# Forecasting Covid-19 mortality with age stratified case data

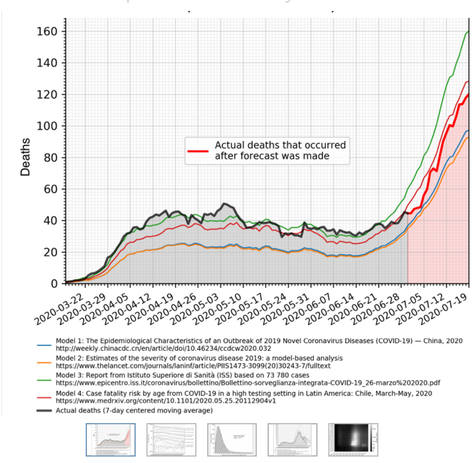
by Marc Bevand

Oct 15, 2020

covid-19-mortality-forecasting

input data: *age, onset of symptoms, age-stratified CFR*
outputs: *future deaths, past deaths*
forecast window: *25 days*
model type: *arithmetic*

a




A simple improvement to forecasting Covid-19 mortality was developed by using age-stratified data [Source Code 1] to address the large variance in case fatality ratio (CFR) across age groups [1].

In Panel A, mortality forecasts using this method showed similar performance in predicting observed deaths compared to representative alternative models for the same region [2-5]. Forecast is based on the age of every single Covid-19 case reported by the Florida Department of Health combined with 4 models of age-stratified CFRs. The forecasting method involves calculating age stratified CFRs, fitting a gamma distribution for time lag between onset of symptoms and death, and finally calculating the proportion and timing of mortality for current cases.

In Panel B, age-stratified CFRs were calculated from existing death data. Moving averages with 35 day (solid lines) and 14 day (dashed lines) windows were applied as well as a correction for right-censoring (missing death records in recent data). Dotted lines represent 14 day moving average without correction for right-censoring.

In Panel C, a gamma distribution was fitted to data pooled across all ages and had a mean onset symptom to death of 25.1 days. Slight differences in age-specific gamma distributions were noticeable, however age stratified gamma distributions did not substantially alter forecasts and thus a single gamma distribution pooled across all age groups was used.

Third, future mortality was then forecasted by calculating the proportion and timing of mortality in current cases in each age group. A more recent forecast for mortality in the state of Florida through October 2020 is provided in Panel D.

A detailed example using the Florida Department of Health (FDOH) open Covid-19 data set [ Dataset 1] includes additional details and explanations [Protocols 1]. Additional scripts and data visualizations that may be helpful to find and understand relevant patterns in the data are also included [Source Code 1]. For example, a script to visualize case counts across time and age brackets as heat maps is shown in Panel E, where each pixel represents 7-day time period and 5 year age bracket.

The FDOH resource [Dataset 1] is to our knowledge the only open dataset documenting the exact age of hundreds of thousands of cases. Health departments and other maintainers of Covid-19 datasets should prioritize inclusion of age parameter and age-stratified CFRs to improve the accuracy of morbidity and mortality forecasts and analyses.

A limitation of our technique specific to the FDOH dataset is that date of death and onset of symptoms (used for gamma distribution model) is not always accurate. We know this is an important issue because another dataset published by FDOH—deaths by day—allows us to infer there is often a delay of multiple weeks between the moment a death occurs, and the moment the death is reflected in the line list. However we cannot identify individual deaths or symptom onset dates to correct for this delay. This limitation contributes to longer tails in the onset-to-death distribution.

Overall, our method offers a fast and simple way to predict Covid-19 mortality using age-stratified case data.

References:

1) https://doi.org/10.3947/ic.2020.52.2.154
2) https://doi.org/10.1016/S1473-3099(20)30243-7
3) https://doi.org/10.1101/2020.05.25.20112904
4) https://doi.org/10.46234/ccdcw2020.032
5) https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_26-marzo%202020.pdf

Protocols:

1) Detailed description and representative outputs of Covid-19 mortality forecasts

Code:

1) Code and Example using Florida Dataset

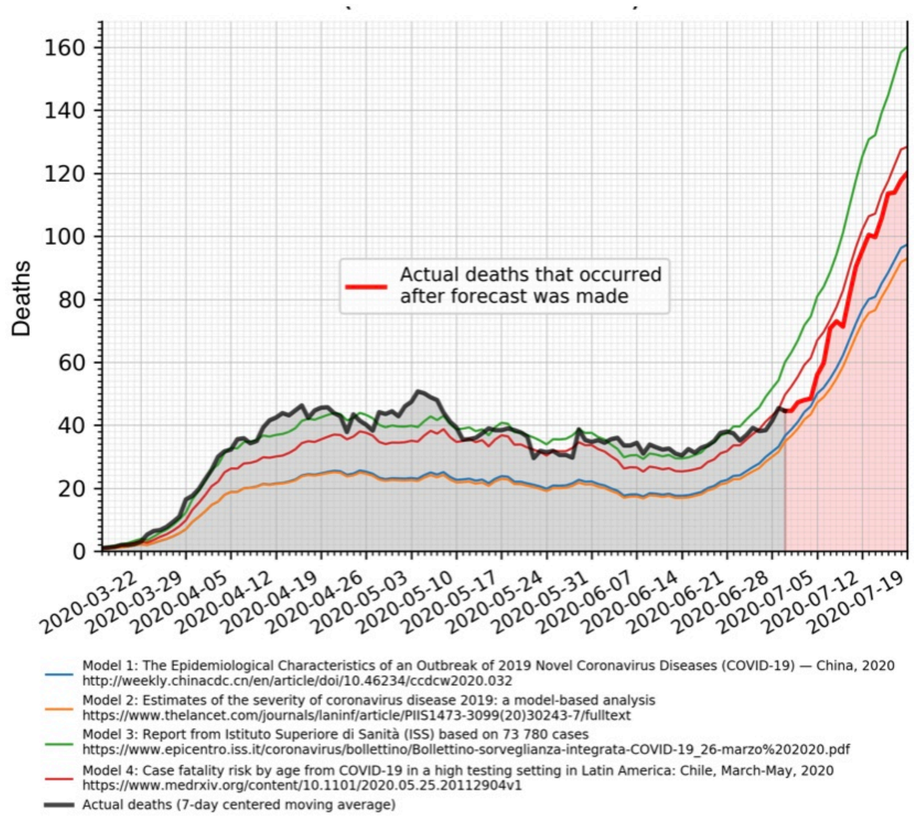
Datasets:

1) Florida COVID-19 cases and deaths

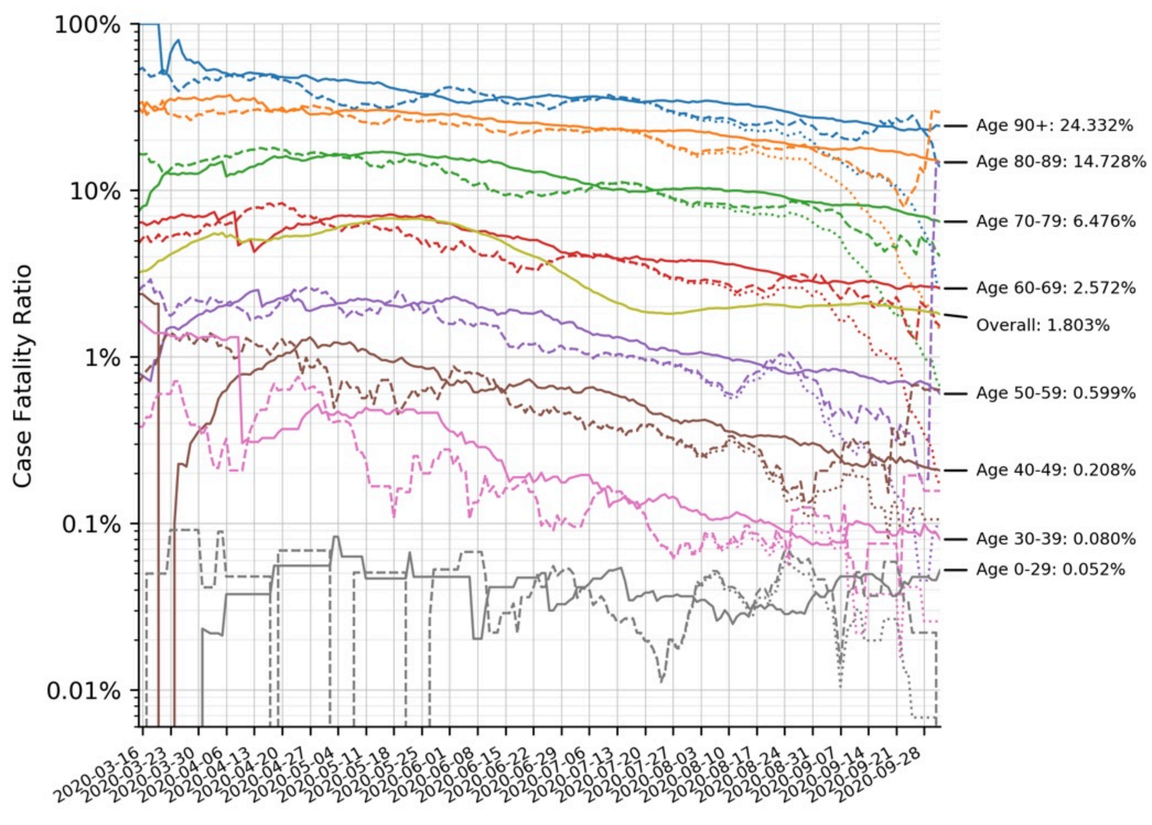


© 2020 FLASHPUB Terms & Conditions All content CC - BY

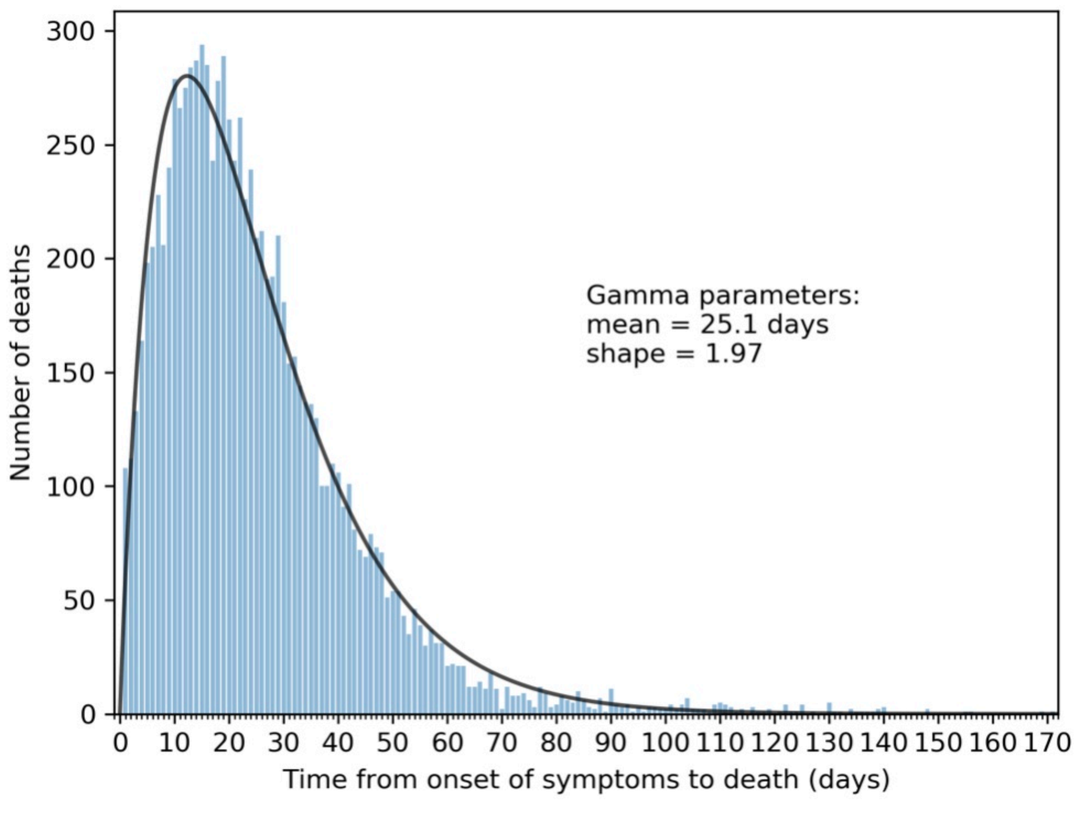
a


Example forecast accuracy for Florida
Deaths
Actual deaths that occurred after forecast was made
Model 1: The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) — China, 2020
http://weekly.chinacdc.cn/en/article/doi/10.46234/ccdcw2020.032
Model 2: Estimates of the severity of coronavirus disease 2019: a model-based analysis
https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30243-7/fulltext
Model 3: Report from Istituto Superiore di Sanità (ISS) based on 73 780 cases
https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_26-marzo%202020.pdf
Model 4: Case fatality risk by age from COVID-19 in a high testing setting in Latin America: Chile, March-May, 2020
https://www.medrxiv.org/content/10.1101/2020.05.25.20112904v1
Actual deaths (7-day centered moving average)

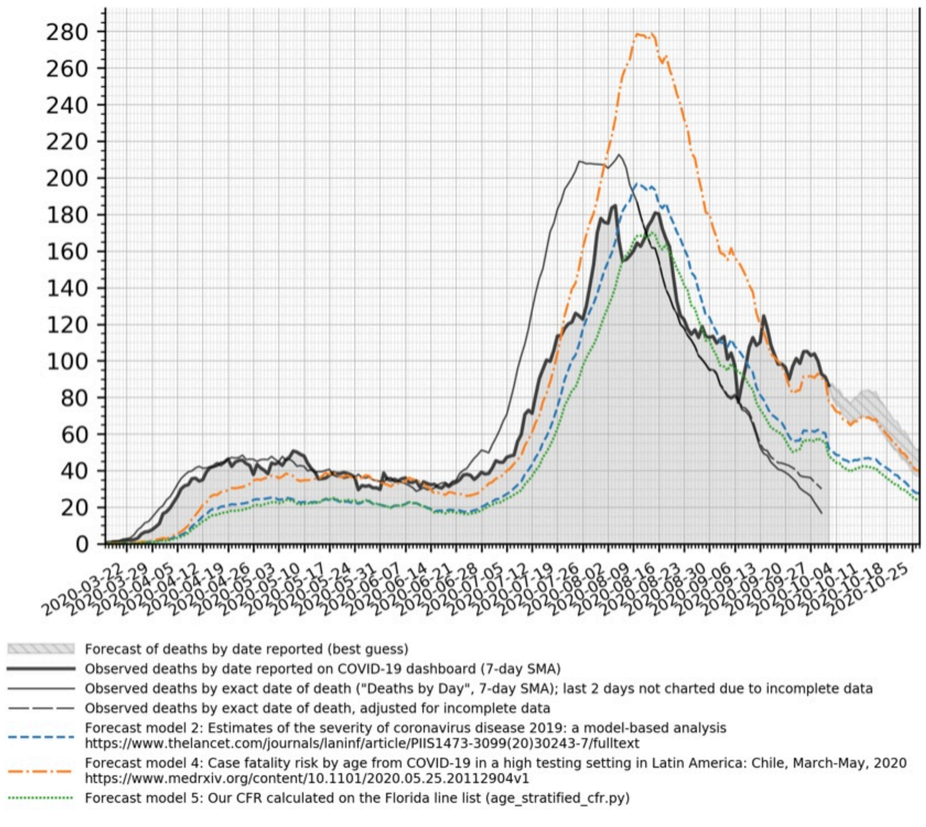
b


Age stratified CFRs for Florida
Case Fatality Ratio
100%
10%
1%
0.1%
0.01%
2020-03-16
2020-03-23
2020-03-30
2020-04-06
2020-04-13
2020-04-20
2020-04-27
2020-05-04
2020-05-11
2020-05-18
2020-05-25
2020-06-01
2020-06-08
2020-06-15
2020-06-22
2020-06-29
2020-07-06
2020-07-13
2020-07-20
2020-07-27
2020-08-03
2020-08-10
2020-08-17
2020-08-24
2020-08-31
2020-09-07
2020-09-14
2020-09-21
2020-09-28
Age 90+: 24.332%
Age 80-89: 14.728%
Age 70-79: 6.476%
Age 60-69: 2.572%
Overall: 1.803%
Age 50-59: 0.599%
Age 40-49: 0.208%
Age 30-39: 0.080%
Age 0-29: 0.052%

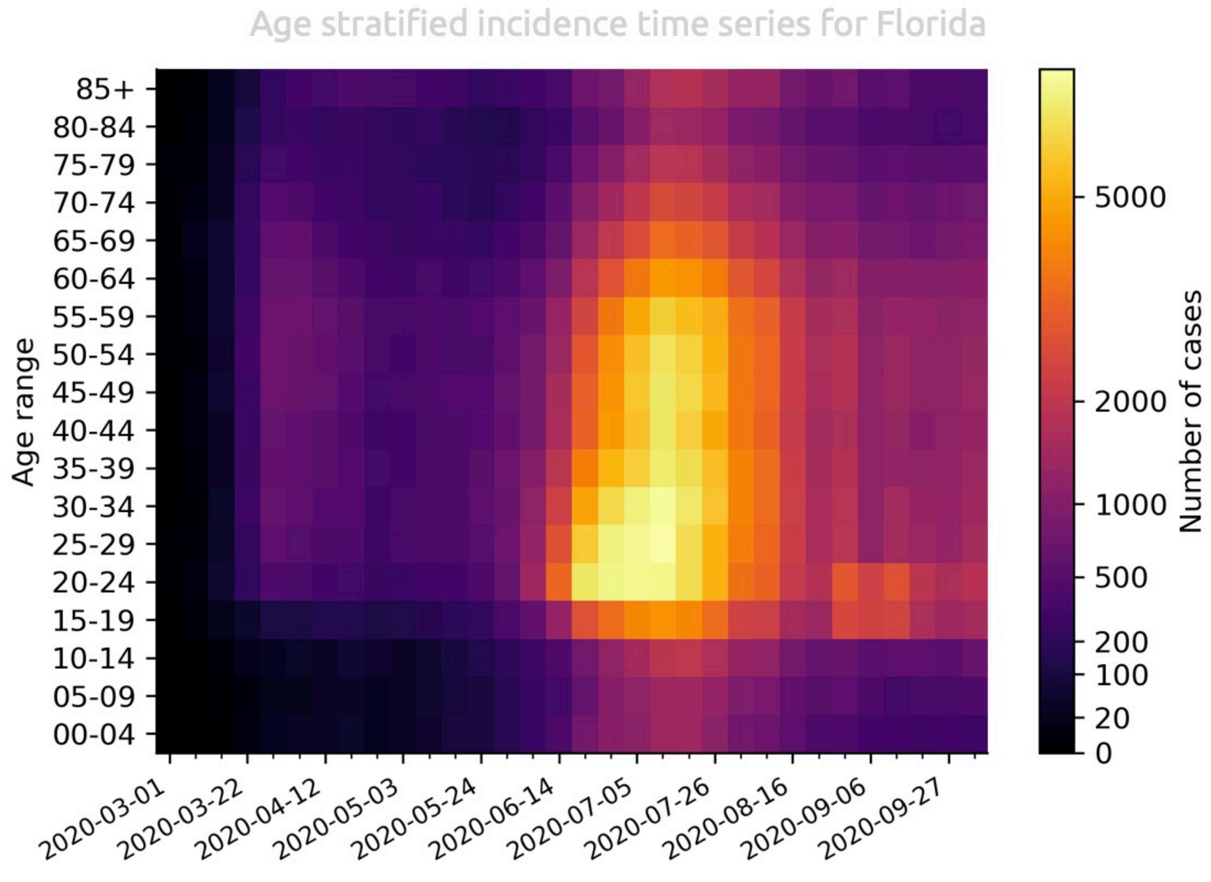
C


Symptom onset to death gamma distribution for Florida
Number of deaths
Time from onset of symptoms to death (days)
Gamma parameters:
mean = 25.1 days
shape = 1.97
0
50
100
150
200
250
300
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170

d


Florida mortality forecast for October 2020
280
260
240
220
200
180
160
140
120
100
80
60
40
20
0
2020-03-22
2020-03-29
2020-04-05
2020-04-12
2020-04-19
2020-04-26
2020-05-03
2020-05-10
2020-05-17
2020-05-24
2020-05-31
2020-06-07
2020-06-14
2020-06-21
2020-06-28
2020-07-05
2020-07-12
2020-07-19
2020-07-26
2020-08-02
2020-08-09
2020-08-16
2020-08-23
2020-08-30
2020-09-06
2020-09-13
2020-09-20
2020-09-27
2020-10-04
2020-10-11
2020-10-18
2020-10-25
Forecast of deaths by date reported (best guess)
Observed deaths by date reported on COVID-19 dashboard (7-day SMA)
Observed deaths by exact date of death ("Deaths by Day", 7-day SMA); last 2 days not charted due to incomplete data
Observed deaths by exact date of death, adjusted for incomplete data
Forecast model 2: Estimates of the severity of coronavirus disease 2019: a model-based analysis
https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30243-7/fulltext
Forecast model 4: Case fatality risk by age from COVID-19 in a high testing setting in Latin America: Chile, March-May, 2020
https://www.medrxiv.org/content/10.1101/2020.05.25.20112904v1
Forecast model 5: Our CFR calculated on the Florida line list (age_stratified_cfr.py)

e


Age stratified incidence time series for Florida
Age range
85+
80-84
75-79
70-74
65-69
60-64
55-59
50-54
45-49
40-44
35-39
30-34
25-29
20-24
15-19
10-14
05-09
00-04
2020-03-01
2020-03-22
2020-04-12
2020-05-03
2020-05-24
2020-06-14
2020-07-05
2020-07-26
2020-08-16
2020-09-06
2020-09-27
Number of cases
5000
2000
1000
500
200
100
20
0