlier, about two weeks before the intervention in Fig 12. Even if the intervention is extended from one to four months no herd immunity is created and, thus, the epidemic rebounds and will run eight months, instead of three (Fig 1) or less (Fig 11).

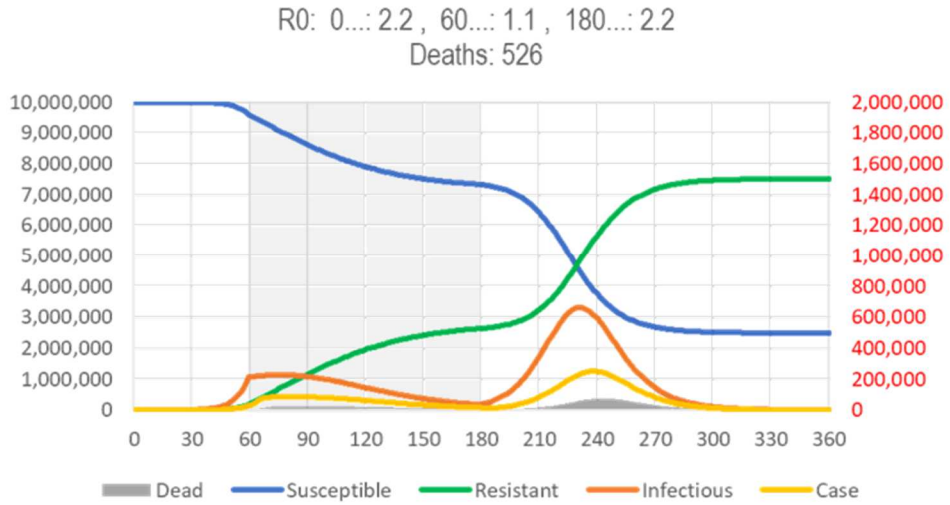


Fig 13: SIR Model of SARS, Effect of Early Social Distancing / Lockdown. (see Fig 1 for legend). It is assumed that a highly effective intervention reduces R_0 by 50% for 4 months, beginning after the appearance of a novel type of cases is noticed. The proportion of symptomatic cases (0.05%/d, i.e., .35% of infected people will become cases and the proportion of cases to die (2%) may change, but the issues discussed here are broadly independent of these assumptions.(spreadsheet for model calculations available from https://app.box.com/s/pa446z1csxcvfksgi13oohjm3bjg86ql)

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

In summary, there is a narrow window-of-opportunity for interventions ("flattening the curve") aiming to improve public health by reducing R_0 to be successful:

- Starting after the peak **prevalence** (of infections) has little effect (not shown). The curve goes down, but is not "flattened".
- Starting at the peak **prevalence** gives the epidemic a "coup de grace", shortening its duration, albeit at the price of reducing the R/S ratio. The curve is narrower, but also not "flattened" (Fig 11).
- Starting at the peak **incidence** "flattens" the curve without broadening it and maximizes the number of deaths prevented during the current epidemic, but reduces herd immunity and, thus, the chance of another epidemic coming sooner (Fig 12).
- Starting before the peak **incidence** "flattens the curve", but also broadens it and causes a rebound, unless the intervention is continued for many more months (Fig 13).

It is herd immunity that stops the spread of an infectious disease, so in general, one would want to let the epidemic initially run its natural course (or even accelerate it, as people have traditionally done with "measles parties") to build immunity as fast as possible.

To reduce the duration of the epidemic and its impact on the economy (and also increase the time until the next epidemic can spread), one would wait until the prevalence of infectious people (I) reaches its peak (in the above model: day 83, red).

Without repeated broad testing, however, this date cannot be directly observed, but it is known that peak prevalence of infected people is followed about a week by peak number of new cases. This is too late to make a decision, but the SIR model shows that this peak is preceded by two weeks by the "turning point" in cases where the curve of the new cases changes from increasing faster to increasing more slowly (day 76). The turning point can be estimated from the observed cases in time to making a decision. (It is also about 50% of the peak number of new cases, which one might be able to predict.) Hence, peak prevalence (of infections) follows the turning point/half peak (in number of cases) by about a week. **The window of opportunity for starting an intervention is the week following the turning point in number of cases per day.**

Discussion

Strengths and shortcomings

A major strength of this analysis of the epidemiological data is that it does not rely on epidemiological models with questionable assumptions. Instead, the results reflect raw incidences over time as reported by the ECDC, depicted by country or region of neighboring countries.

A shortcoming of such an entirely data-based approach is that it lacks the sophistication and potential additional insights that could come from fitting, e.g., differential equation models. The only modeling assumption made is that curves should be "smooth" (except when reporting artifacts are suspected), but even then, data were redistributed only to the directly neighboring day.

Still, the evidence is strong enough to draw qualitative conclusions about possible scenarios for the spread of SARS-CoV-2 in the near future. Also, the results suggest strategies to explore the variability of the SARS-CoV-2 virus strains and to select prevention strategies.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Evidence for (at least) two different strains of SARS-CoV-2

During the 2003 SARS epidemic the number of new cases peaked about three weeks after the initial increase of cases was noticed and then declined by 90% within a month. Table 1 shows the relevant timepoints for the 2020 COVID-19 epidemic. The SARS-C0V-2 data also suggest that it takes at least a month from the first case entering the country (typically followed by others) for the epidemic to be detected, about three weeks for the number of cases to peak and a month for the epidemic to "resolve". This data is consistent with the results from the SIR model (see Epidemiological Models).

Table 2: Epidemiological Timepoints by Country Top: 2003 SARS, Bottom: 2020 SARS-CoV-2

	First cases	Gap [d]	Begin (>.1/1M)	Gap [d]	Peak	Gap [d]	End (<.1/1M)	Begin to End [d]	
CA/SGH	02-23		03-02	18	03-20	28	04-18	28	(Svoboda 2004)
CA/NYG	04-20		04-27	30	05-27	6	06-03	36	(Svoboda 2004)
Guangdong	2002	>50	01-19	22	02-11	81	05-02	103	(Cao 2019)
Shanxi			03-15	34	04-19	30	05-19	64	(Cao 2019)
Beijing			03-05	49	04-24	30	05-24	69	(Zhou 2003; Cao 2019)
Mongolia			03-22	22	04-14	40	05-24	62	(Cao 2019)
Hebei			04-04	20	04-24	24	05-19	44	(Cao 2019)
Tianjin			04-14	8	04-22	12	05-04	20	(Cao 2019)
HK/PWH	02-21		03-11	6	03-17				
HK/AG			03-20	4	03-24	29	04-13		(Zhou 2003; Leung 2004)
HK/					04-12	54	06-06		
Taiwan	02-24	23	03-17	38	04-25	50	06-14	88	(Small 2003; Yeh 2004)
Singapore			03-15	20	04-05	24	04-29	44	(Small 2003; Goh 2006)
Vietnam			02-23	11	03-04	28	04-06	39	(Shi 2003)
Median 2003			03-15	20	04-13	29	05-19	44	
China (Hubei)	2019	>50	01-17	18	02-05	27	03-18	60	0 in Hubei
S- Korea	01-20	31	02-19	12	03-01	>25	>03-25	>37	Ongoing low level
Iran	02-20	2	02-22	14	03-07				Several waves?
Italy	01-31	22	02-22	30	03-22				2nd wave?
Germany	01-28	30	02-27						Reporting system
France	01-25	31	02-27	>30	?04-01		~04-15?		
US	01-25	40	03-05	>30	?04-o7		~05-01?		
Median 2020		31	02-22	25	03-25		~04-15?		

The 2003 SARS and the 2020 SARS-CoV-2 are not only similar with respect to genetics (79% homology),(Lu 2020) immunology,(Ahmed 2020) involvement of endocytosis (also with influenza and syncytial viruses),(Behzadi 2019) seasonal variation (same season in the northern hemisphere also with influenza, syncytial, and metapneumo viruses)(Olofsson 2011), evolution (origin in bats, 88% homology),(Benvenuto 2020; Malik 2020) but also with respect to the duration between emergence and peak of cases as well as between this peak and resolution of the epidemic (Table 1). Based on these similarities, one could predict the COVID-19 epidemic to end before 04-15 in Europe and about two weeks later in the U.S.

The time and height of the peak incidence of cases in the different countries are consistent with the hypothesis that SARS-CoV-2 moved step-by-step westward from China, via other Asian countries to Middle East (Iran. Qatar, and Bahrain), Southern Europe (Italy, followed by its neighbors CH/F/ES/AT/SI), central and northern Europe, and, finally, the US.

Viruses improve their "survival" if they develop strategies to coexist with the (human) host.(Woolhouse 2007) Multiple coronaviruses have been found to coexist in bat populations.(Ge 2016) The emerging COVID-19 data is consistent with the hypothesis that (at least) two SARS-CoV-2 strains have developed. One strain, which traveled through South Korea, remained more infectious, while the

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

other strain, which traveled through other Asian countries lost more of its infectiousness. The strain that passing through South Korea and then Iran and Italy (SKII strain) showed high lethality in Iran and Italy, but less lethality when it traveled to Italy's neighbors, either because of differences in health systems, because the strain mutated back, or because a strain arriving directly from Asia had the advantage of spreading first. Only sequencing samples from these countries can help to answer these questions.

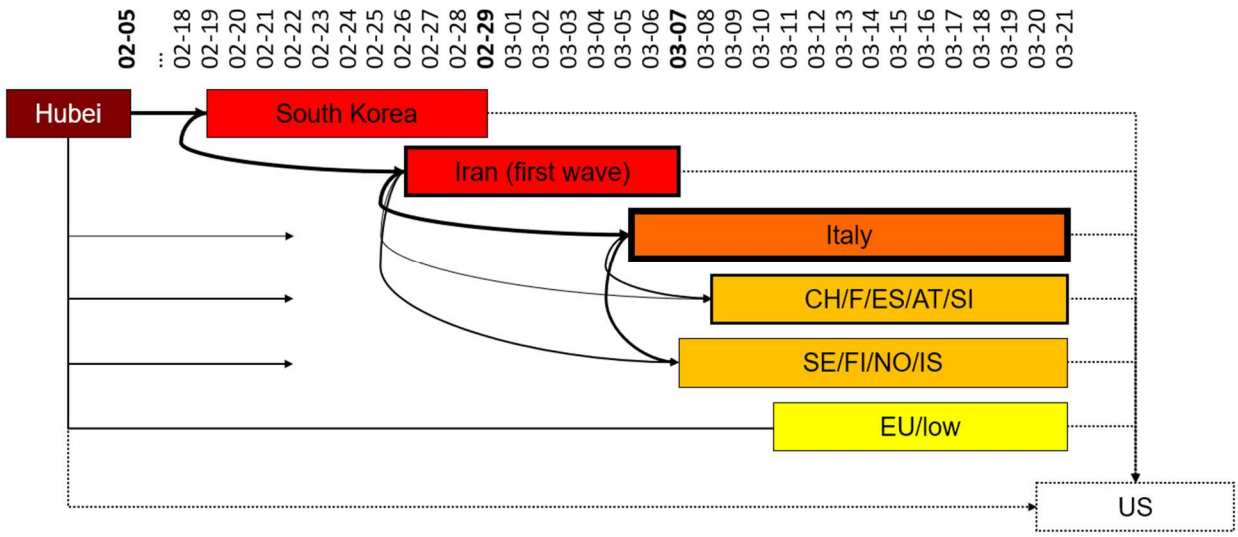


Fig 14: Hypothetical virus transmission pathways. Connection width: number of contacts, box colors: infectiousness, box borders: lethality, dotted connections/borders: unknown. The end date of a box indicates the date of peak incidence, if known (bold date).

Changes in infectivity and lethality between China and Europe

Mainland China is not reporting relevant numbers of novel cases anymore and Hubei reports no new cases since 03-19. The number of new cases in South Korea also has declined to low levels since its peak around 02-30. Maritime Southeast Asia continues to show low levels of new cases only (<3/M/d), with the possible exception of Singapore and Malaysia/Brunei. Japan and the Philippines are slowly approaching a 2/M/d incidence.

The data are consistent with the same "SKII" strain traveling from China via South Korea and Iran to Italy. Iran was hit about a week after South Korea (around 03-07), with a similar peak incidence, but higher lethality (red bars in Fig 2). The data also suggests a second wave of infections in Iran, which may have peaked on 03-15. Italy was hit a week after the first wave in Iran (which peaked around 03-07/08, Fig 4a). Incidence in Italy reached a substantially higher incidence than reported in Iran. The peak incidence in Italy may have been reached on 03-22 (at about 100/M/d).

Without sufficiently detailed genetic data, it is not clear whether the high lethality in Italy is due to genetic variations in the virus or to Italy having the second oldest populations in the world (after Japan). A 03-20 report by the Istituto Superiore di Sanita,[(COVID-19 Surveillance Group 2020)] however, implicates that age and comorbidities played a role – among 3200 deaths, mostly in Lombardy and Emilia-Romana, median age was 80 years (IQR 73-85, only were 36 below the age of 50), 98.8% had at least one comorbidity (hypertension: 74%, diabetes: 34%, ischemic heart disease: 30%, atrial fibrillation, 22%, chronic renal failure: 20%, …).

The epidemiological data does not support the hypothesis that SARS-CoV-2 spread from Munich in Germany to Italy.[(Kupferschmidt 2020)] Instead, the virus may have spread from Italy to its neighboring

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

countries, Switzerland, France, Spain, Austria, and Slovenia, within just a few days of arriving from Iran. The top incidence seems to be less than half of that in Italy and the lethality is lower, too. While Italy has many people 65 years and older (23%, second only to Japan data.worldbank.org/indicator/SP.POP.65UP.TO.ZS), the relatively small differences or age distribution within Europe (e.g., Germany: 21%) are unlikely to account for much of this difference. A possible explanation (indicated in Fig 6) is that the less virulent strain(s) arriving from other parts of Asia may have had a head start in those countries, so that imported infections from Italy met subjects who had already developed (cross) resistance against both strains.

The parts of Europe not directly hit by the SKII strain leveled off at (about 40/M/d). Overall, incidence in Europe seems to have leveled off at about 60/M/d (and recent reporting irregularities seem to have less of an effect). There is strong evidence for incidence now decreasing.

Predictions for COVID-19 in North America

From Table 1, SARS-CoV-2 has arrived in the U.S. almost a week after it arrived in Europe. The incidence is still low (currently at about 60/M/d) but already shows signs of leveling off. If incidence in the U.S. were to peak at about 75/M/d, higher than Europe as a whole, but less than Italy and its neighbors (Fig 4), one would expect new cases to peak at up to 25,000 per day and the cumulative incidence could reach 600/M (3 times the number of cases per million people in South Korea to account for a longer course because of the size of the countries) or a total of 200,000 cases, and, at twice of the currently the observed 1.7% lethality in the US (compared to South Korea's 1.5%), about 6,000 deaths, 2,502 of which have already been reported. In South-Korea, the epidemic there completed its course, economy and political system are comparable to the US, and the proportion of people older than 65 years is similar (14% vs 16%). On the other hand, the numbers could double if the SKII strain should have hit the US as the first strain. A number of 6000–12000 U.S. death over the course of the epidemic compares to an expected number of 16,000–78,000 influenza deaths per season from pneumonia and respiratory/circulatory complications alone, which also occur predominantly among people at 65 years of age and older.(Rolfes 2018).

The precise number of people dying depends on (a) which virus strain got to the US first and (b) how early people are being treated against severe complications (e.g., pneumonia).

A historical perspective

This is not the first, and likely not the last time, that well-intentioned public health policies are inconsistent with our understanding of how epidemics spread. For instance, during much of the HIV epidemic, there was widespread fear that HIV could establish itself in the population as a whole, even though the data (including data showing absence of transmission to the wives of hemophiliacs)(Wittkowski 1995a) and models(Wittkowski 1992; Seydel 1994) contradicted this fear.(Wittkowski 1995b; 1996) These results have been repeatedly confirmed.(Centers for Disease Control and Prevention 2019; Haddad 2019) In the case of heterosexual transmission of HIV one could argue that there was little risk associated with a the public health policy promoting condom use, but in the case of COVID-19 prevention, ignoring models and data may carry substantial risk.

During the AIDS epidemic, epidemiologists had the advantage that, in addition to the date of report, the date of diagnosis was available for analysis so that variations in reporting delays, such as mid-February in China, 03-20 in Germany, and 03-26 in Canada, could be accounted for. Unfortunately, the public COVID-19 data lacks that information.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Implications for prevention

A major problem with respiratory diseases is that one cannot stop all chains of infections within families, friends, neighbors, … . Even after a couple of weeks of "lockdown" there will be a few infectious persons, and as long as there are enough susceptible people in the society, this is enough to re-start the epidemic until there are enough immune people in the society to create "herd immunity". Hence, one would expect the cases to appear in waves (Fig 13, the period of the "lockdown" corresponds to March to May, 2020 in the U.S.). Such waves of cases have been seen in different countries and the longer than expected duration of the epidemic supports the hypothesis that the social distancing / lockdown interventions had some effect, albeit at a high cost for approx. 10% of deaths saved.

The longer time from onset to peak in Europe (4 weeks) vs China, South Korea, and the first wave in Iran (two weeks) is also consistent with premature interventions (to "flatten the curve") prolonging the epidemic ("broadening the curve"). In Oceania, incidence rose form 1/M/d on 03-13 to a peak of 15/M/d on 03-27, within the two weeks predicted by the model for epidemics taking its natural course. Other than travel restrictions (14 day quarantine for non-residents and travel restrictions between states, similar to what other countries have enacted) the AU government has provided funds for (https://www.health.gov.au/news/health-alerts/novel-coronavirus-2019-ncov-health-alert/government-response-to-the-covid-19-outbreak, accessed 03-30)

- delivering an AU$20.7B (AU$800 per person) support package (investment, jobs, health)
- opening fever clinics and funding home delivery of prescription medicines.

A peak incidence of 15/M/d and a total incidence of 155/M are similar to South Korea (14/189) and comparable to less densely populated European countries (Slovakia: 20/168, Finland: 16/203, Greece: 10/116). Hence, the epidemic of Oceania, together with those of China, South Korea, and Iran are consistent with the model results suggesting that a natural Covid-19 epidemic peaks two weeks after the first cases are seen, and then declines with financial and medical assistance from the government to prevent deaths to reduce the burden on the health system and damage to the economy. As in Sweden, no "social distancing" rules were imposed.

This analysis of the publicly available data suggests that at the time Italy imposed quarantine on the Lombardy and adjacent regions on 03-08, the SKII virus strain had already reached the adjacent countries (Switzerland, France, Spain, Austria, Slovenia). Even though the lockdown started early (03-08), which may have caused a rebound consistent with a decline in compliance.

In the US, the "turning point" (50% of peak incidence) may have been reached around 03-27. Still, New York, New Jersey, and Connecticut restaurants were ordered to closed from 03-20; the shutdown of California was ordered on 03-19 (Executive order N-33-20). As social distancing was ordered before the epidemic reached its turning point, a "flattened curve" is to be expected, but the curve will also be broader. (The virus will stay longer in the population.)

Some containment strategies could even be counterproductive in other ways. For instance, the simple model used in Fig 13 does not account for age-stratification. In diseases such as COVID-19, children develop mostly mild forms, elderly people have a high risk of dying.(Zimmermann Curtis 2020) Hence, containment of high-risk groups, like elderly people in nursing homes (see the Washington State example) is highly effective in protecting them from becoming infected and in reducing the pool of children and young adults that would have to be infected to reach herd immunity. A substantial increase in the duration of the epidemic by preventing immunity to develop among the young, however, might make effective containment of the elderly more difficult and, thus, increase the number of deaths among the elderly.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

In the U.S. as a whole, the "turning point" in new cases cannot be earlier than 03-25. Case data for New York City is inconsistent, but the New York Times reports that New York and Detroit reached the turning point on 03-19. Under this assumption, the optimal time point for starting a New York public health intervention to reduce duration and impact of the COVID-19 epidemic was around 03-27. Social distancing in NY, one of the epicenters of the U.S. epidemic, started about 03-17 (day 73 in the model) with restaurants being closed, and intensified on 03-22 (day 78) with all non-essential businesses being closed. Restrictions have been discussed to be lifted on 04-12 (day 100, because of the Easter holiday) or 04-30 (day 117). The model predicts that an intervention until Easter would reduce the number of deaths in New York and Detroit (and possibly some other parts of the U.S., but the virus would still linger on for another three months, so that it would not be safe for (elderly) high-risk people to participate in this year's Easter activities. Extending the social distancing until the end of April would substantially reduce, but not eliminate, the number of infectious people remaining in the population (Fig 15).

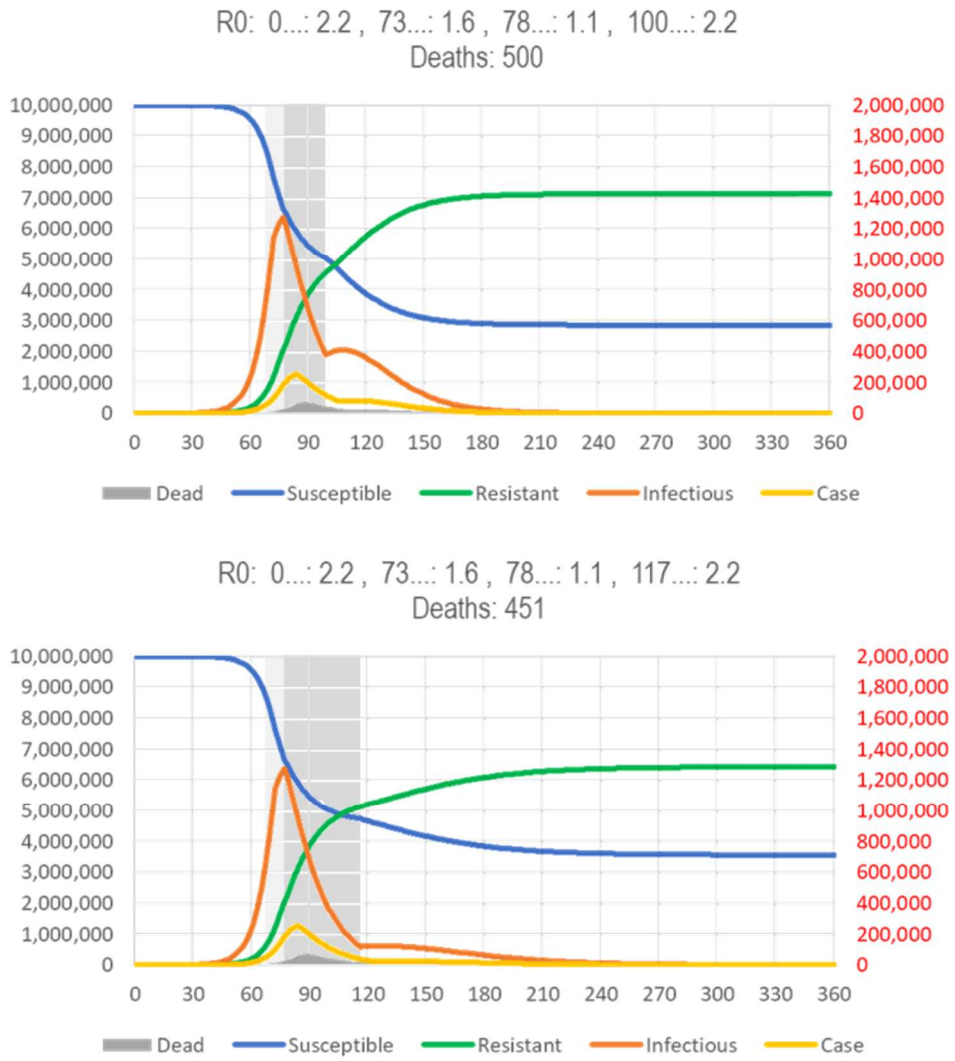


Fig 15: SIR Model of SARS, Phased in Restrictions. (see Fig 1 for legend). The gray areas indicate the periods of low (5 d) and high intensity (22/40 d) restrictions)

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Conclusions

Until a vaccine will become available, the only pharmacological strategy to reduce the number of deaths is to reduce the damage the infection (and immune system) does, e.g., by reducing the initial viral load,[(Chu 2004)] and making sure that people get treated at the earliest signs of pneumonia.

As the curve of new cases is flattening in the U.S. at about 20,000 per day and the proportion of cases dying is stabilizing at ~2% (with regional differences), the number of deaths is stabilizing at about 400/d (a quarter of them in NYC) and beginning to decrease (a worldwide characteristic of COVID-19) with a one-week delay. With an expected total of less than 10,000 deaths, social distancing may have saved 1000 deaths at a cost of $2T ($2B per case) so far, but prolonging the epidemic may also have increased the number of deaths, because preventing exposure of young people early on means that more of the elderly become infected before herd immunity is reached.

Aside from separating susceptible populations (elderly and high-risk subjects, e.g., in nursing homes) from the epidemic, which is effective as long as virus is circulating, public health intervention aiming to contain a respiratory disease need to start within a narrow window of opportunity starting at or a week after the curve of the new cases changes from increasing faster to increasing more slowly. Only if stopping the epidemic from generating a sufficient number of immune people is avoided can containment efforts stop after about a month or two (depending on late or early start, respectively), when the ratio of infectious vs immune people is low enough for preventing the disease from rebounding. When the window of opportunity has been missed, containment has only limited impact on the course of the epidemic, but high impact on economy and society.

As epidemics in maritime Asia and the southern hemisphere remain a possibility and some develop within the next weeks, the use of data and scientifically sound models may help to time interventions to optimize their effect.

To determine that time point, case data collected and reported needs to contain not only the date of report, but also the date of “diagnosis” and whether the patient had clinical symptoms or was merely tested positive and whether the patient was positive for circulating virus RNA/DNA (currently infectious) or antibodies (already immune).

References

Ahmed SF, Quadeer AA, et al. (2020). Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies. *Viruses* **12**(3).

Behzadi MA, Leyva-Grado VH (2019). Overview of Current Therapeutics and Novel Candidates Against Influenza, Respiratory Syncytial Virus, and Middle East Respiratory Syndrome Coronavirus Infections. *Frontiers in microbiology* **10**: 1327.

Benvenuto D, Giovanetti M, et al. (2020). The global spread of 2019-nCoV: a molecular evolutionary analysis. *Pathog Glob Health*: 1-4.

Berry RJ (1985). Natura non facit saltum. *Biological Journal of the Linnean Society* **26**(4): 301-5.

Cao W, Fang L, et al. (2019). What we have learnt from the SARS epdemics in mainland China? *Global Health Journal* **3**(3): 55-9.

Cascella M, Rajnik M, et al. (2020). Features, Evaluation and Treatment Coronavirus (COVID-19). *StatPearls*. Treasure Island (FL), StatPearls Publishing.

Centers for Disease Control and Prevention (2019). Estimated HIV incidence and prevalence in the United States, 2010-2016. *HIV Surveillance Supplemental Report* **24**(1).

Chu CM, Poon LL, et al. (2004). Initial viral load and the outcomes of SARS. *Cmaj* **171**(11): 1349-52.

COVID-19 Surveillance Group (2020). Characteristics of COVID-19 patients dying in Italy Report based on available data on March 20th, 2020. Rome, Italy, Istituto Superiore di Sanita.

Dietz K (1993). The estimation of the basic reproduction number for infectious diseases. *Stat Methods Med Res* **2**(1): 23-41.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.28.20036715; this version posted April 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Ge XY, Wang N, et al. (2016). Coexistence of multiple coronaviruses in several bat colonies in an abandoned mineshaft. *Virol Sin* **31**(1): 31-40.

Goh KT, Cutter J, et al. (2006). Epidemiology and control of SARS in Singapore. *Ann Acad Med Singapore* **35**(5): 301-16.

Haddad N, Robert A, et al. (2019). HIV in Canada-Surveillance Report, 2018. *Can Commun Dis Rep* **45**(12): 304-12.

Jernigan DB (2020). Update: Public Health Response to the Coronavirus Disease 2019 Outbreak - United States, February 24, 2020. *MMWR Morb Mortal Wkly Rep* **69**(8): 216-9.

Kermack WO, McKendrick AG (1991). Contributions to the mathematical theory of epidemics--I. 1927. *Bull Math Biol* **53**(1-2): 33-55.

Kupferschmidt K (2020). Genome analyses help track coronavirus' moves. *Science* **367**(6483): 1176-7.

Lauer SA, Grantz KH, et al. (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann Intern Med*.

Leung GM, Hedley AJ, et al. (2004). The epidemiology of severe acute respiratory syndrome in the 2003 Hong Kong epidemic: an analysis of all 1755 patients. *Ann Intern Med* **141**(9): 662-73.

Lu R, Zhao X, et al. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**(10224): 565-74.

Malik YS, Sircar S, et al. (2020). Emerging novel coronavirus (2019-nCoV)-current scenario, evolutionary perspective based on genome analysis and recent developments. *Vet Q* **40**(1): 68-76.

Olofsson S, Brittain-Long R, et al. (2011). PCR for detection of respiratory viruses: seasonal variations of virus infections. *Expert Rev Anti Infect Ther* **9**(8): 615-26.

Ren SY, Gao RD, et al. (2020). Fear can be more harmful than the severe acute respiratory syndrome coronavirus 2 in controlling the corona virus disease 2019 epidemic. *World J Clin Cases* **8**(4): 652-7.

Rogers K, Jakes L, et al. (2020). Trump Calls It the 'Chinese Virus.' Critics Say That's Racist and Provocative. *New York Times*. New York, A.G. Sulzberger**:** 11.

Rolfes MA, Foppa IM, et al. (2018). Annual estimates of the burden of seasonal influenza in the United States: A tool for strengthening influenza surveillance and preparedness. *Influenza Other Respir Viruses* **12**(1): 132-7.

Seydel J, Kramer A, et al. (1994). Backcalculation of the number infected with human immunodeficiency virus in Germany. *J Acquir Immune Defic Syndr* **7**(1): 74-8.

Shi Y (2003). Stochastic dynamic model of SARS spreading. *Chinese Science Bulletin* **48**: 1287-92.

Small M, Shi P, et al. (2003). Plausible models for propagation of the SARS virus. *Trans Fundam Electron, Commun Comput Sci* **E87-A**.

Svoboda T, Henry B, et al. (2004). Public health measures to control the spread of the severe acute respiratory syndrome during the outbreak in Toronto. *N Engl J Med* **350**(23): 2352-61.

Wittkowski KM (1992). *Estimates of HIV prevalence in Germany and the U.S. – Sensitivity to assumptions concerning reporting delays and incubation time*. VIII International Conference on AIDS / III STD World Congress, Amsterdam, NL, Congrex Holland.

Wittkowski KM (1995a). Entwicklung der HIV-Epidemie bei Hämophilen anhand der Zahlen des Robert-Koch-Instituts (Berlin). *25 Hämophilie-Symposion, Hamburg 1994*. I Scharrer and W Schramm. Berlin, D, Springer**:** 20-8.

Wittkowski KM (1995b). [Epidemiology of HIV in Germany]. *Gesundheitswesen* **57**(5): 291-8.

Wittkowski KM (1996). [Epidemiology of HIV in Germany (status: 30 June 1995)--transmission among heterosexuals and value of HIV screening programs]. *Gesundheitswesen* **58**(4): 229-33.

Woolhouse M, Gaunt E (2007). Ecological origins of novel human pathogens. *Crit Rev Microbiol* **33**(4): 231-42.

Yeh SH, Wang HY, et al. (2004). Characterization of severe acute respiratory syndrome coronavirus genomes in Taiwan: molecular epidemiology and genome evolution. *Proc Natl Acad Sci U S A* **101**(8): 2542-7.

Zhou G, Yan G (2003). Severe acute respiratory syndrome epidemic in Asia. *Emerg Infect Dis* **9**(12): 1608-10.