Supplementary Appendix

Using the COVID-19 to influenza ratio to estimate the numbers of symptomatic COVID-19 cases in Wuhan prior to the lockdown

Zhanwei Du[1], Benjamin J. Cowling[2] and Lauren Ancel Meyers[1,3*]

1. The University of Texas at Austin, Austin, Texas 78712, The United States of America
2. The University of Hong Kong, 7 Sassoon Rd, Hong Kong SAR, China
3. Santa Fe Institute, Santa Fe, New Mexico, The United States of America

[*] Correspondence: Lauren Ancel Meyers (laurenmeyers@austin.utexas.edu)

Table S1. Model Parameters and Data Sources. Parameters with an age indicator (a) have separate values for the 30-59 and 60+ age ranges.

Symbol	Description	Values	Sources
$H_{d,a,\tau}$	Number of COVID-19 outpatients in age group a in district d over time period τ	Estimated	
r	Ratio of ILI outpatients that are COVID-19 positive versus influenza positive (adults over 30)	1.28 [95% CI: 0.56-3.26]	Ref. (1): 9 COVID-19 positive and 7 influenza positive among 45 tested ILI throat swab samples taken from adult outpatients at two hospitals in central Wuhan, December 30, 2019-January 19, 2020

Appendix: Using the COVID-19 to influenza ratio to estimate Wuhan cases

N_d	Age-stratified population sizes in district d	2010 population scaled by the ratio of the 2019 to the 2010 total Wuhan population (Table S1)	Sixth National Census of the People's Republic of China in 2010 (2) and total population of Wuhan in 2019 (11.08 million) (3)
Ω_τ	Number of outpatient visits (all causes) in Wuhan across all ages over time period τ	42274, 38702 and 39245 over three weeks, respectively	China CDC weekly reports of outpatient visits in Wuhan, December 30, 2019-January 19, 2020 (1)
Θ_τ^a	Number of ILI outpatients in age group a in Wuhan over time period τ	Age 30-59: 47, 43 and 79 Age 60+: 14, 4 and 28, for each of the three weeks, respectively	China CDC weekly reports of ILI outpatients in Wuhan, December 30, 2019-January 19, 2020 (1)
Φ_τ	Percent of influenza positive tests	25%, 28.6% and 34.3% for each of the three weeks, respectively	Ref. (1): 25%, 28.6% adult (30+) influenza positive among 160 ILI throat swab samples, from December 30, 2019 to January 12, 2020. On average, the percent of influenza in adults (30+) is 48.1% of those under 30 years old, indicating 34.3% of influenza positive among ILI adults (30+) in the third week, given 71.4% of influenza positive among ILI people (0-29).
T_d	Epidemic doubling time	7.3 [95% CrI: 6.3–9.7] days 5.2 [95% CrI: 4.6–6.1] days	Ref. (4) and (5)

Appendix: Using the COVID-19 to influenza ratio to estimate Wuhan cases

Table S1. Estimated COVID-19 adult infections in the 13 central districts of Wuhan, from December 30, 2019 to January 19, 2020. Values are medians and 95% bounds across 1,000 Monte Carlo samples.

District	Population (≥30)	Infections		
		all (≥30)	30-59	60+
Wuchang*	723,612	711 (353-1,262)	543 (273-982)	164 (80-300)
Jiang'an*	586,685	576 (286-1,023)	448 (225-810)	125 (61-229)
Hongshan*	583,563	573 (285-1,017)	463 (233-837)	107 (52-196)
Huangpi	540,473	531 (264-942)	388 (195-701)	140 (69-257)
Qiaokou*	500,163	491 (244-872)	398 (200-719)	90.7 (44-166)
Xinzhou	493,939	485 (241-861)	365 (184-661)	117 (57.2-214)
Jianghan*	445,042	437 (217-776)	349 (176-631)	86 (42-157)
Hanyang*	378,570	372 (185-660)	290 (146-525)	80 (40-146)
Caidian	367,205	361 (179-640)	274 (138-496)	85 (41-155)
Jiangxia	348,505	342 (170-608)	266 (134-482)	74 (36-136)
Qingshan	332,212	326 (162-579)	241 (121-436)	84 (41-153)
Dongxihu	281,403	276 (137-491)	223 (112-404)	51 (25-94)
Hannan	76,057	75 (37-133)	59 (30-106)	16 (8-29)
Wuhan	5,657,429	5,558 (2,761-9,864)	4,307 (2,169-7,790)	1,218 (595-2,232)

* These 7 districts are located in central Wuhan; the other six are suburban.

Data

We analyzed the age-stratified numbers of outpatient visits for ILI in Wuhan city (1), COVID-19 positive and influenza positive patients among tested ILI throat swab samples at two hospitals from December 30, 2019 to January 19, 2020 (1), and the age-stratified population sizes in each of the 13 Wuhan districts (2), which we scaled by the growth in overall Wuhan population between 2010 and 2019.

Appendix: Using the COVID-19 to influenza ratio to estimate Wuhan cases

Method

The age-specific risks of COVID-19 and influenza virus infection are assumed to be identical in all 13 central districts of Wuhan. Therefore, the ratio of COVID-19 to influenza adult outpatients (r) estimated from the subset of outpatients sampled in ref. (1) can be used to estimate the number of COVID-19 infections across all of central Wuhan (Figure A.1).

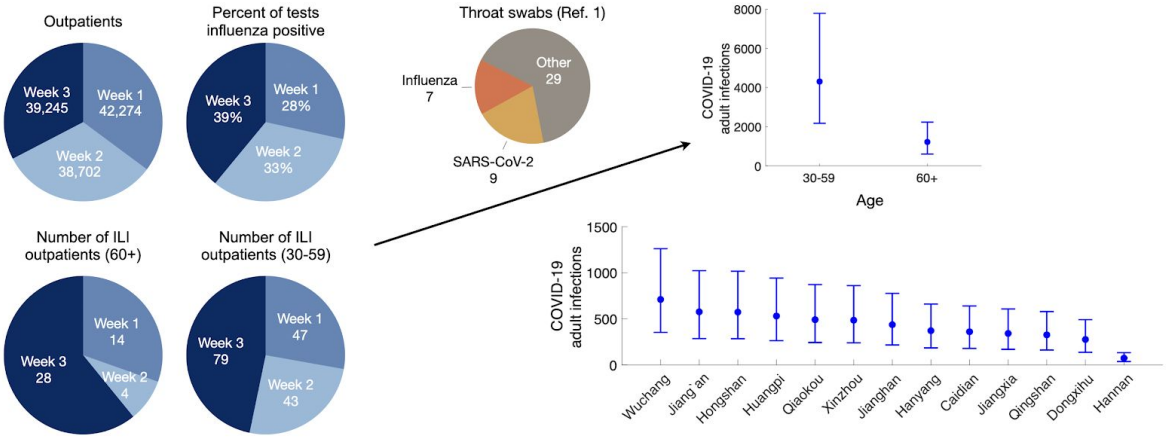

Figure A.1. Estimating adult COVID-19 infections based on the ratio between patients retrospectively testing positive for COVID-19 and influenza in two hospitals in Central Wuhan from December 30, 2019 to January 19, 2020. First we use influenza surveillance data (number of outpatients, percent positive influenza tests, and number of ILI outpatients reported for the Wuhan region by the Chinese CDC) to estimate the proportion of adult outpatients (all cause) testing positive for influenza from December 30, 2019 to January 19, 2020 (left graphs). Second, we estimate the ratio of COVID-19 positive to influenza positive patients among adult outpatients with ILI, based on a recent retrospective study in two Wuhan hospitals (1.28 [95% CI: 0.56-3.26]) (1). We then estimate the number of symptomatic COVID-19 infections among adults across Wuhan during this time period based on the proportion of influenza positive outpatients and the ratio of COVID-19 to influenza positive outpatients, using Monte Carlo sampling to incorporate uncertainty in our estimates of both quantities (upper right). Finally, we estimate the age-specific COVID-19 adult infections for the 13 central districts in central Wuhan based on the district level population sizes for each age group. Given that the nine detected COVID-19 cases lived in six different districts of Wuhan in ref. (1), we assumed that risk was uniform across all 13 districts during the 21-day time period.

Estimating COVID-19 adult infections in Wuhan

Appendix: Using the COVID-19 to influenza ratio to estimate Wuhan cases

We take a Bayesian approach to estimating the number of COVID-19 infections in district d among adult in age range a (30-59 or 60+ years) during the focal nine-day period τ ($H_{d,a,\tau}$). We first estimate the ratio (r) of COVID-19 cases to influenza cases among adult from the retrospective sample (1) using the Katz-log method (6), and then use this to predict the outpatients in each district and age range by assuming:

$$H_{d,a,\tau} \mid N_d, \Theta_\tau^a, \Omega_\tau, \Phi_\tau, r \sim \mathrm{B}\left(N_d, \frac{\Theta_\tau^a}{\Omega_\tau} \cdot \Phi_\tau \cdot r\right)$$

where N_d is the number of people of all ages in district d; Θ_τ^a is the number of ILI outpatients in age group a in over a period of time τ; Ω_τ is the number of all cause outpatients of all ages in Wuhan over a period of time τ; Φ_τ is the percent of influenza tests that are positive in the South Provinces of China during time period τ; the r is the ratio of adult COVID-19 outpatients to adult influenza outpatients. In other words, we multiply the number of influenza-related outpatients by the ratio ILI outpatients that have COVID-19 versus influenza, both estimated for Wuhan from December 30, 2019 to January 19, 2020.

We take a Bayesian approach to derive r as the following posterior distribution. Let N denote the total number of adults in the sample. Let x_c and x_f denote the observed number of adults with COVID-19 and influenza, respectively. Then

$$x_c \mid p_c, N \sim \mathrm{B}(N, p_c) \text{ and } x_f \mid p_f, N \sim \mathrm{B}(N, p_f).$$

If we assume uninformative priors on p_c and p_f,

$$p_c \sim \mathrm{Beta}(1,1) \text{ and } p_f \sim \mathrm{Beta}(1,1)$$

then the posterior distributions are known in closed form(7):

$$p_c \mid x_c, N \sim \mathrm{Beta}(1 + x_c, 1 + N - x_c) \sim \mathrm{Beta}(10, 37)$$

$$p_f \mid x_p, N \sim \mathrm{Beta}(1 + x_f, 1 + N - x_f) \sim \mathrm{Beta}(8, 39)$$

We use MCMC to take draws from p_c and p_f, and then calculate $r = p_c/p_f$ to obtain the distribution for r. We thereby estimate that the ratio of COVID-19 to influenza adult

hospitalizations across central Wuhan during December 30, 2019 to January 19, 2020 was 1.28 [95% CI: 0.56-3.26].

Using 1,000 draws from the distribution of r, we estimate $H_{d,a,\tau}$ for each district and age group. For each draw, we also sample from Ω_a, assuming triangular distributions with bounds and mode given in Table S1, and point estimates for all other parameters given in Table S1. We report the means and 95% credible intervals of the resulting posterior predictive distribution for each district.

Estimating COVID-19 adult infections prior to the Wuhan lockdown

To project the number of adult infections in Wuhan prior to the closing on January 23, 2020 (H_{cum}), we assume

$$H_{\text{cum}} = \sum_{i=t_0}^{L} h_0 \cdot 2^{i/T_d}$$

where T_d is the epidemic doubling time, t_0 is the day of the first adult infection in Wuhan, and L corresponds to January 22, 2020 (the day before the Wuhan lockdown). We use our age- and district-stratified estimates for adult COVID-19 infections for December 30, 2019 to January 19, 2020 to estimate this quantity, under the assumption that the values reflect cumulative incident infections during that nine-day period (Figure A.2).

We use Monte Carlo sampling to incorporate the uncertainty in both the epidemic doubling rate in Wuhan during this period (4) and adult infections from December 30, 2019 to January 19, 2020 ($H_{a,d,\tau}$). We take draws from the distribution of $H_{a,d,\tau}$ and T_d (summarized in Table S1) to estimate the time since the first adult infection by

$$\delta = T_D \left(\log_2 \left(\frac{H_\tau}{\sum_{i=0}^{20} 2^{i/T_d}} \right) \right).$$

That is, the estimated date of the first COVID-19 infection in Wuhan (t_0) is δ days prior to December 30, 2019. We then estimate H_{cum} according to the equation above to project the cumulative COVID-19 adult infections preceding the Wuhan lockdown.

Appendix: Using the COVID-19 to influenza ratio to estimate Wuhan cases

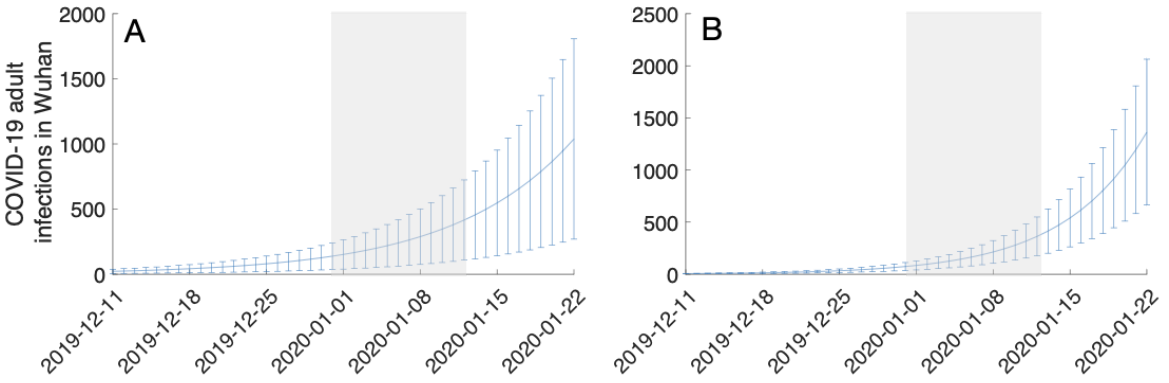

Figure A.2. Estimating the number of symptomatic COVID-19 in adults across the 13 central districts of Wuhan prior to January 23, 2020. We assume an epidemic doubling time of either (A) 7.3 [95% CrI: 6.3–9.7] days or (B) 5.2 [95% CrI: 4.6–6.1] days and that the numbers of adult COVID-19 infections estimated for the 13 districts (Table S1) are equal to the sum of the daily number of incident infections from December 30, 2019 to January 19, 2020. Using an exponential model of epidemic growth we estimate that the first adult COVID-19 infection occurred on (A) November 9, 2019 [95% CI: October 25 - November 22, 2019] or (B) November 29, 2019 [95% CI: November 22 - December 6, 2019], and then project the daily adult COVID-19 infections until January 23, 2020. Gray shading indicates the time period of our initial estimates.

Appendix: Using the COVID-19 to influenza ratio to estimate Wuhan cases

References

1. Kong W-H, Li Y, Peng M-W, Kong D-G, Yang X-B, Wang L, et al. SARS-CoV-2 detection in patients with influenza-like illness. Nat Microbiol [Internet]. 2020 Apr 7; Available from: http://dx.doi.org/10.1038/s41564-020-0713-1

2. CNBo S. The Sixth National Population Census of the People's Republic of China. China: China National Bureau of Statistics. 2010;

3. Wuhan Statistics Bureau. Wuhan city [Internet]. 2019 [cited 2020 Mar 13]. Available from: http://tjj.wuhan.gov.cn/details.aspx?id=83

4. Zhanwei Du, Lin Wang, Simon Cauchemez, Xiaoke Xu, Xianwen Wang, Benjamin J. Cowling, et al. Risk for Transportation of 2019 Novel Coronavirus Disease from Wuhan to Other Cities in China. Emerging Infectious Disease journal [Internet]. 2020;26(5). Available from: https://wwwnc.cdc.gov/eid/article/26/5/20-0146_article

5. Wu JT, Leung K, Bushman M, Kishore N, Niehus R, de Salazar PM, et al. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. Nat Med [Internet]. 2020 Mar 19; Available from: https://doi.org/10.1038/s41591-020-0822-7

6. Katz D, Baptista J, Azen SP, Pike MC. Obtaining Confidence Intervals for the Risk Ratio in Cohort Studies. Biometrics. 1978;34(3):469–74.

7. Berger JO, Yang R. A catalog of noninformative priors. ISDS Discussion Paper, Duke Univ; 1997.