The pandemic gap: the critical distinction between “infected” and “sick”; understanding mortality rates; context around the Spanish flu in a pre-antibiotic world

A lot of data is being made available on the coronavirus, but most of it requires careful analysis before drawing conclusions. Here’s what is clear: the rate of reported infections outside China is now accelerating at a similar rate to what was happening a month and a half ago in China itself. In response, many parts of the world have adopted control measures such as quarantine/lockdown, school closures, etc.

That’s where the clear part ends. The complex thing about pandemics is that early mortality rate estimates tend to decline over time. Why? Here are four simple measures that matter in the context of a pandemic:

(a) the population of a given geographical area
(b) the total number of infected individuals, including both asymptomatic people and people that get sick
(c) the total number of people that are infected, get sick and self-report
(d) the total number of people that die

During the haze of a pandemic, the best estimates that entities like the World Health Organization often derive are based on (a), (c) and (d), and even things like (d) are complicated by pandemics affecting older individuals with pre-existing conditions. They do not know (b) upfront, and sometimes it is never known, or only known with the passage of time. Take the Swine Flu (H1N1/2009) as an example. Early estimates in the fall of 2009 from the WHO\(^1\) pegged the H1N1 mortality rate at 1.0%-1.3%, since they were dividing (d) by (c). Four years later, a study from the WHO and the Imperial College of London\(^2\) estimated H1N1 mortality as a function of total infections, including both the asymptomatic and the sick. Their revised H1N1 mortality rate using (b) as a denominator: just 0.02%.

So, please treat estimated infection rates and mortality rates with care, since they can mean very different things. Marc Lipsitch from Harvard has estimated that 40% - 70% of the world’s population could become infected\(^3\). Lipsitch himself makes it clear that this number is an example of (b) and not (c) and that there is an enormous gap between the two, so please do not multiply population by 40%-70% and then multiply by a mortality rate assumption. The vast majority of infected people will likely not become sick, and around 80% of people who get sick develop mild infections rather than severe ones.

---

1. WHO Situation Report, Pandemic (H1N1) 2009, Update 76, November 22, 2009.
To get a sense for the possible spread of COVID-19, let’s look at China now that its trajectory of reported cases is in decline\(^4\). The table shows (a), (c) and (d), since (b) and all statistics derived from it are unknown. **Outside Wuhan/Hubei, China mortality and infection rates are much lower**, even in poor provinces that neighbor Hubei itself. As illustrated on the next page, China’s ex-Hubei mortality rates are not that different from seasonal flu mortality rates for individuals over 65 in developed countries like the US.

COVID-19: Population, infections, fatalities and derived statistics in select Chinese provinces

Columns B, C and D are reported based on cumulative figures to-date (March 8)

<table>
<thead>
<tr>
<th>Disease</th>
<th>Region</th>
<th>Population</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>C/A</th>
<th>D/B</th>
<th>B/A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(asymptomatic + sick)</td>
<td>Infected (sick only)</td>
<td>Infected (sick only)</td>
<td>Deaths</td>
<td>Death rate of the sick</td>
<td>Death rate, all infected</td>
<td>Total infections to population</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Wuhan</td>
<td>11,000,000</td>
<td>Unknown</td>
<td>49,871</td>
<td>2,370</td>
<td>4.75%</td>
<td>0.45%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Hubei</td>
<td>59,170,000</td>
<td>Unknown</td>
<td>67,707</td>
<td>2,986</td>
<td>4.41%</td>
<td>0.11%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Guangdong</td>
<td>113,460,000</td>
<td>Unknown</td>
<td>1,352</td>
<td>7</td>
<td>0.52%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Henan</td>
<td>96,050,000</td>
<td>Unknown</td>
<td>1,272</td>
<td>22</td>
<td>1.73%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Zhejiang</td>
<td>57,370,000</td>
<td>Unknown</td>
<td>1,215</td>
<td>1</td>
<td>0.08%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Hunan</td>
<td>68,990,000</td>
<td>Unknown</td>
<td>1,018</td>
<td>4</td>
<td>0.39%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Anhui</td>
<td>63,240,000</td>
<td>Unknown</td>
<td>990</td>
<td>6</td>
<td>0.61%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Jiangxi</td>
<td>46,480,000</td>
<td>Unknown</td>
<td>935</td>
<td>1</td>
<td>0.11%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Shandong</td>
<td>100,470,000</td>
<td>Unknown</td>
<td>758</td>
<td>6</td>
<td>0.79%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Jiangsu</td>
<td>80,510,000</td>
<td>Unknown</td>
<td>631</td>
<td>-</td>
<td>0.00%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Chongqing</td>
<td>31,020,000</td>
<td>Unknown</td>
<td>576</td>
<td>6</td>
<td>1.04%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Sichuan</td>
<td>83,410,000</td>
<td>Unknown</td>
<td>539</td>
<td>3</td>
<td>0.56%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Heilongjiang</td>
<td>37,730,000</td>
<td>Unknown</td>
<td>481</td>
<td>13</td>
<td>2.70%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Beijing</td>
<td>21,540,000</td>
<td>Unknown</td>
<td>428</td>
<td>8</td>
<td>1.87%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Shanghai</td>
<td>24,240,000</td>
<td>Unknown</td>
<td>342</td>
<td>3</td>
<td>0.88%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Hebei</td>
<td>75,560,000</td>
<td>Unknown</td>
<td>318</td>
<td>6</td>
<td>1.89%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Fujian</td>
<td>39,410,000</td>
<td>Unknown</td>
<td>296</td>
<td>1</td>
<td>0.34%</td>
<td>0.00%</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

What might reported UK government central estimates of 100,000 deaths imply? Something very different than what’s reportedly happening in China itself

Estimate 1: 38% of population infected; 40% of infected people get sick; 1% mortality rate as % of sick population

COVID-19 UK 66,440,000 25,000,000 10,000,000 100,000 1.00% 15.05% 0.400% 38%

Estimate 2: 33% of population infected; 10% of infected people get sick; 4.4% mortality rate as % of sick population

COVID-19 UK 66,440,000 22,000,000 2,258,960 100,000 4.43% 3.40% 0.455% 33%

Sources: World Health Organization Coronavirus Situation Reports, Health Commission of Hubei Province, Sunday Times, JPMAM. 2020

That’s what makes the UK’s reported estimate of 100,000 deaths very puzzling\(^5\). To arrive at such an outcome, we had to assume that 38% of the entire UK population is infected (i.e., similar to the 1918 Spanish flu), and that 40% of infected people get sick and then experience 1% mortality; or we had to assume that only 10% of infected people get sick but then experience 4.4% mortality that’s equal to the epicenter of the virus outbreak in Wuhan. Even after accounting for Chinese infection/death underreporting and the difficulty Western countries might have replicating what China has done (the largest lockdown/quarantine in the history of the world, accomplished via AI, big data and different privacy rules\(^6\)), both of our modeled UK outcomes would be magnitudes worse than what’s occurring in China and South Korea. This is another pandemic gap that we are still trying to understand and reconcile.

Why are Hubei mortality rates so much higher? Likely answer: a collapse in Hubei’s healthcare infrastructure given a flood of infections. This prevented Hubei doctors from providing round-the-clock care that other cities provided to keep patients alive until their immune systems could fight the disease.

---

\(^4\) Column C shows total infections to-date and does not reflect recoveries. Provinces such as Anhui, Jiangsu, Fujian, Qinghai and Tibet have now been declared as infection-free by the Chinese government.

\(^5\) “Coronavirus: Boris Johnson keeps calm but Whitehall plans for worst”, Sunday Times, March 8, 2020

\(^6\) “China suppressed Covid-19 with AI and big data” (via smartphones), Asia Times, March 3, 2020
Mortality rates for COVID-19 so far. With that context, it should be clearer how difficult it is to compare mortality rates in the heat of a pandemic given uncertainties on numerators/denominators. That said, even an uncertain exercise can reveal important trends. As shown below, mortality rates in Hubei are heavily influencing both China and global measures. Outside Hubei province, China mortality rates are much lower. Italy’s mortality is high right now; we will have to wait and see if it declines once self-reported infections increase in the weeks ahead. South Korea has conducted the most widespread testing and its mortality rate may be closer to the "deaths to true infections" ratio discussed on the first page. Iran is in the early stages as well, and has a very weak healthcare system further compromised by sanctions and a steady brain drain of medical personnel.7

Mortality rates: COVID-19 vs other diseases, and the impact of Wuhan/Hubei
Mortality rate; bubble size indicates relative number of fatalities for COVID-19 only

Mortality rates shown for all countries with at least 1,000 infections to date. Sources: CDC, China National Health Commission, Center for Health Protection (HK), Global Health Data Exchange, World Health Organization, Netherlands Institute for Health Services Research, Imperial College of London, Mayo Clinic, JPMAM. 2020.

Healthcare Access and Quality Index

Physicians per 1,000 people


7 "Iran Faces Serious Shortage Of Doctors Due To Emigration", Radio Farda, Iran In-Depth, June 30, 2019
The Spanish Flu (1918-1920) in context, and why it’s a poor proxy for COVID-19

- **No vaccine was ever developed to combat the Spanish Flu.** In contrast, a SARS vaccine was developed in response to the 2002 outbreak but was never used since public health measures (closing workplaces, people working at home, etc) got the disease under control by May 2003 before the vaccine was ready. Since that time, Harvard scientists have found the antibody which blocks SARS and MERS from entering human cells, which were used to develop antibody therapies (which are different from vaccines, which are the treatment of choice and much cheaper to produce). The new COVID-19 virus shares 86% genetic similarity with SARS, so scientists aren’t starting from ground zero.

- **There were no antibiotics in 1918** to treat secondary bacterial infections associated with influenza. From a paper marking the 100th anniversary of Spanish Flu: “in 1918, most severe influenza-associated pneumonias were associated with secondary bacterial infections… high pandemic case fatality during the fall 1918 pandemic resulted primarily from increased frequency, and not increased severity, of secondary bacterial pneumonias, especially in young adults” (see chart). Furthermore, without secondary bacterial pneumonia, “experts generally believed that most patients would have recovered”. The first antibiotic was discovered in 1929 but mass antibiotic production did not occur until the 1940’s.

- **The US CDC reports** that lab experiments with recombinant influenza viruses containing genes from the 1918 virus suggest that the 1918 and 1918-like viruses would be as sensitive as other virus strains to FDA-approved anti-influenza drugs rimantadine and oseltamivir.

- **As further indication of a world without antibiotics and other healthcare innovations such as anti-virals, ICU-level hospital care, ventilators, etc,** the US life expectancy for men and women ranged from 50-55 years before the onset of the Spanish Flu in 1918.

---

8 China began clinical trials of a SARS vaccine in November 2003, while in the US, the first human SARS trials began in December 2004, conducted by the National Institute of Allergy and Infectious Diseases.

9 Harvard professor Wayne Marasco identified a single antibody out of a 27-billion antibody library that blocked the SARS virus from entering human cells. Marasco is actively testing new antibodies in search of one that will have the same effect on SARS-CoV-2 (COVID-19).

10 “The 1918 influenza pandemic: 100 years of questions answered and unanswered”, Taubenberger et al, Viral Pathogenesis and Evolution Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Science Translational Medicine, July 2019.

11 Male life expectancy at birth was ~25 during the Roman Era; rose to ~33 by the Middle Ages; and hovered between 30 and 40 until the late 1800’s. Starting in the 1920’s, the innovations cited above ushered in the most remarkable improvement in life expectancies in the history of the world.
Appendix charts: quarantine, SARS vs COVID, high frequency indicators in China, OECD tourism

- China’s quarantine/lockdown is the largest in the history of the world, as per a paper on comparing COVID-19 and SARS transmission trends. Evidence of this is shown in the first chart: China’s manufacturing and service sector declines are larger than in 2008. The second chart shows a small recovery in activity from the lows. We have reasons to believe that the electricity numbers overstate actual activity.

- In 2003, SARS was eventually contained by surveillance, isolation of patients, strict enforcement of quarantine of all contacts, and in some areas community-level quarantine. By interrupting human-to-human transmission, SARS was eradicated. Isolation was effective for SARS because peak infectiousness occurred after patients were already very ill with respiratory symptoms and could be easily identified. Although asymptomatic patients were reported for SARS, no known transmission occurred from these patients.

- The new virus SARS-CoV-2 (which causes the COVID-19 disease) has 86% similarity with the 2002 SARS-CoV virus, and both have median incubation times of ~5 days and basic reproductive numbers of ~2.2. The first paper linked below estimates the mean serial interval of COVID-19 at 7.5 days (the time it takes for an infected person to become contagious to others), similar to the SARS virus. However, a separate paper from the International Journal of Infectious Diseases may explain why SARS-CoV-2 is spreading more rapidly: the authors estimate the SARS-CoV-2 serial interval at just 4.5 days, which is less than its incubation period (i.e., when symptoms occur). That means that asymptomatic individuals could be contagious before they know they have the virus. If that’s the case, that’s quite different than SARS, since isolation of severely ill COVID-19 patients at the time they show up at health-care facilities would be too late.

---

12 “Can we contain the COVID-19 outbreak with the same measures as for SARS?”, Wilder-Smith et al, London School of Hygiene and Tropical Medicine, March 5, 2020

13 “Serial interval of novel coronavirus infections”, International Journal of Infectious Diseases, March 4, 2020