Analyzing Trends of COVID Outcomes in Five NYC Boroughs by Fitting a Richard's Growth Model Using Newton-Raphson Algorithm

Zhengwei Song

April 2023

Contents

1 Background

The COVID-19 pandemic has affected many aspects of our lives and the distribution of COVID-19 vaccines is a critical step to control the pandemic. However, at the beginning of the COVID-19 outbreak, little was known about its mode of transmission, the risk to different populations, etc. Therefore, tracking and studying the COVID-19 epidemiological data at all levels empowered COVID-19 research, and further helped policymakers make timely and informed decisions to identify the priority groups for vaccination after the COVID-19 vaccines were authorized.

This study has 2 objectives. One is to fit pandemic growth curves for each NYC borough from February 29, 2020 to December 11, 2020, and the other is to make a distribution plan for COVID-19 vaccines across NYC based on the predicted trends of all boroughs.

The open-source public dataset was used in this study: the citywide and borough-specific daily counts of COVID cases, hospitalizations, and deaths from February 29, 2020 to December 11, 2020 published by the New York Department of Health (NYDOH).

2 Richard Growth Function

Richard's growth function is a four-parameter nonlinear S-shaped function extended from the logistic model, since it does not enforce symmetry before and after the inflection point. It has been commonly applied to predict the spread of various infectious diseases in a population over time, such as influenza, dengue fever, and COVID-19. Let $N(t)$ be the number of a population at time t, Richard's growth function is defined by

$$N(t)=\frac{a}{\{1+d\exp\{-k\left(t-t_0\right)\}\}^{1/d}}$$

where t is the time since the beginning of a population,

$$\boldsymbol{\theta}=(a,k,d,t_0)$$

are parameters with specific geometric meanings. The parameter a is the upper bound of the function, i.e., $a=\lim_{t\to\infty}N(t)$, the largest population it could reach); The parameter k is the growth rate, which controls the slope at an inflection; The parameter t_0 is the time at an inflection, where the curve changes from convex to concave; and finally, the parameter d is another shape parameter.

The Richard growth model used in this study took the form of

$$Y_i=N\left(t_i,\boldsymbol{\theta}\right)+\epsilon_i$$

where $Y_i=\sum_{k=1}^{i}y_k$ was the cumulative number of cases by time t_i, $\boldsymbol{\theta}=(a,k,d,t_0)$ was the parameters in the Richard's function, and ϵ_i was the random error with mean zero.

3 A Modified Newton-Raphson Algorithm for Curve Fitting

I implemented a modified Newton-Raphson Algorithm to fit the Richard's growth curve to the data, which consists of daily COVID counts of diagnosed cases, hospitalizations, and deaths for all 5 NYC boroughs. The objective is to maximize $-RSS=f(\theta)=\sum_{Ii=1}^{t_i}(Y_i-N(t_i,\theta))^2$. RSS stands for Residual Sum of Squares. I compute the gradient of the objective function

$$\nabla f(\boldsymbol{\theta})=\left(\frac{\partial f(\boldsymbol{\theta})}{\partial a},\frac{\partial f(\boldsymbol{\theta})}{\partial k},\frac{\partial f(\boldsymbol{\theta})}{\partial d},\frac{\partial f(\boldsymbol{\theta})}{\partial t_0}\right)'$$

and the Hessian matrix of the objective function $\nabla^2 f(\boldsymbol{\theta})$ with i,j element equal to $\frac{\partial^2 f(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}$, using the pracma package in R.

The Newton-Raphson algorithm finds a least-squares estimator of $\boldsymbol{\theta}=(a,k,d,t_0)$ by solving the system of equations $\nabla f(\boldsymbol{\theta})=\mathbf{0}$ iteratively. Specifically, by approximating the gradient $\nabla f(\boldsymbol{\theta})$ with its first-order Taylor expansion, i.e., in the i^{th} iterative step, $\nabla f(\boldsymbol{\theta}_i) \approx \nabla f(\boldsymbol{\theta}_{i-1}) + \nabla^2 f(\boldsymbol{\theta}_{i-1})(\boldsymbol{\theta}_i - \boldsymbol{\theta}_{i-1})$ given that $\boldsymbol{\theta}_{i-1}$ is sufficiently close to $\boldsymbol{\theta}_i$. Setting this approximation form to 0 and solving for $\boldsymbol{\theta}_i$, I obtain the iterative step $\boldsymbol{\theta_i} = \boldsymbol{\theta_{i-1}} - (\nabla^2 f(\boldsymbol{\theta_{i-1}}))^{-1} \nabla f(\boldsymbol{\theta_{i-1}})$. In this form, the step size λ is 1. However, λ can be adjusted to achieve reduction of the RSS in each iteration. Detailed steps of the Newton-Raphson algorithm implemented for fitting the Richard's growth curve to the COVID data are as follows:

Initialization step (i = 0). At this step, the set of starting values $\boldsymbol{\theta}_0$ were used to compute the initial gradient and Hessian. The starting step size λ was predetermined as 1. The starting RSS, RSS_0, was set to ∞. The tolerance value ϵ was set as 1^{-10}.

Iterative process (i = 1, 2, 3, ...). While the $|RSS_i - RSS_{i-1}| > \epsilon$:

1. Direction check: First, I checked whether $-(\nabla^2 f(\boldsymbol{\theta}_{i-1}))^{-1} \nabla f(\boldsymbol{\theta}_{i-1}) > 0$, i.e., whether the current step is in ascent direction. Note that the ascent direction, instead of the descent direction, is intended because the objective is to maximize -RSS. If the current step is not in ascent direction, I force an ascent direction by replacing the Hessian matrix H_{i-1} by $H_{i-1} - \gamma I$, where γ is a number larger than H_{i-1}'s largest eigenvalue. The replacement matrix is negative definite because all its eigenvalues are negative, which guarantees an ascent direction.

2. Bound check: Next, I update the parameter values

$$\boldsymbol{\theta_i} = \boldsymbol{\theta_{i-1}} - \lambda(\nabla^2 f(\boldsymbol{\theta_{i-1}}))^{-1} \nabla f(\boldsymbol{\theta_{i-1}})$$

   where $\lambda = 1$. I then check separately if each element of the updated $\boldsymbol{\theta}_i = (a, d, t, t_0) > 0$. If not, the parameter value was not updated and was set to its value in the previous step. The bound checking step ensures that the updated parameters make sense in the context of pandemic modeling.

3. Step-halving: After obtaining updated parameters $\boldsymbol{\theta}_i$ within interpretable bounds, I compute the RSS corresponding to the updated parameters. While $RSS_i \geq RSS_{i-1}$, I reduce the default step size $\lambda = 1$ by half, re-update the parameters, re-perform the bound check, and re-calculate the RSS. Step-halving is repeated until $RSS_i < RSS_{i-1}$.

The iterative process terminates with $|RSS_i - RSS_{i-1}| \leq \epsilon$.

4 Choosing Starting Values

The Newton-Raphson algorithm is sensitive to starting values. Therefore, it is crucial to choose starting values sufficiently close to the minimizer of the RSS function for the Newton-Raphson algorithm to achieve convergence. In the project, the starting values were chosen empirically.

First, by inspecting the plot of the daily counts and the cumulative counts for each COVID outcome over time for each borough, I observed two pandemic waves for all outcome-borough combinations. Therefore, I decided to divide the data into two pandemic waves and choose starting values for each wave separately. Specifically, I defined day 1 to day 150 as the first wave and day 151 to the last day in the data set, day 287, as the second wave. The aforementioned plots for each borough are as follows:

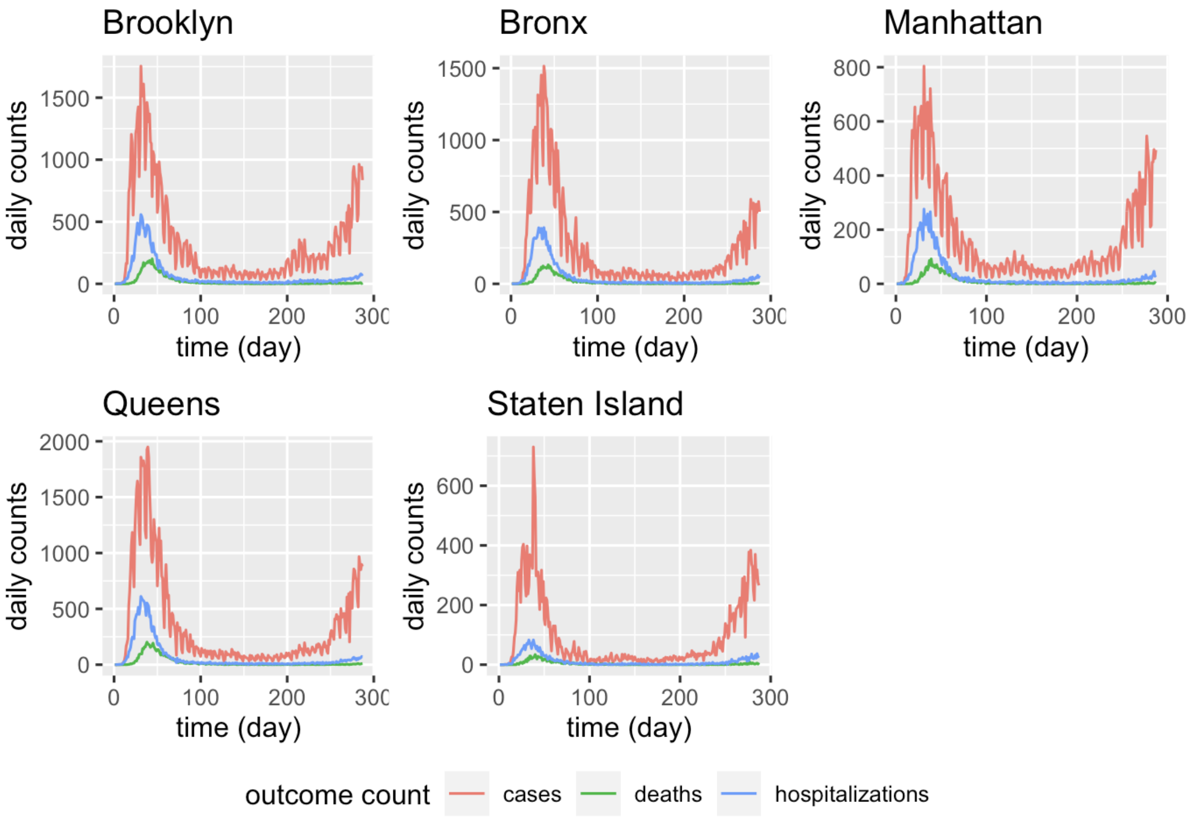


Figure 1: Daily Counts by Outcome and Borough

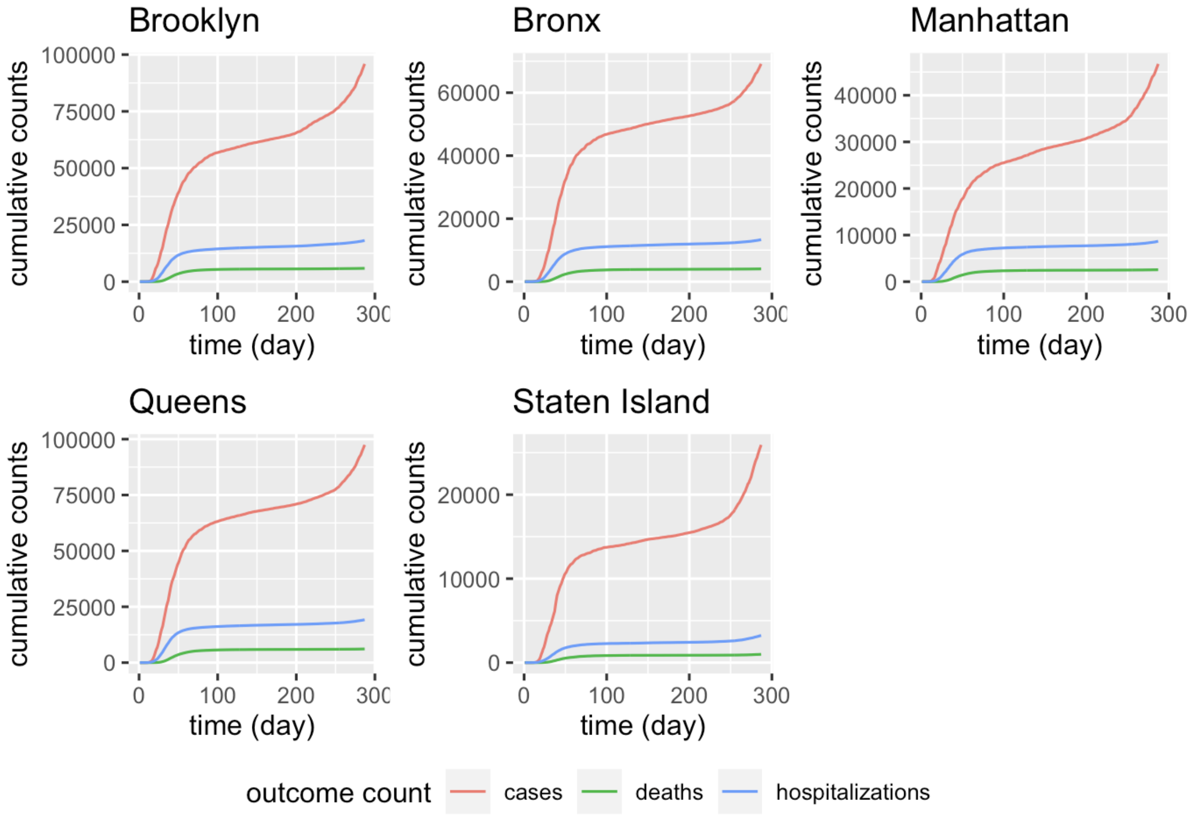


Figure 2: Cumulative Counts by Outcome and Borough

The starting value of a was chosen based on the maximum cumulative counts for each wave. For wave 1, the starting value of a was chosen as a number slightly larger (20%) than the cumulative counts on the last day of wave 1. I chose the starting value of a in wave 2 to be slightly larger (20%) than 2 times the cumulative case on the last recorded day of wave 2 (December 11th, 2020). This was based on the assumption that the daily counts for all COVID outcomes (diagnosis, hospitalizations, and deaths) in wave 2 peaked on December 11th, 2020 and would start decreasing after that date. I believe that this is a reasonable assumption because the COVID vaccine was approved for use on December 11th, 2020, and vaccination is expected to reduce the daily diagnosis, hospitalization, and death counts.

d is a positive number. Specifically, d = 0 corresponds to the Gompertz growth curve

[1], and d = 1 corresponds to the logistic growth curve [2]. I chose d = 0.8 as the starting value for all borough-outcome combinations.

I choose t_0 as the the day where the number of daily counts of the COVID outcome reaches the highest. That is, t_0 corresponds to the peak in daily counts, or the inflection point in cumulative counts.

After determining the starting values of a, d, and t_0, I obtained the starting value of k by linear regression because of the following linear relationship:

$$
\begin{aligned}
m = N(t) &= \frac{a}{(1+de^{-k(t-t_0)})^{\frac{1}{d}}} \\
\Rightarrow \frac{a}{m} &= (1+de^{-k(t-t_0)})^{\frac{1}{d}} \\
\Rightarrow \log\frac{a}{m} &= \frac{1}{d}\log(1+de^{-k(t-t_0)}) \\
\Rightarrow \log(\frac{a}{m})^d &= \log(1+de^{-k(t-t_0)}) \\
\Rightarrow (\frac{a}{m})^d &= 1+de^{-k(t-t_0)} \\
\Rightarrow (\frac{a}{m})^d - 1 &= de^{-k(t-t_0)} \\
\Rightarrow \log((\frac{a}{m})^d - 1) &= \log(de^{-k(t-t_0)}) = \log d - k(t-t_0) \\
\Rightarrow \log(\frac{(\frac{a}{m})^d - 1}{d}) &= -k(t-t_0)
\end{aligned}
$$

Therefore, by plugging in the empirically selected a, d, and t_0, plugging the observed COVID count outcomes in m, and plugging the day index in t, the starting value of k can be obtained from the linear regression model with dependent variable $\log(\frac{(\frac{a}{m})^d-1}{d})$ and independent variable $-(t-t_0)$.

The starting values for all borough-outcome combinations in wave 1 and wave 2 are in the following table, respectively:

	Cumulative Cases (K)					Cumulative Hospitalizations (K)					Cumulative Deaths (K)				
	BK	BX	MN	QN	SI	BK	BX	MN	QN	SI	BK	BX	MN	QN	SI
a	65	70	30	70	15	17	15	10	20	5	6	5	4	7	2
k	0.0432	0.0483	0.0450	0.0483	0.0444	0.0407	0.0364	0.0306	0.0349	0.0249	0.0414	0.0317	0.0271	0.0395	0.0192
d	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
t_0	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40

Table 1: Starting Values for Wave 1 (Outcome Counts in Thousand)

	Cumulative Cases (K)					Cumulative Hospitalizations (K)					Cumulative Deaths (K)				
	BK	BX	MN	QN	SI	BK	BX	MN	QN	SI	BK	BX	MN	QN	SI
a	80	42	42	65	30	10	5	4	5	2	1.2	0.5	0.4	0.6	0.3
k	0.0266	0.0235	0.0253	0.0263	0.0292	0.0221	0.0205	0.0216	0.0242	0.0247	0.0232	0.0180	0.0177	0.0199	0.0271
d	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
t_0	137	137	137	137	137	137	137	137	137	137	137	137	137	137	137

Table 2: Starting Values for Wave 2 (Outcome Counts in Thousand)

5 Result

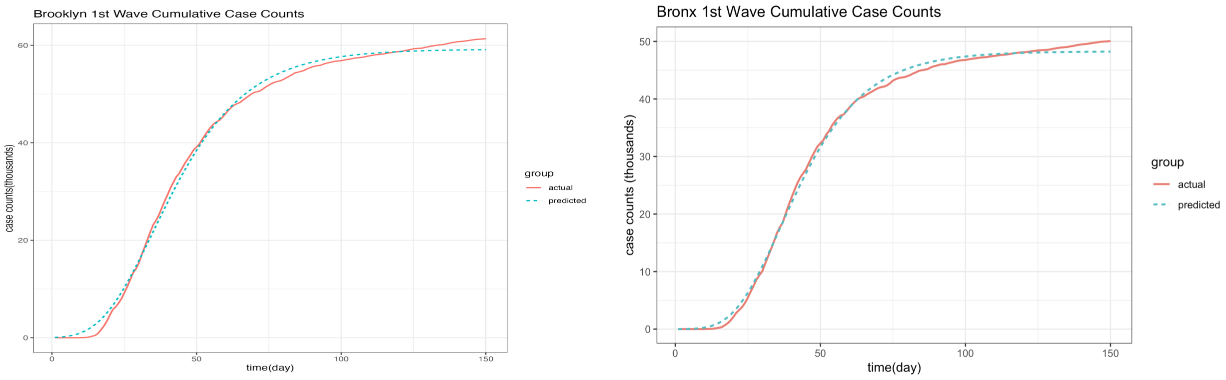
By applying the Newton-Raphson algorithm, the estimated four parameters are obtained (from the given starting values) and then to be fitted in Richard's function, where the starting values are chosen based on the selection rule (Section 4). Also, among the four parameters, parameter k, which controls the growth rate, is to be used as an indicator[1] to assess and compare the severity of COVID-19 of case counts, hospitalization counts and death counts.

	Cumulative Cases (K)					Cumulative Hospitalizations (K)					Cumulative Deaths (K)				
	BK	BX	MN	QN	SI	BK	BX	MN	QN	SI	BK	BX	MN	QN	SI
$\hat{a}$	59.1925	65.3846	27.0454	65.3846	13.9874	15.7101	11.882	7.6964	16.3018	2.4750	5.6007	3.911	2.4564	5.7782	0.8855
$\hat{k}$	0.0560	0.0617	0.0511	0.0617	0.0723	0.0618	0.0665	0.0667	0.0882	0.0574	0.0752	0.069	0.0655	0.0767	0.0639
$\hat{d}(\times 10^{-12})$	41.822	1.5317×10^4	1.5387	1.5317×10^4	19.554	15.020	32.01	18.989	21.557	1.4090	12.862	22.32	7.5658	22.050	251.25
$\hat{t_0}$	35.0033	34.8640	33.7299	34.8640	32.9436	33.5572	33.835	32.5439	31.8764	34.5511	40.8587	41.267	41.5707	41.0972	40.8286

Table 3: Estimated Values for Wave 1 (Outcome Counts in Thousand)

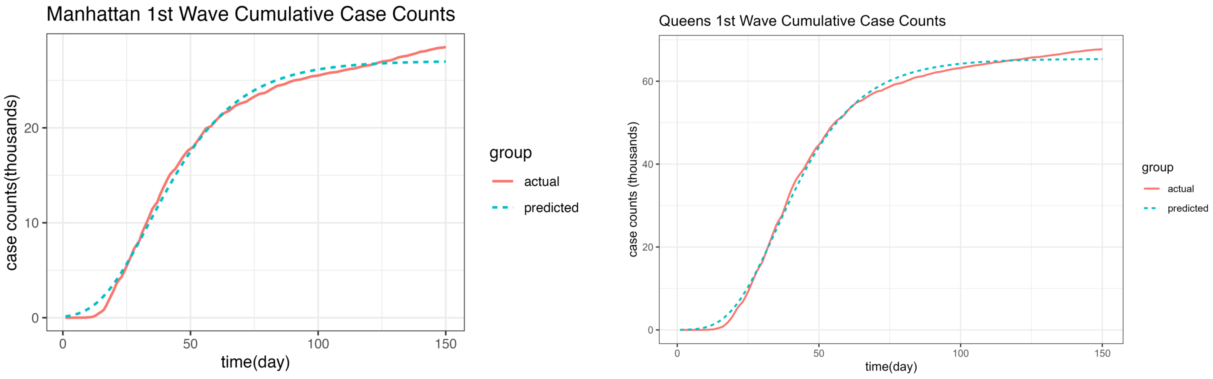
5.1 First Wave

5.1.1 Cases



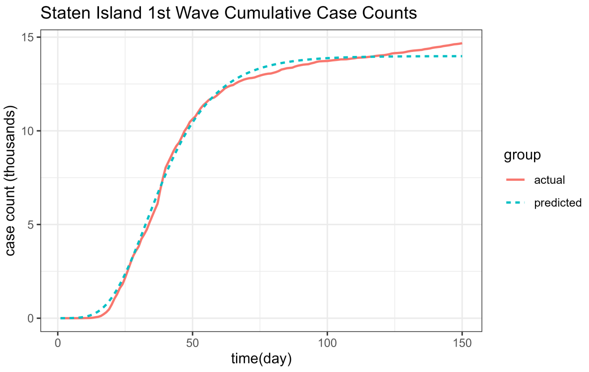
(a) Brooklyn

(b) Bronx



(c) Manhattan

(d) Queens



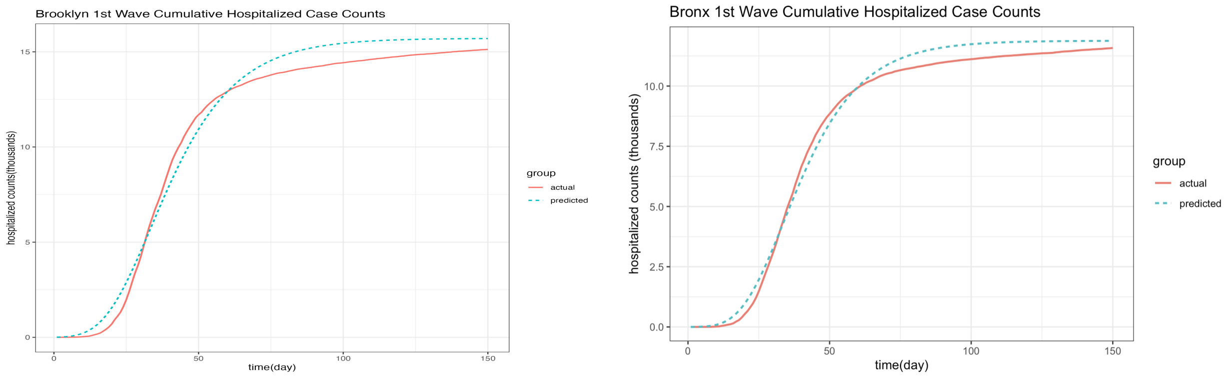
(e) Staten Island

Figure 3: Richard's fitted curves (red) vs. actual data (dashed blue) of Wave 1 Cumulative Cases (in Thousand) by Time (Day) per Borough

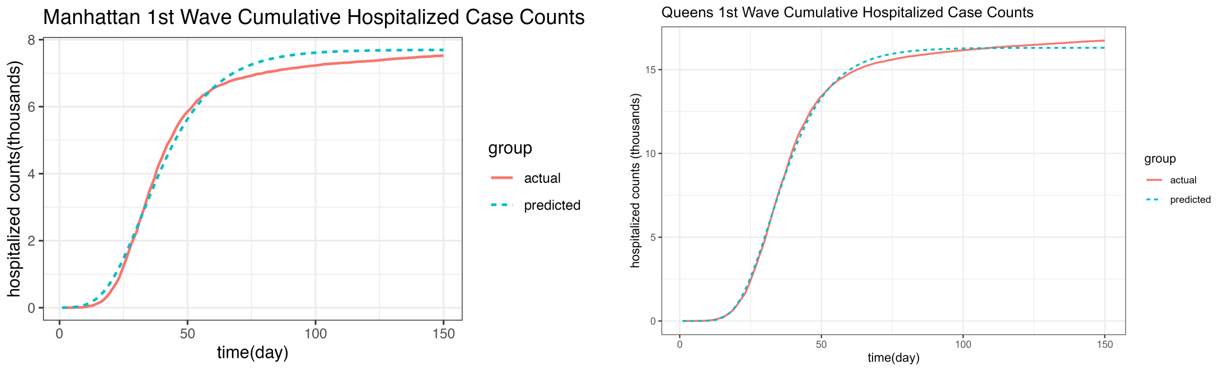
For cumulative case counts shown in Richard's model, in each five of New York City boroughs, the respective four optimized parameters are shown in Table 3. Taking Brooklyn as an example, the starting value used as 65, 0.0432, 0.8 and 40 for a, k, d and t_0, respectively. Then, the estimated parameters are 59.1925, 0.0560, 4.1822×10^{-11} and 35.0033 for $\hat{a}$, $\hat{k}$, $\hat{d}$ and $\hat{t_0}$, respectively. The result indicates that the number of COVID-19 cases would reach approximately 59.193 thousand people in Brooklyn, with a growth rate of 0.056. Also, the inflection time was 35 days, which means the case outbreak would speed down after 35 days since the record starts. Finally, the fitted Richard's model by these parameters is highly consistent with the actual data (Fig. 3).

Also, among the five boroughs, the result $\hat{k}$ (=0.0723) shows Staten Island recorded the largest growth rate, followed by Bronx and Queens (0.0617), and Brooklyn (0.0560), Manhattan (0.0511). Thus, Staten Island is the most severe borough on case counts in the first wave of the pandemic.

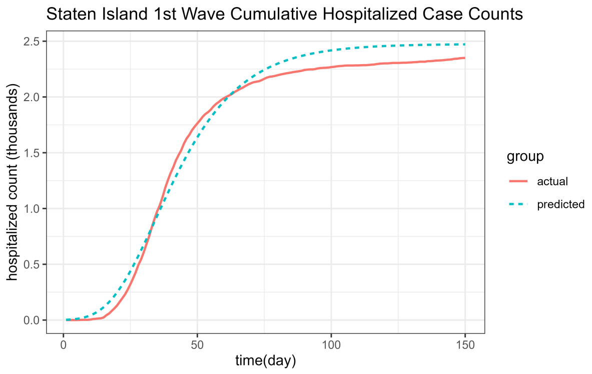
5.1.2 Hospitalizations



(a) Brooklyn (b) Bronx



(c) Manhattan (d) Queens



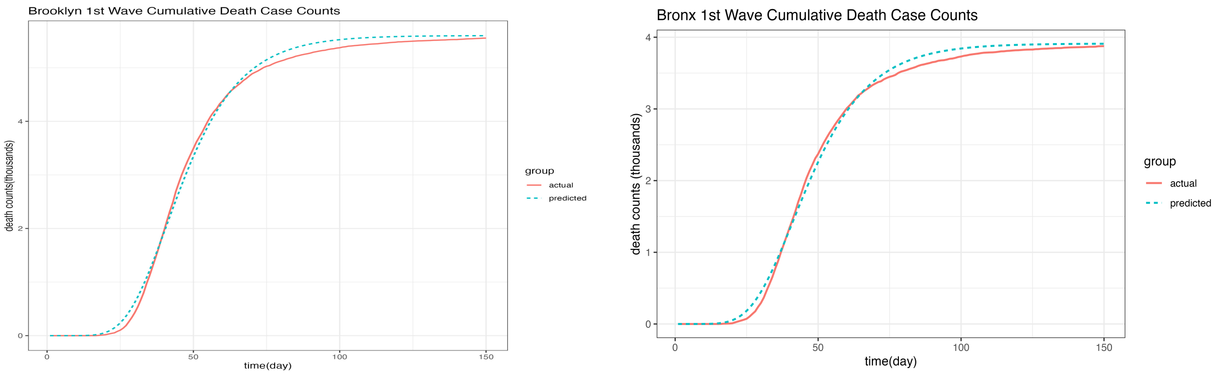
(e) Staten Island

Figure 4: Richard's fitted curves (red) vs. actual data (dashed blue) of Wave 1 Cumulative Hospitalizations (in Thousand) by Time (Day) per Borough

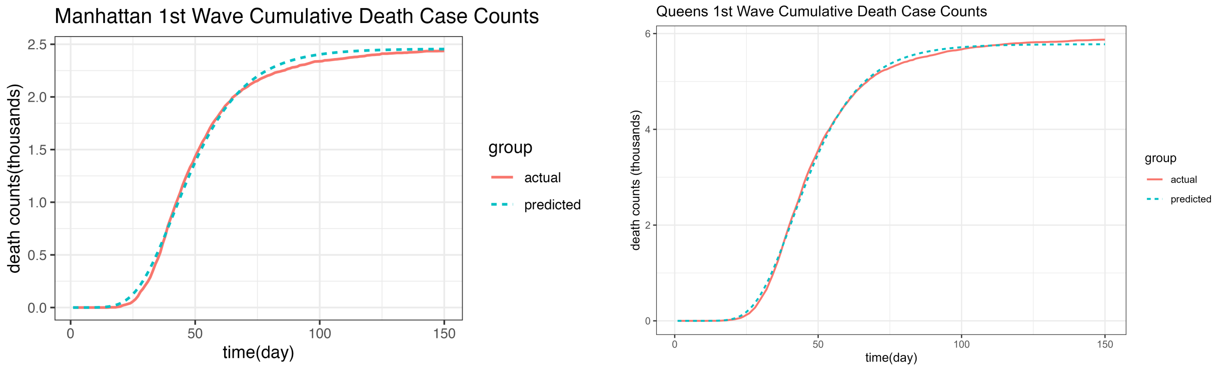
For cumulative hospitalization counts shown in Richard's model, in each five of New York City boroughs, similarly, the four optimized parameters are shown in Table 3. Finally, the fitted Richard's model by these parameters is highly consistent with the actual data (Fig. 4).

Also, among the five boroughs, the result $\hat{k}$ (=0.0882) shows Queens recorded the largest growth rate, followed by Manhattan (0.0667), Bronx (0.0665), Brooklyn (0.0618), and Staten Island (0.0574). Thus, Queens is the most severe borough on hospitalization counts in the first wave of the pandemic.

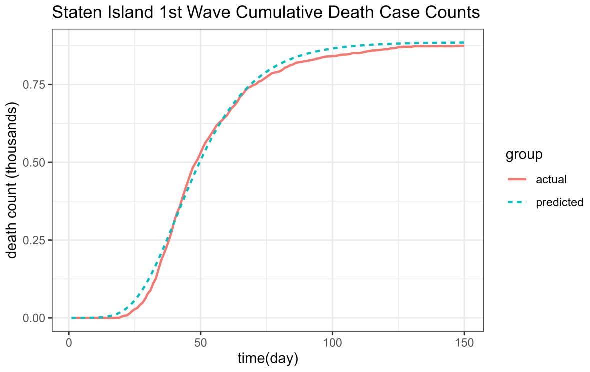
5.1.3 Deaths



(a) Brooklyn (b) Bronx



(c) Manhattan (d) Queens



(e) Staten Island

Figure 5: Richard's fitted curves (red) vs. actual data (dashed blue) of Wave 1 Cumulative Deaths (in Thousand) by Time (Day) per Borough

For cumulative death counts shown in Richard's model, in each five of New York City boroughs, similarly, the four optimized parameters are shown in Table 3. Finally, the fitted Richard's model by these parameters is highly consistent with the actual data (Fig. 5).

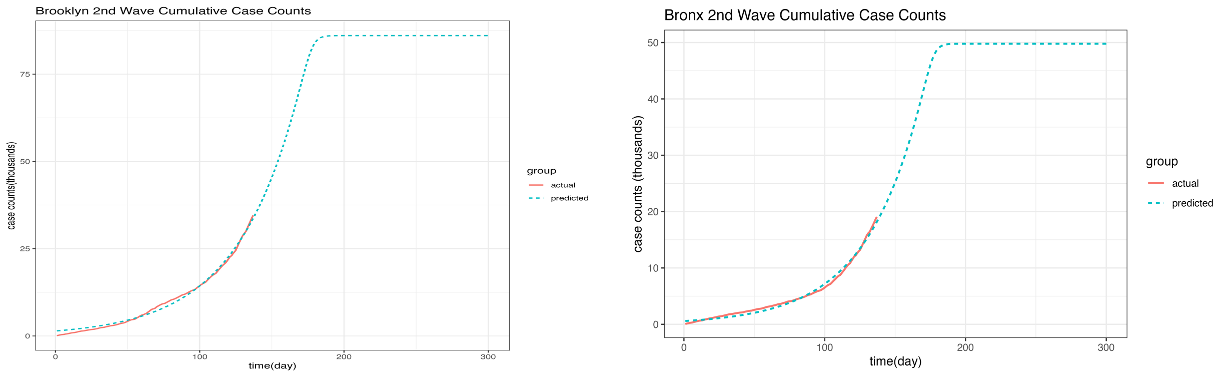
Also, among the five boroughs, the result $\hat{k}$ (=0.0767) shows Queens recorded the largest growth rate, followed by Brooklyn (0.0752), Bronx (0.069), Manhattan (0.0655), and Staten Island (0.0639). Thus, Queens is the most severe borough on death counts in the first wave of the pandemic.

5.2 Second Wave

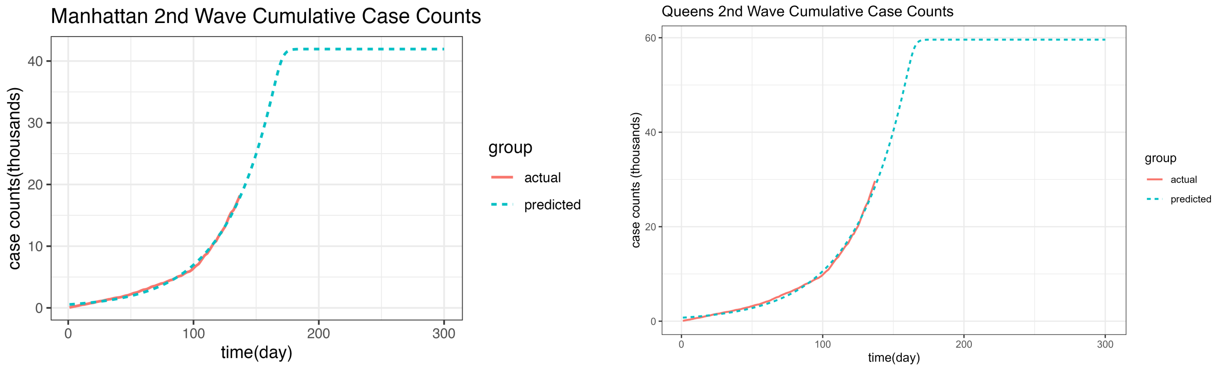
	Cumulative Cases (K)					Cumulative Hospitalizations (K)					Cumulative Deaths (K)				
	BK	BX	MN	QN	SI	BK	BX	MN	QN	SI	BK	BX	MN	QN	SI
$\hat{a}$	85.5067	49.789	41.9403	59.5907	34.2817	9.4513	6.736	2.5723	9.1255	1.9815	2.6286	0.524	0.3741	0.5203	0.3938
$\hat{k}$	0.3379	0.365	0.4047	0.5719	0.381	0.0185	0.166	0.3801	0.1956	0.5729	0.0089	0.0081	0.2022	0.3125	0.4208
$\hat{d}$	14.6456	14.574	15.8552	21.3935	11.0869	0.6664	8.615	17.2515	8.7343	16.915	0.1412	2.695×10^{-9}	11.1460	17.1367	9.7594
$\hat{t_0}$	169.4569	169.813	163.5019	159.2813	162.7768	169.543	198.097	169.7831	186.8791	156.6111	238.2228	154.253	189.5283	173.8689	160.3436

Table 4: Estimated Values for Wave 2 (Outcome Counts in Thousand)

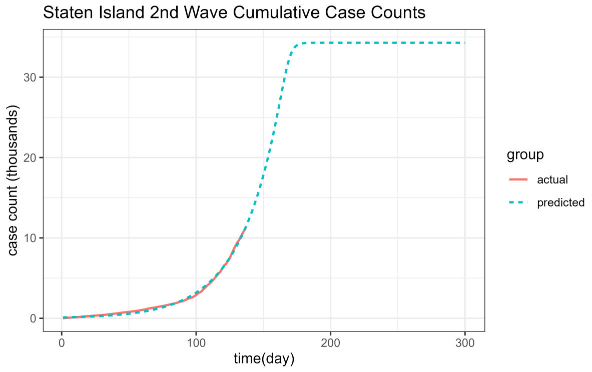
5.2.1 Cases



(a) Brooklyn (b) Bronx



(c) Manhattan (d) Queens



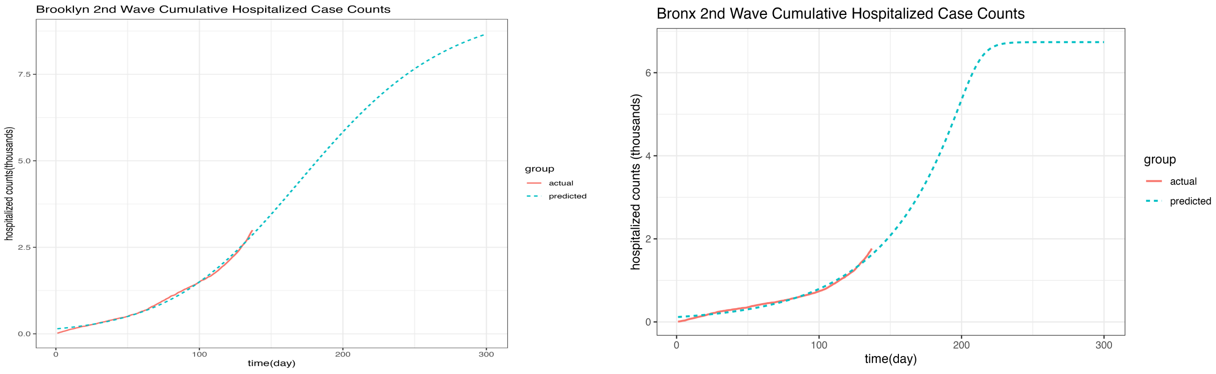
(e) Staten Island

Figure 6: Richard's fitted curves (red) vs. actual data (dashed blue) of Wave 2 Cumulative Cases (in Thousand) by Time (Day) per Borough

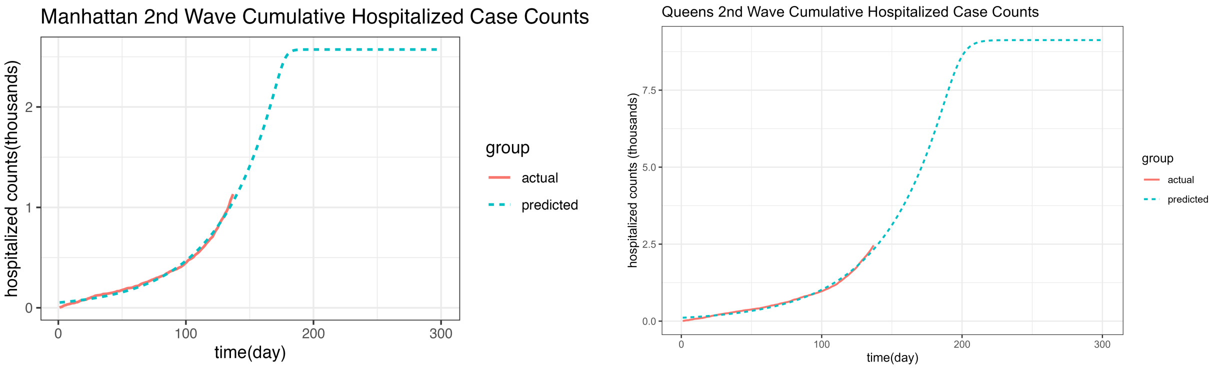
For cumulative case counts shown in Richard's model, in each five New York City boroughs, the four optimized parameters are shown in Table 4. Take Brooklyn for example, the starting value used as 80, 0.0266, 0.8 and 137 for a, k, d and t_0, respectively. Then, the estimated parameters are 85.5067, 0.3379, 14.6456 and 169.4569 for $\hat{a}$, $\hat{k}$, $\hat{d}$ and $\hat{t_0}$, respectively. The result indicates that the number of COVID-19 cases would reach approximately 85.507 thousand people in Brooklyn, with a growth rate 0.3379. Also, the inflection time was 169 days, which means the case outbreak would speed down after 169 days since the second wave starts. Finally, the fitted Richard's model by these parameters is highly consistent with the actual data (Fig. 6). Additionally, it gives prediction for the case counts after the record ending day based on the fitted Richard's model.

Also, among the five boroughs, the result $\hat{k}$ (=0.5719) shows Queens recorded the largest growth rate, followed by Manhattan (0.4047), Staten Island (0.381), and Bronx (0.365), and Brooklyn (0.3379). Thus, Queens is the most severe borough on case counts in the second wave of the pandemic.

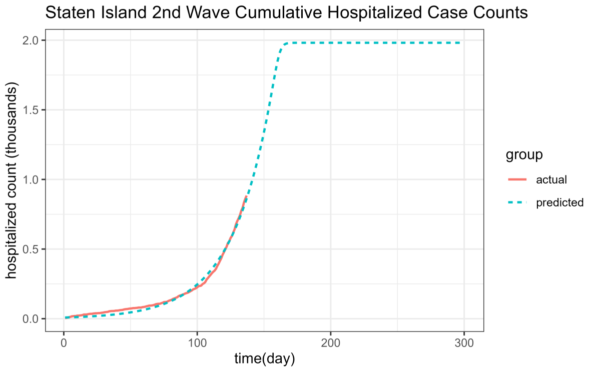
5.2.2 Hospitalizations



(a) Brooklyn (b) Bronx



(c) Manhattan (d) Queens



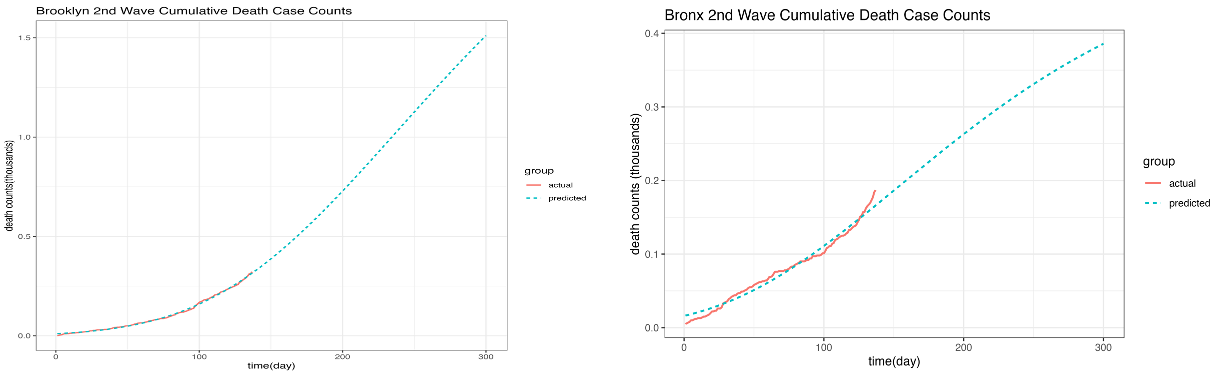
(e) Staten Island

Figure 7: Richard's fitted curves (red) vs. actual data (dashed blue) of Wave 2 Cumulative Hospitalizations (in Thousand) by Time (Day) per Borough

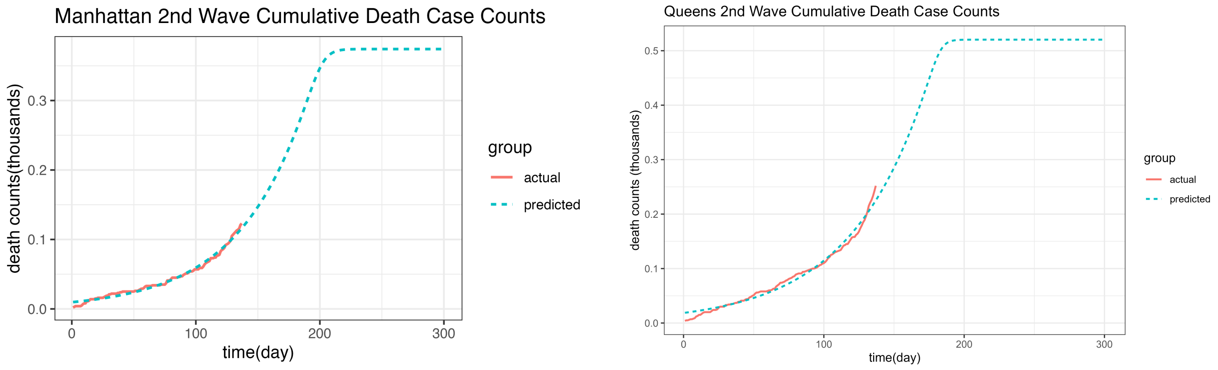
For cumulative hospitalization counts shown in Richard's model, in each five New York City boroughs, similarly, the four optimized parameters are shown in Table 4. Finally, the fitted Richard's model by these parameters is highly consistent with the actual data (Fig. 7). Additionally, it gives prediction for the hospitalization counts after the record ending day based on the fitted Richard's model.

Also, among the five boroughs, the result $\hat{k}$ (=0.5729) shows Staten Island recorded the largest growth rate, followed by Manhattan (0.3801), Queens (0.1956), Bronx (0.166), and Brooklyn (0.0185). Thus, Staten Island is the most severe borough on hospitalization counts in the second wave of the pandemic.

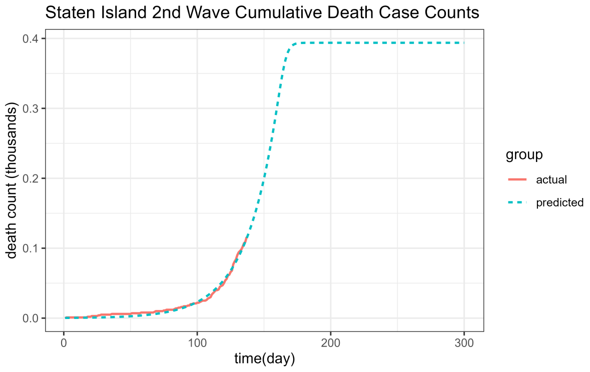
5.2.3 Deaths



(a) Brooklyn (b) Bronx



(c) Manhattan (d) Queens



(e) Staten Island

Figure 8: Richard's fitted curves (red) vs. actual data (dashed blue) of Wave 2 Cumulative Deaths (in Thousand) by Time (Day) per Borough

For cumulative death counts shown in Richard's model, in each five of New York City boroughs, similarly, the four optimized parameters are shown in Table 4. Finally, the fitted Richard's model by these parameters is highly consistent with the actual data (Fig. 8). Additionally, it gives prediction for the death counts after the record ending day based on the fitted Richard's model.

Also, among the five boroughs, the result $\hat{k}$ (=0.4208) shows Staten Island recorded the largest growth rate, followed by Queens (0.3125), Manhattan (0.2022), Brooklyn (0.0089), and Bronx (0.0081). Thus, Staten Island is the most severe borough on death counts in the second wave of the pandemic.

5.3 Between Waves

The severity between the two waves is also measured by $\hat{k}$ among cases, hospitalizations, and deaths.

Table 5 indicates four main results:

1. For cumulative cases, compared to the first wave, all boroughs have obviously increased growth rates for case counts, which means the cases are growing faster in all boroughs in the second wave.

	Case		Hospitalization		Death	
Borough	Wave 1	Wave 2	Wave 1	Wave 2	Wave 1	Wave 2
BK	0.0560	0.3379	0.0618	0.0185	0.0752	0.0089
BX	0.0617	0.365	0.0665	0.166	0.069	0.0081
MN	0.0511	0.4047	0.0667	0.3801	0.0655	0.2022
QN	0.0617	0.5719	0.0882	0.1956	0.0767	0.3125
SI	0.0723	0.381	0.0574	0.5729	0.0639	0.4208

Table 5: Vaccine Distribution Plan

2. For cumulative hospitalizations, compared to the first wave, Bronx, Manhattan, Queens show obviously increased growth rates, while Brooklyn and Staten Island show decreased growth rates, which means, in the second wave, the hospitalizations are growing faster in Bronx, Manhattan, Queens while growing slower in Brooklyn and Staten Island.

3. For cumulative deaths, compared to the first wave, however, Manhattan, Queens and Staten Island have sharply increased growth rates, while Brooklyn and Bronx have sharply decreased growth rates, which means, in the second wave, the deaths are growing faster in Manhattan, Queens and Staten Island while growing slower in Brooklyn and Bronx.

4. Additionally, the result shows the Bronx recorded an increased hospitalization growth rate but decreased death growth rate. On the contrary, Staten Island recorded a decreased hospitalization growth rate but increased death.

5.4 Predictions and Recommendations

With the fitted Richard growth functions for the second wave, I can get the predicted pandemic trends after December 11 for five boroughs' cumulative cases, hospitalizations and deaths respectively.

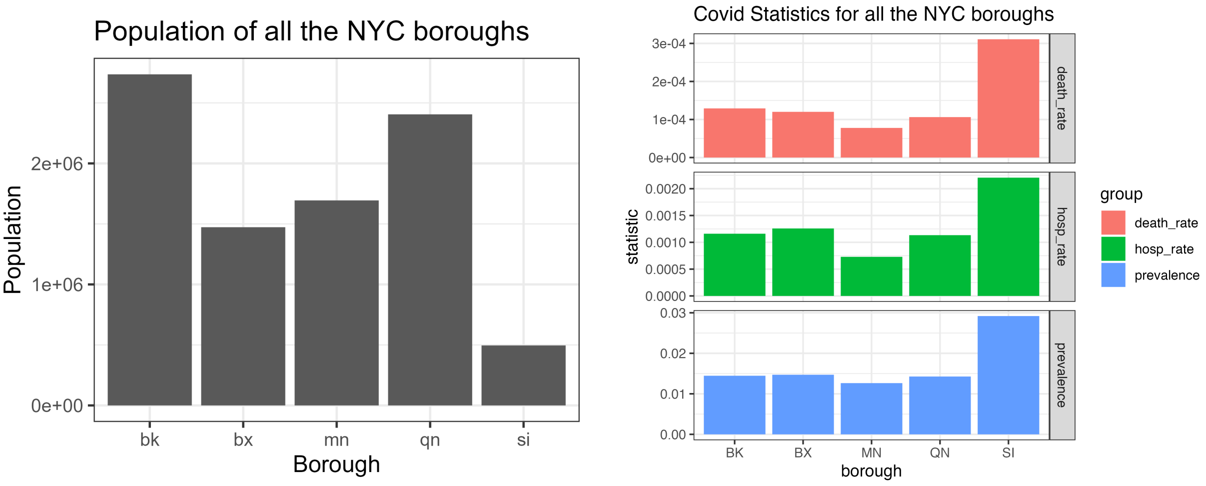
The distribution of COVID-19 vaccines is a critical step to control the pandemic. To give some suggestions for NYDOH regarding the vaccination roll-out plan, I utilize the predicted number of cumulative cases, hospitalizations and deaths on December 18, which is one week after the assumed middle point of the second wave.

For each borough, I consider its populations together with three measures of disease: prevalence, hospitalization rate and death rate.

$$\text{Prevalence} = \frac{\#\text{cumulative cases}}{\#\text{populations}}$$

$$\text{Hospitalization rate} = \frac{\#\text{cumulative hospitalization cases}}{\#\text{populations}}$$

$$\text{Death rate} = \frac{\#\text{cumulative death cases}}{\#\text{populations}}$$



(a)

(b)

Figure 9: (a) The population and (b) measuers of disease for five boroughs in NYC.

Fig. 9(a) shows the population of five NYC boroughs. Brooklyn has the highest population, followed by Queens, Manhattan, the Bronx and Staten Island. While the population of Staten Island is disproportionately lower than others, it has the highest prevalence, hospitalization rate and death rate (Fig. 9(b)). It can also be seen from Fig. 9(b) that Manhattan has the lowest values in all the three measures of disease.

The vaccination distribution strategy varies for different purposes. If the government prioritizes controlling prevalence, the proportion of vaccines distributed to each borough is the standardized statistics of population × prevalence. This is similar for hospitalization rate and death rate. Regarding how fast to roll out the vaccinations, I rank boroughs based on the fitted growth rate, that is the k in the Richard growth function. I suggest that the borough with the largest k should receive the vaccines first.

Table 6 gives the standardized statistics used to make a vaccination roll-out plan. For example, if the policymakers emphasize controlling prevalence, I distribute 30% of the vaccinations to Brooklyn, 16% to the Bronx, 16% to Manhattan, 26% to Queens and 11% to Staten Island. Vaccinations should be given out to Queens first since it has the largest k, which is the growth rate of the Richard curve fitted from the second wave of cumulative cases. It was followed by Manhattan, Staten Island, the Bronx and Brooklyn.

	Prevalence		Hospitalization		Death	
Borough	%	k rank	%	k rank	%	k rank
BK	30%	5	31%	5	33%	4
BX	16%	4	18%	4	17%	5
MN	16%	2	12%	2	12%	3
QN	26%	1	27%	3	24%	2
SI	11%	3	11%	1	14%	1

Table 6: Vaccine Distribution Plan

6 Conclusion

In Task 1, Richard's curves were fitted using a modified Newton-Raphson algorithm for the 1st wave of the daily case count in New York City's five boroughs. Residual sum of squares was minimized to fit the optimal models, and the selection of starting value was crucial. In the study, I found a reasonable and useful strategy in selecting the starting values which guarantees the convergence of optimization. The findings demonstrate that Richard's curve is a valuable tool for quantifying COVID-19 growth in various populations, facilitating inter-population comparisons.

In Task 2, I extended the analysis to both 2 waves of daily COVID-19 cases, hospitalizations, and deaths. After fitting Richard's curves, I was able to compare the same wave across different boroughs and compare the different waves for each borough. The analysis revealed that the growth rate of daily cases in the second wave was greater than that of the first wave across all five boroughs, indicating that the second wave was more severe than the first. Furthermore, the fitting results for the second wave can be leveraged for future case count predictions after December 11.

In Task 3, I established a vaccine roll-out plan based on both census data and epidemiological measures of interest, including prevalence, hospitalization rate, and death rate, which gave flexibility for the policymakers to control the target measure. A combination of multiple measures would also be a good choice to calculate the proportion of vaccine distribution.

7 Discussion

In this study, I applied Richard's curve to fit and predict COVID-19 daily cases, hospitalizations and deaths data across New York City's five boroughs from February 29, 2020 to

December 11, 2020. A modified Newton-Raphson algorithm was applied to fit the Richard's curves for all two wave COVID data. Direction check and step-halving steps helped to converge to the local minimum. A bound check step was added to ensure the updated parameters make sense. Initially, I applied Richard's curve to model the growth rate of the COVID-19 population, and observed that it exhibited good fitness in the first wave which was defined as the first 150 days since February 29. However, since the latter part data for the second wave was not completely provided in the dataset, further investigation is required to determine if the growth rate of the second wave follows Richard's curve.

When selecting the starting values, some parameters, such as the shape parameter d, are not explicitly defined, and are therefore selected somewhat arbitrarily. As the starting value of the parameter k depends on the value of d, a poorly chosen d value can lead to several problems. To address this, I attempted to improve the model fit by trying a series of starting values for d, fitting the model using different parameter sets, and comparing the results. However, the most reliable way to select an appropriate starting value for d is to gain a more comprehensive understanding of the properties of the COVID-19 curve.

Once I had established suitable starting values for the model, I was able to obtain well-fitting curves for the first wave. Since the data for the last day in the first wave was available, I was able to make accurate estimations of the starting values. Likely due to the availability of data for all time points in this wave, I found that the results of the first wave were not particularly sensitive to the starting values. However, this was not the case for the second wave. In experiments, I observed that even slight changes in the starting values could lead to completely different final parameter sets, highlighting the need for more solid and reasonable assumptions in selecting the starting value for the second wave.

Regarding the vaccine roll-out plan, I proposed three potential plans based on different epidemiological measures of interest. However, relying solely on one measure would be impractical. Instead, I recommend defining a combination of measures to improve the plan. Additionally, socioeconomic factors in the community, such as the proportion of the elderly population, race, and income level, should be taken into consideration to create a more comprehensive vaccine distribution plan.

References

[1] B. Gompertz, "Xxiv. on the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. in a letter to francis baily, esq. frs &c," *Philosophical transactions of the Royal Society of London*, no. 115, pp. 513–583, 1825.

[2] A. Tsoularis and J. Wallace, "Analysis of logistic growth models," *Mathematical biosciences*, vol. 179, no. 1, pp. 21–55, 2002.

Code Implementation of Newton-Raphson

Function that Returns RSS, Gradient, and Hessian of the RSS Function:

```
rich_func = function(dat, param_vec){
  cumy = dat$cum_case_count
  datt = dat$t

  a = param_vec[1]
  k = param_vec[2]
  d = param_vec[3]
  t0 = param_vec[4]

  yhat = a/(1+d*exp(-k*(datt-t0)))^(1/d)
  rss = -1 * sum((cumy-yhat)^2)

  # define the rss function to compute gradient and hessian
  rss_f = function(param_vec){
  a = param_vec[1]
  k = param_vec[2]
  d = param_vec[3]
  t0 = param_vec[4]

  return(-1*sum((cumy - a/(1+d*exp(-k*(datt-t0)))^(1/d))^2))
  }

  # compute gradient
  gradient = grad(rss_f, param_vec)


  # compute hessian
  hessian = hessian(rss_f, param_vec)

  return(list(rss = rss, gradient = gradient, hessian = hessian))
}
```

Modified Newton-Raphson Algorithm:

```
rich_newton_raphson <- function(dat_list, func, start_vec, tol=1e-10, maxiter = 1000){
  iter <- 0
  current_param <- matrix(start_vec, ncol = 1, nrow = 4)
  model_info <- func(dat_list, current_param)
  result <- c(0, -model_info$rss, current_param)
  prev_rss <- Inf
  while (iter < maxiter && abs(model_info$rss - prev_rss) > tol){
    iter <- iter + 1

    prev_rss <- model_info$rss
    previous_param <- current_param

    # adding the ascent direction step
    ## check whether the current step is in ascent direction
```

```
## if no, choose a constant gamma such that hessian - gamma*I is negative definite
direction <- -solve(model_info$hessian) %*% model_info$gradient

if (t(direction) %*% model_info$gradient < 0){
  hessian_eigen <- eigen(model_info$hessian)$values
  max_eigen <- max(hessian_eigen)
  gamma <- max_eigen + 1
  model_info$hessian <- model_info$hessian - gamma * diag(nrow(model_info$hessian))
}

current_param <- previous_param - 1*solve(model_info$hessian)
%*% model_info$gradient

# Bound check
if(current_param[1] < 0){
  current_param[1] <- previous_param[1]
}

if(current_param[2] < 0){
  current_param[2] <- previous_param[2]
}
if(current_param[3] < 0){
  current_param[3] <- previous_param[3]
}

if(current_param[4] < 0){
  current_param[4] <- previous_param[4]
}

previous_model_info <- model_info
model_info <- func(dat_list, current_param)

step_size = 1
rss_diff = -prev_rss - (-model_info$rss)

# Step-halving
while(rss_diff < 0){
  step_size = step_size/2
  prev_rss = previous_model_info$rss
  current_param <- previous_param - step_size *solve(previous_model_info$hessian)
  %*% previous_model_info$gradient
  if(current_param[1] < 0){
  current_param[1] <- previous_param[1]
}

## bound check
if(current_param[2] < 0){
  current_param[2] <- previous_param[2]
}
if(current_param[3] < 0){
  current_param[3] <- previous_param[3]
}
```

```
    if(current_param[4] < 0){
      current_param[4] <- previous_param[4]
    }
      model_info <- func(dat_list, current_param)
      rss_diff = -prev_rss - (-model_info$rss)

    }

    result <- rbind(result,
                    c(iter = iter,
                      rss = -model_info$rss,
                      current_param = current_param))
  }

  return(result)
}
```