

## Using mathematical models to assess the number of total deaths and deaths per 10000 people from the COVID-19 epidemic

\*Qingqing Sun<sup>1</sup>, Chunyi Wang<sup>2</sup>

1. Tianjin University, 2. Mathematics Faculty, University of Waterloo

Death tolls and death tolls per 10000 people of 161 countries and regions as 13 June, 2020 were predicted by "random forest" (<https://github.com/qingqingsun-code/covid-19-prediction>). The reference indexes are (1) past and present flight routes, (2) change rate of population, world share, population density, net change population density, migrant density, land area, (3) fertility rate, density, median age, (4) urban population proportion, urban areas, urban population density, (5) 65+ older population density, proportion of 65+ older population, (6) numbers of hospital beds per 1,000 people, intensive care unit (ICU) beds per 10000 people, density of medical doctors (per 10000 population), density of nursing and midwifery personnel (per 10000 population), density of pharmacists (per 10000 population), (7) gross domestic product (GDP) Nominal (\$) per people, GDP per chasing power parity (PPP) (Int. \$) per people, (8) out of pocket expenditure as percentage of current health expenditure % (US\$), direct financial plan on fighting COVID-19, current health expenditure (CHE) as percentage of gross domestic product (GDP)(%), CHE per capital (US\$), domestic general government health expenditure (GGHED) as percentage of general government expenditure (GGE)(%), (9) average of international health regulation score capacity scores, universal health coverage (UHC) service coverage index. Based on the heatmap of correlations, the highest indexes were annual rate of population change, urban area, and GGHED as percentage of GGE. Predicted death tolls (441,692) were 4.64% high than the statistical death toll (422,090). Predicted mean death toll per 10,000 was 1.14, higher than the statistical one 0.78. Because patients with basic diseases are less resistant to COVID-19, there is a possible statistical bias in the number of deaths in some country and region. It is also not ruled out that some of the dead patients have not been diagnosed. Quarantine order is the main reason for the difference between simulated data and statistical data. As a global community, we need solidarity, disclosure information, cooperation in the development of vaccines, sound medical relief supplies to limit the spread of COVID-19.

Keywords: COVID-19, Random forest, Death toll



