# Analysis of the Relationship between Numbers of Cases and Mortality Rate for COVID-19 in New York City for the Month of April 2020

**Ken Biegeleisen**

Retired, New York, USA
Email: kb@notahelix.net

## Abstract

During the month which began just prior to the start of April 2020, the New York City mortality rate for COVID-19 infection increased continuously and gradually, from 1% - 2% at the beginning of the month, to a peak of about 11% by early May. During this same period, the total number of cases increased from about 20,000 to about 200,000. Since there is no reason, *a priori*, to expect the mortality rate of any single disease to have any particular mathematical relationship to the number of cases, a statistical analysis would appear to be in order. This analysis is presented, and one possible interpretation is suggested.

## Keywords

COVID-19, Pandemic, Mortality, Corona, Pneumococcal Pneumonia, Influenza

## 1. Introduction

The author of the present manuscript maintained a daily Excel spreadsheet for case numbers and mortality rate of COVID-19 infection in New York City, from March 27 to April 26, 2020. As Table 1 shows, the mortality rate just prior to the month of April, specifically on March 27, 2020, at which time the Department of Health reported a total of 23,112 cases, was between 1% - 2%. This would be the expected rate for any flu-like illness, and comparable to the COVID-19 mortality rate on the Diamond Princess Cruise ship [1], the latter of which might perhaps be the closest we shall get to a proper "laboratory" for studying the behavior of this virus in a closed system. The New York City mortality rate, however, immediately began increasing thereafter, so that by the end of the month of April,

*there had been a total of only four dates in which the daily mortality rate had not been higher than the day before.* The author ceased maintaining the spreadsheet after the month of April, but a follow up entry, on June 9, reveals that the mortality rate had continued to increase during the month of May, peaking at about 11%.

It thus appears that the mortality rate of this disease is inextricably-linked to the case numbers. One might expect a great deal of discussion about this rather remarkable phenomenon, but the author is not aware of such discussion in any reputable medical/scientific journal, or in the news media. The author has therefore undertaken a review of the April 2020 mortality statistics, and a consideration of the possible interpretation thereof.

## 2. Methods

The raw data for this statistical review is presented in Tables 1 and 2. Table 1 is an accumulation of the daily COVID-19 morbidity/mortality data for the City of New York, taken directly from the ongoing daily report of the New York City Department of Health (hereafter "NYCDH") (https://www1.nyc.gov/site/doh/covid/covid-19-data.page#download). Between March 27 and April 26, the author entered the NYCDH daily numbers into the above-referenced Excel spreadsheet. Table 1 shows the data from that spreadsheet.

The Table includes two additional dates; May 9 and August 21. These additional dates were added to document the fact that the New York City mortality rate had not yet peaked on April 26, but rather some time in early May, and that by August 21 the mortality rate was already declining, and was therefore past its peak.

The August 21 entry in Table 1 in a double-entry, encompassing both the counting method employed by the NYCDH during the month of April ("confirmed" cases, 8.3%), and the new method introduced by the NYCDH in later months ("confirmed" + "probable" cases, 10.4%). Only the first of these entries, 8.3%, can be properly compared to the April statistics, because no category of "probable" cases had yet been defined by the NYCDH in the month of April 2020.

Table 2 consists of data on false-positive COVID-19 test results. The numbers in the Table were reproduced directly from the single, un-numbered table in a June 2020 *Annals of Internal Medicine* article by Oran and Topol [2]. This data has been modified here by the addition of a line labeled "Total", which was conspicuously absent from the reference. To Table [2] the author has further added two additional species of "total", namely "average of averages" and "weighted average", as defined below.

The stated purpose of the authors of reference [2] was to present data on *false-positive* COVID-19 test results in sixteen different cohorts. Each cohort was the subject of a prior published report, and the various reports were merely consolidated by Oran and Topol [2]. The publication dates of the sixteen cohort

Table 1. New York City mortality rate of Covid-19 as a function of case numbers.

| Date | Cases | Deaths | Mortality rate | |
|---|---|---|---|---|
| 3/27/20 | 23,112 | 365 | 1.6% | |
| 3/28/20 | 26,697 | 450 | 1.7% | |
| 3/29/20 | 30,765 | 672 | 2.2% | |
| 3/30/20 | 33,474 | 776 | 2.3% | |
| 3/31/20 | 38,087 | 914 | 2.4% | |
| 4/1/20 | 41,771 | 1096 | 2.6% | |
| 4/2/20 | 48,462 | 1397 | 2.9% | |
| 4/3/20 | 49,707 | 1562 | 3.1% | |
| 4/4/20 | 56,289 | 1867 | 3.3% | |
| 4/5/20 | 58,570 | 2061 | 3.5% | |
| 4/6/20 | 60,850 | 2254 | 3.7% | |
| 4/7/20 | 64,813 | 2496 | 3.9% | |
| 4/8/20 | 68,776 | 2738 | 4.0% | |
| 4/9/20 | 80,204 | 4260 | 5.3% | |
| 4/10/20 | 87,725 | 4778 | 5.4% | |
| 4/11/20 | 94,409 | 5429 | 5.8% | |
| 4/12/20 | 98,715 | 5742 | 5.8% | |
| 4/13/20 | | (Not recorded) | | |
| 4/14/20 | 104,410 | 6182 | 5.9% | |
| 4/15/20 | 107,263 | 6589 | 6.1% | |
| 4/16/20 | 111,424 | 6840 | 6.1% | |
| 4/17/20 | 117,565 | 7563 | 6.4% | |
| 4/18/20 | 121,967 | 8006 | 6.6% | |
| 4/19/20 | 126,368 | 8448 | 6.7% | |
| 4/20/20 | 129,788 | 8811 | 6.8% | |
| 4/21/20 | 132,467 | 9101 | 6.9% | |
| 4/22/20 | 135,451 | 9522 | 7.0% | |
| 4/23/20 | 138,435 | 9944 | 7.2% | |
| 4/24/20 | 141,754 | 10,290 | 7.3% | |
| 4/25/20 | 146,165 | 10,626 | 7.3% | |
| 4/26/20 | 150,576 | 10,961 | 7.3% | |
| (No records were kept between April 26 and June 9) | | | | |
| 6/9/20 | 204,253 | 21,877 | 10.7% | |
| 8/21/20 | 227,724 | 19,003 | 8.3% | "Confirmed" cases |
| 8/21/20 | 227,724 | 23,639 | 10.4% | "Confirmed" + "probable" cases |

Source: https://www1.nyc.gov/site/doh/covid/covid-19-data.page#download.

Table 2. Prevalence of asymptomatic COVID-19 Infections from sixteen cohorts.

| Location of cohort: | Number tested: | Number positive: | Number not sick: | Percentage not sick: |
|---|---|---|---|---|
| Iceland | 13,080 | 100 | 43 | 43% |
| Vo', Italy | 5155 | 102 | 43 | 42% |
| Diamond Princess cruise ship | 3711 | 712 | 331 | 46% |
| Boston homeless shelter | 408 | 147 | 129 | 88% |
| NYC obstetric patients | 214 | 33 | 29 | 88% |
| USS Theodore Roosevelt | 4954 | 856 | 500 | 58% |
| Japanese citizens in Wuhan | 565 | 13 | 4 | 31% |
| Greeks in UK, Spain & Turkey | 783 | 40 | 35 | 88% |
| Charles de Gaulle aircraft carrie | 1760 | 1046 | 500 | 48% |
| LA homeless shelter | 178 | 43 | 27 | 63% |
| Nursing facility, Washington | 76 | 48 | 3 | 6% |
| Prisoners, 4 states | 4693 | 3277 | 3146 | 96% |
| NJ university/hospital staff | 829 | 41 | 27 | 66% |
| Indiana residents | 4611 | 78 | 35 | 45% |
| Argentine cruise ship | 217 | 128 | 104 | 81% |
| San Francisco residents | 4160 | 74 | 39 | 53% |
| TOTAL: | 45,394 | 6738 | 4995 | 74% |

Source: Oran & Topol [2]. Average of averages (as defined in the text): 59%; Weighted average (as defined in the text): 54%.

reports ranged from the beginning of the worldwide pandemic in early 2020, to the date of the Oran and Topol [2] publication, June 3, 2020. Reference [2] was thus not a research article *per se*, but rather a review article, encompassing data from 16 previous epidemiologic studies, which emerged from many different nations and settings.

Returning to Table 1, what we see there is a clear and unequivocal relationship between *case count* and *mortality rate* for COVID-19 infection. The data from Table 2 provides the basis for a possible explanation of the Table 1 data.

## 3. Results

A mere glance at Table 1 shows that the deadliness, *i.e.*, mortality rate, of COVID-19 infection, apparently increased alarmingly and daily during the spring of 2020, *pari passu* with the case count. During the period extending from March 27 to August 21, the daily mortality rate increased from a starting point of 1% - 2%, peaking at about 11%, then declining to about 8%. The relationship between case count and mortality rate is overwhelmingly evident from the Table itself. Whereas the explanation for the relationship between case count and mortality rate may be non-obvious, the numbers themselves can be readily understood by anyone with a grammar school education. A glance at the beginning and peak

numbers tells the entire story:

| Date | Cases | Deaths | Mortality rate |
|---|---|---|---|
| 03/27/2020 | 23,112 | 365 | 1.6% |
| 06/09/2020 | 204,253 | 21,877 | 10.7% |

During the period under consideration, the case count increased nearly 10-fold. Simultaneously, the mortality rate increased by a factor which, although it cannot be precisely specified, clearly also lies between 5-fold and 10-fold.

These numbers are striking, yet, as to an explanation thereof, none has been forthcoming.

We look next at the false-positive test result data in Table 2, from the review article of Oran and Topol [2]. We shall totally ignore the *obiter dictum* of their article, and focus solely on the numbers therein, which are summarized in Table 2. "False-positive" was defined by those authors as meaning that the COVID-19 test result was "positive", but that the subject was not sick. The question, of whether people who are *not* sick can nevertheless spread the disease, is an interesting question, but we shall not occupy ourselves with it, because we are only interested in the numbers.

Table 2 shows that the overall false-positive rate, for all sixteen cohorts, was 74%. The rates of false-positivity, for all sixteen cohorts, ranged from a low of 6% to a high of 96%. If we were to presume that each of these sixteen different rates was meaningful and specific for its own particular cohort, regardless of the size thereof, then we could thereby meaningfully speak of an "average" rate of false-positivity (which the present author has designated "average of averages" in Table 2), which would be 59%. In order to give more weight to the size of the cohort, the author has also somewhat-arbitrarily defined a "weighted average". This was obtained by multiplying the [number of cases in each cohort] by the [percentage "not sick" in that same cohort], then adding those numbers together (giving a sum of 24,268.5). That sum was then divided by the total number of cases in all sixteen cohorts (45,394). This gave a "weighted average" of 54%.

While there is an unequivocal vagueness about all this, one thing is certain: The COVID-19 test kits in use in April 2020 had false-positive rates of somewhere between 50% - 100%. The author has no way of adjudicating these discrepancies, so he shall, for the sake of making the primary argument below, presume that the overall 74% false-positive rate; *i.e.*, the rate for all 16 cohorts combined; is the one most likely to be correct. If, perchance, that rate turns out to be closer to 50%, the argument below still stands in its entirety, because it is a qualitative argument, requiring no presumption other than that the test has a significant rate of false-positive results.

## 4. Discussion

The purpose of this article is to increase awareness of the question: *What is the*

*nature of the self-evident relationship between case numbers and mortality rate in COVID-19 infection*? The author has raised this question in discussions with many of his colleagues, and has patiently awaited a discussion thereof in the news media, but no answer to the question is ever given. As far as the news media are concerned, the question itself is never even asked.

The author shall now propose an answer to the question. It is not the purpose of this article to provide a definitive answer, but merely to propose a single possible answer in a setting in which no other answer has even been proposed.

In suggesting an answer to the question, the author shall dismiss, at the outset, the possibility that the virus "morphed" into a more lethal form in April, then "morphed" back to a less lethal form in May. In addition to the fact that there is no published evidence for any such thing, there is, moreover, a darwinian problem therein. If the virus "morphed" into a more lethal form, then it and its host would be buried, and that particular strain of virus would cease propagating. There cannot be any evolutionary survival advantage for such a virus. This, of course, is a logical argument, but if anyone wishes to propose it nevertheless, it is incumbent upon that person to provide a plausible explanation for such an unlikely metamorphosis.

A better explanation can be found in a consideration of the actions that might have taken place in any hospital emergency room at the height of this pandemic. Suppose that on some day, 4 patients arrived, all febrile and coughing. It is an undeniable fact that only one disease was likely to have been tested for: COVID-19. If the test results had been positive, the 4 patients would surely have been admitted with a diagnosis of "COVID-19", and if any of them died in the hospital, the death certificates would surely have given "COVID-19" as the cause of death. In addition to the likelihood that no other disease would have been tested for, the author believes that he stands on safe ground when he asserts that government health departments have issued "guidelines" for the diagnosis of COVID-19, even in the complete absence of any test at all.

The author also thinks that he stands on safe ground when he asserts that the regular annual United States death toll (and, by extension, the annual New York City death toll) from influenza, pneumococcal pneumonia, Legionnaire's disease, tuberculosis, smokers' emphysema/bronchitis, and a host of "other" pulmonary diseases, is very likely the same in 2020 as in any previous year. In the absence of any data to the contrary, it would be foolish to presume otherwise. However, for the year 2020, the author cannot locate public health statistics for any deaths from any of these "other" diseases, and we shall very likely never know how many people did die from them in 2020, because the hospitals reported having been "overwhelmed" by the COVID-19 pandemic, wherefore these "other" diseases were rarely tested for, if indeed they were tested for at all.

In the United States, public health records show plainly and openly that the death tolls in previous years, from all these "other" respiratory diseases combined, regularly and reliably add up to some number between 100,000 - 200,000.

This year, however, if we are to take the absence of reporting as "evidence", then we might be compelled to believe that there were actually *no deaths* from any of these "other" diseases. It is neither reasonable nor rational to believe such a thing, but at this point we shall never know the exact numbers of 2020 deaths from these "other" respiratory diseases, because they were not tested for.

The important question is a statistical one: *When people arrive at the emergency room coughing, what is the effect of presuming that they all have COVID-19, when such a diagnosis is based upon the current test?*

We know now that the test has a 50% - 100% false-positive rate. It hardly matters whether it is 50% or 100%; *the test is not accurate.* As stated previously, the author shall, for the sake of argument, accept the Annals of Internal Medicine (reference [2]) overall false-positive rate of 74% to be the correct number, admitting without protest that it might be as low as 50%, because the argument the author shall make is primarily qualitative.

When 4 patients arrive at the emergency room coughing, and test "positive" for COVID-19, there is no scientific basis for excluding the possibility that 74% of them, *i.e.*, about 3/4 of them, have a false-positive test result. That means that, statistically-speaking, one of the 4 really does have COVID-19, but the other 3, having false-positive test results, but coughing nevertheless, must have a different disease. Whether it is influenza (which kills 50,000 Americans per year), pneumococcal pneumonia (which also kills 50,000 Americans per year), Legionnaire's Disease (which kills tens of thousands per year—exact numbers not certain), or any other potentially-fatal pulmonary disorder, will very likely never be known. But one thing can be known with virtual certainty: statistically speaking, it is very unlikely that all four have been coughing from COVID-19 infection.

Nevertheless, all of them will be admitted as "COVID-19" victims. Next, please consider the hospital ward. In bed #1, the patient may indeed have COVID-19. But in the bed (or room) next to him, the 2nd patient may have had a false-positive test result, and may actually be suffering from influenza. The 3rd patient in the bed next to patient #2 may actually have pneumococcal pneumonia, and in the next bed the 4th patient may actually have Legionnaire's Disease.

What then? Each of these patients will then be seriously exposed to the pathogens from the adjacent beds, and it is entirely unrealistic to presume anything other than that there will be a highly-significant risk of super-infection. Thus, an unknown number of patients will enter the hospital with one disease, but become super-infected with one, two or even more additional pulmonary diseases, from the infected droplets or touched surfaces of these "other" cases.

In many discussions of COVID-19 mortality, we hear much talk about "co-morbidity". It is widely-conceded, even in nightly television news reports, that most of the COVID-19 deaths have involved the co-morbidities of old age and/or pre-existing serious underlying illness. Thus, it is not in doubt, in any quarter, that the mortality rate from COVID-19 infection is increased in the

presence of other simultaneously-occurring diseases.

The presence, in a single hospital ward, of multiple patients suffering from multiple respiratory illnesses, and dying at 5 - 10 times the mortality rate of COVID-19 at the beginning of the pandemic, would must assuredly be accounted "co-morbidity". The hypothetical scenario laid out above can therefore be quite plausibly invoked to adequately and fully explain why the mortality rate from COVID-19 appears to go up with the case count.

## 5. Conclusions

Ordinarily, any disease is associated with its own particular mortality rate. That rate may, of course, vary with all sorts of accompanying conditions and circumstances. But a disease which becomes more virulent simply because the case count has gone up is, to the best of my knowledge, unprecedented.

The analysis which culminated in the present manuscript began with a simple question: "Why did the mortality rate of COVID-19 in New York City increase, *pari passu* with the case count, throughout the month of April 2020?" This manuscript offers a simple answer to the simple question: *co-morbidity.* Is the answer correct? We cannot know with certainty. The purpose of this manuscript is not to tender any definitive answers, but mainly to focus more attention on the question itself, to which there may be multiple other answers. But if there are multiple other answers, they should be made public, because the question is medically-significant.

## Acknowledgements

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

[1] Russell, T.W., Hellewell, J., Jarvis, C.I., van Zandvoort, K., Abbott, S., Ratnayake, R., CMMID COVID-19 Working Group, Flaschel, S., Eggol, R.M., Edmunds, W.J. and Kucharski, A.J. (2020) Estimating the Infection and Case Fatality Ratio for Coronavirus Disease (COVID-19) Using Age-Adjusted Data from the Outbreak on the Diamond Princess Cruise Ship, February 2020. *Eurosurveillance*, **25**. https://doi.org/10.2807/1560-7917.ES.2020.25.12.2000256

[2] Oran, D.P. and Topol, E.J. (2020) Prevalence of Asymptomatic SARS-CoV-2 Infection: A Narrative. *Annals of Internal Medicine*. https://doi.org/10.7326/M20-3012