Title

Trends and prediction in daily incidence of novel coronavirus infection in China, Hubei Province and Wuhan City: an application of Farr's law.

Journal

American Journal of Translational Research, 12(4)

ISSN

1943-8141

Authors

Xu, Jie
Cheng, Yajiao
Yuan, Xiaoling
et al.

Publication Date

2020

Peer reviewed

Am J Transl Res 2020;12(4):1355-1361
www.ajtr.org /ISSN:1943-8141/AJTR0109790

Original Article

Trends and prediction in daily incidence of novel coronavirus infection in China, Hubei Province and Wuhan City: an application of Farr's law

Jie Xu[1*], Yajiao Cheng[1*], Xiaoling Yuan[1], Wei Vivian Li[2], Lanjing Zhang[3,4,5,6]

[1]Department of Infectious Disease, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; [2]Department of Biostatistics and Epidemiology, Rutgers School of Public Health, Piscataway, NJ, USA; [3]Department of Pathology, Princeton Medical Center, Plainsboro, NJ, USA; [4]Department of Biological Sciences, Rutgers University Newark, NJ, USA; [5]Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, USA; [6]Department of Chemical Biology, Rutgers Ernest Mario School of Pharmacy, Piscataway, NJ, USA. [*]Equal contributors and co-first authors.

Received February 24, 2020; Accepted March 16, 2020; Epub April 15, 2020; Published April 30, 2020

Abstract: Background: The recent outbreak of novel coronavirus (2019-nCoV) has infected tens of thousands of patients in China. Studies have forecasted future trends of the incidence of 2019-nCoV infection, but appeared unsuccessful. Farr's law is a classic epidemiology theory/practice for predicting epidemics. Therefore, we used and validated a model based on Farr's law to predict the daily-incidence of 2019-nCoV infection in China and 2 regions of high-incidence. Methods: We extracted the 2019-nCoV incidence data of China, Hubei Province and Wuhan City from websites of the Chinese and Hubei health commissions. A model based on Farr's law was developed using the data available on Feb. 8, 2020, and used to predict daily-incidence of 2019-nCoV infection in China, Hubei Province and Wuhan City afterward. Results: We observed 50,995 (37,001 on or before Feb. 8) incident cases in China from January 16 to February 15, 2020. The daily-incidence has peaked in China, Hubei Providence and Wuhan City, but with different downward slopes. If no major changes occur, our model shows that the daily-incidence of 2019-nCoV will drop to single-digit by February 25 for China and Hubei Province, but by March 8 for Wuhan city. However, predicted 75% confidence intervals of daily-incidence in all 3 regions of interest had an upward trend. The predicted trends overall match the prospectively-collected data, confirming usefulness of these models. Conclusions: This study shows the daily-incidence of 2019-nCoV in China, Hubei Province and Wuhan City has reached the peak and was decreasing. However, there is a possibility of upward trend.

Keywords: Trend, incidence, novel coronavirus, China, pandemic

Introduction

The recent outbreak of the novel coronavirus (2019-nCoV) in China has infected about tens of thousands of patients, and claimed thousands of lives [1-5]. The virus and disease were recently renamed as Severe acute respiratory syndrome-related coronavirus-2 (SARS-CoV-2) and Coronavirus disease 2019 (COVID-19), respectively [6, 7]. Studies have forecasted future trends of the infection incidence of coronavirus [2], but appeared unsuccessful, likely owing to the recently-implemented aggressive interventions [8].

Farr's law is a classic epidemiology theory/practice for predicting epidemics, and has been successfully used in predicting mortality and incidence of various diseases [9-13]. However, to our knowledge, Farr's law has not been applied to predicting daily incidence of 2019-nCoV. Therefore, we used and validated a model based on Farr's law to predict the daily-incidence of 2019-nCoV infection in China and 2 regions of high-incidence.

Methods

We extracted the 2019-nCoV daily incidence data of China, Hubei Province and Wuhan City from websites of the Chinese and Hubei health commissions, respectively [4, 5]. The model was developed based on the data available on Feb. 8, 2020, and the predictions were com-

pared with the prospectively collected data afterward. Following the terminologies in previous work [10, 14], the ratio 1 was the ratio of a given day's incidence over that of the day before. Dividing the ratios 1 of a given day and the day before that day resulted in the ratio 2. In most of the previous epidemiologic studies, ratio 2 remains constant, and was rarely examined for its normality. We then examined the possible normality of the ratio 2 using Skewness and Kurtosis tests after various data transformations. After identifying the best-fit data transformation format, we examined the potential linear and log-linear associations of incidence with time. Assuming the future ratio 1's and 2's would be the same as the mean of the last five days, and ratio 2's the same as the mean of the last 10 days, we predicted the ratio 1's and subsequently daily incidence. Statistical analyses were conducted using Stata (version 15, Stata Corp, Collage Station, TX) and Joinpoint (NCI). All *P* values were two-sided and a $P<0.05$ was considered statistically significant.

Results

The daily-incidence was available for China from January 16 to February 23, 2020, for Wuhan and Hubei from January 11 to February 23, 2020. Among various data transformation formats, only natural-logarithm-transformed incidence data were of normal distribution. We did not identify any significant log-linear or linear association of the ratio 2 with time (day). The median/mean of ratio 2 before February 8, 2020 was 0.944 (quartiles 0.886-1.051) for China, 0.982 (quartiles 0.744-1.317) for Wuhan and 0.948 (quartiles 0.768-1.207) for Hubei Province. The future incidence was predicted based on the mean or quartiles of ratio 2 and subsequently inferred ratio 1. The daily-incidence significantly decreased after February 8 in all regions, and would continue decreasing until reaching zero on February 25. Given the lower quartile (25%) of ratio 2, the incidence may reach zero by February 20, while based on the upper quartile (75%) of ratio 2, the daily-incidence will keep increasing and reach 3000 on February 25 (**Figure 1**). Hubei Province and Wuhan City overall had a similar downward trend (**Table 1**). The prospectively collected daily-incidence after February 8 appeared to fall in the predicted quartiles in all 3 selected regions. Specifically, the predicted incidence in Wuhan City was very close to the prospectively collected data, while the predicted incidence in China and Hubei province was slightly lower than the reported data.

Discussion

Using Farr's law and normality analysis on the ratio 2, we estimated the quartiles of 2019-nCoV daily incidence in China, Hubei Province and Wuhan City, respectively, and predicted the daily-incidence of these regions. The predicted and prospectively collected daily incidence was similar in Wuhan and slightly different in China and Hubei Province, while the prospectively collected data were still within the quartiles of predicted daily incidence. These predictive models primarily based on Farr's law were therefore reasonably accurate.

The predicted daily incidence after February 13, 2020 was lower than the reported data in China and Hubei Providence. It was likely attributable to the changes in diagnostic criteria. We also speculate that the underpredicted incidence in China and Hubei province was attributable to the much slower downward trend in Wuhan City. A possible reason is that, due to larger denominator in China and Hubei Province, those regions' ratio 2's were more stable than that of Wuhan City. However, the Wuhan City is the largest contributor to the new cases in Hubei Province and China. Its smaller downward trends thus would determine or de-accelerate the downward trends in daily incidence of China and Hubei Province, particularly on later days. Future modeling in overall trends thus should consider or be adjusted for the major contributor's trends. It is also noteworthy, despite the underestimation, the reported daily incidence in all three regions was all within estimated 25% to 75% quartiles. It suggests that our model was reasonably accurate and acceptable.

Our results imply that the daily-incidence has reached its peak and will likely decrease continuously in China, Hubei Province and Wuhan City, while the predicted 75% CI of their daily-incidence had an upward trend. Therefore, governments, healthcare providers and residents should be cautiously optimistic, and recognize the potential upward trend of daily-incidence. A rapid growth in daily-incidence may occur

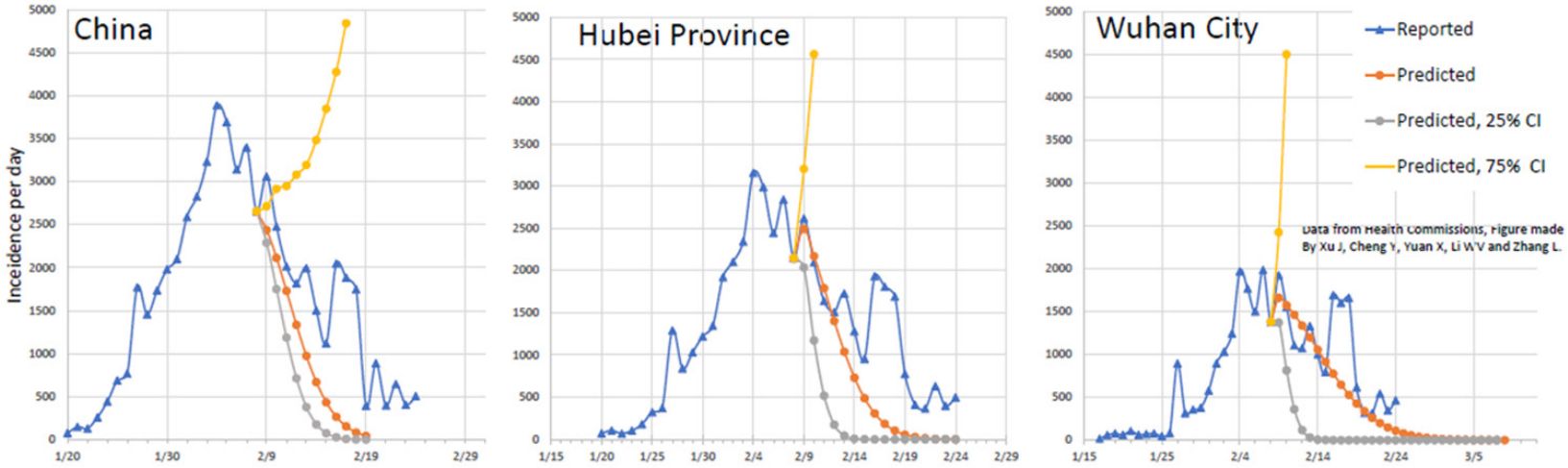

Figure 1. The reported and predicted daily incidence of 2019 novel coronavirus infection in China, Hubei Province and Wuhan City. The predicted daily incidence would reach zero on February 25 for China and Hubei Province, but on March 8 for Wuhan City, due to different ratio 2's. CI, confidence interval.

Table 1. Daily incidence of diagnosed novel coronavirus cases in China and selected areas, Jan. 17 to Feb. 11, 2020, with prediction to March 8, 2020 By Xu J, Cheng Y, Yuan X, Li WV and Zhang L

Date	China				Hubei province				Wuhan City, Hubei			
	Incidence (per day)	Predicted	Ratio 1[a]	Ratio 2[b]	Incidence (per day)	Predicted	Ratio 1[a]	Ratio 2[b]	Incidence (per day)	Predicted	Ratio 1[a]	Ratio 2[b]
1/17/2020									17			
1/18/2020									59		3.471	
1/19/2020									77		1.305	0.376
1/20/2020	77				72				60		0.779	0.597
1/21/2020	149		1.935		105		1.458		105		1.750	2.246
1/22/2020	131		0.879	0.454	69		0.657	0.451	62		0.590	0.337
1/23/2020	259		1.977	2.249	105		1.522	2.316	70		1.129	1.912
1/24/2020	444		1.714	0.867	180		1.714	1.127	77		1.100	0.974
1/25/2020	688		1.550	0.904	323		1.794	1.047	46		0.597	0.543
1/26/2020	769		1.118	0.721	371		1.149	0.640	80		1.739	2.911
1/27/2020	1771		2.303	2.060	1291		3.480	3.030	892		11.150	6.411
1/28/2020	1459		0.824	0.358	840		0.651	0.187	315		0.353	0.032
1/29/2020	1737		1.191	1.445	1032		1.229	1.888	356		1.130	3.200
1/30/2020	1982		1.141	0.958	1220		1.182	0.962	378		1.062	0.940
1/31/2020	2102		1.061	0.929	1347		1.104	0.934	576		1.524	1.435
2/1/2020	2590		1.232	1.162	1921		1.426	1.292	894		1.552	1.019
2/2/2020	2829		1.092	0.886	2103		1.095	0.768	1033		1.155	0.744
2/3/2020	3235		1.144	1.047	2345		1.115	1.019	1242		1.202	1.041
2/4/2020	3887		1.202	1.051	3156		1.346	1.207	1967		1.584	1.317
2/5/2020	3694		0.950	0.791	2987		0.946	0.703	1766		0.898	0.567
2/6/2020	3143		0.851	0.895	2447		0.819	0.866	1501		0.850	0.947
2/7/2020	3399		1.081	1.271	2841		1.161	1.417	1985		1.322	1.556
2/8/2020[c]	2656		0.781	0.723	2147		0.756	0.651	1379		0.695	0.525
2/9/2020	3062	2439	0.918	0.944	2618	2494	0.965	0.982	1921	1659	0.965	0.982
2/10/2020	2478	2115	0.867	0.944	2097	2171	0.947	0.982	1552	1571	0.947	0.982
2/11/2020	2015	1730	0.818	0.944	1638	1792	0.930	0.982	1104	1462	0.930	0.982
2/12/2020	1820	1336	0.772	0.944	1508	1402	0.913	0.982	1072	1335	0.913	0.982
2/13/2020	1995	974	0.729	0.944	1728	1040	0.897	0.982	1330	1198	0.897	0.982
2/14/2020	1503	670	0.688	0.944	1282	731	0.881	0.982	1001	1055	0.881	0.982
2/15/2020	1121	435	0.649	0.944	955	487	0.865	0.982	793[d]	912	0.865	0.982
2/16/2020	2048	266	0.613	0.944	1933	308	0.849	0.982	1690	775	0.849	0.982

2/17/2020	1886	154	0.578	0.944	1807	185	0.834	0.982	1600	646	0.834	0.982
2/18/2020	1749	84	0.546	0.944	1693	105	0.819	0.982	1660	529	0.819	0.982
2/19/2020	394	43	0.515	0.944	775	56	0.804	0.982	615	425	0.804	0.982
2/20/2020	889	21	0.486	0.944	411	29	0.790	0.982	319	336	0.790	0.982
2/21/2020	397	10	0.459	0.944	366	14	0.775	0.982	314	260	0.775	0.982
2/22/2020	648	4	0.433	0.944	630	6	0.761	0.982	541	198	0.761	0.982
2/23/2020	409	2	0.409	0.944	398	3	0.748	0.982	348	148	0.748	0.982
2/24/2020	508	1	0.386	0.944	499	1	0.734	0.982	464	109	0.734	0.982
2/25/2020		0	0.364	0.944		0	0.721	0.982		78	0.721	0.982
2/26/2020		0	0.344	0.944		0	0.708	0.982		56	0.708	0.982
2/27/2020		0	0.324	0.944		0	0.695	0.982		39	0.695	0.982
2/28/2020		0	0.306	0.944		0	0.683	0.982		26	0.683	0.982
2/29/2020		0	0.289	0.944						18	0.670	0.982
3/1/2020		0	0.273	0.944						12	0.658	0.982
3/2/2020		0	0.257	0.944						8	0.646	0.982
3/3/2020		0	0.243	0.944						5	0.635	0.982
3/4/2020		0	0.229	0.944						3	0.623	0.982
3/5/2020		0	0.216	0.944						2	0.612	0.982
3/6/2020		0	0.204	0.944						1	0.601	0.982
3/7/2020		0	0.193	0.944						1	0.590	0.982
3/8/2020		0	0.182	0.944						0	0.579	0.982

Note: the case number may be changed due to clerical error or change of the diagnosis in early days. a. Proportional change in the daily from the day before, and was calculated as incidental cases in the current day/those in the day before. b. The proportional change was calculated as ratio 1 in the current day/ratio 1 in the day before. c. The completion date of the modelling with Farr's law. All daily incidence since Feb. 8, 2020 was prospectively collected, and compared with the predicted daily incidence. d. Not directly available, but calculated by subtracting estimated clinical diagnosis cases (85% of Hubei Province based on prior proportions) from the total reported cases in Wuhan city on Feb. 15, 2020.

when the residents of high-prevalence regions, such as Hubei Province, return to their workplaces in low-prevalence regions at the end of Chinese New-Year holidays or emergency state of fighting the virus' outbreak. Extreme caution thus should be used to prevent such an upward trend.

Several strengths of this study are noteworthy. First, this study is valuable because its comparison and validation with prospectively collected data over 16 data-points (days). Second, the large numbers of cases in China, Hubei Providence and Wuhan City provide acceptable statistically power for predicting future trends. Third, we relied on approximately 23 data points to estimate the trends, which appeared to peak on February 1, 2020. Thus, the sufficient number of data points and sufficient length of observation time helped increase the prediction accuracy. Finally, the normality of ratio 2's was analyzed and provides more rigorous evaluation of this critical parameter. As described before [14], Farr's law uses similar principle of other models used in infectious diseases such as IDEA and others. One limitation of the Farr's law was the assumption that the ratio 2 was constant without rigorous statistical analysis, although often so in real world [11-13]. Our analysis and validation of the normality distribution for ratio 2 is a novel yet important step to quantify ratio 2 more precisely.

A limitation of this study is possible oversimplification of the disease's natural history by this model. However, our prospectively-collected data and recent works proved the usefulness of the Farr's law [14]. This simple yet powerful method has also successfully predicted the trends in incidence of opioid overdose in the U.S. [10]. In addition, our model was supported by an acceptable prediction accuracy as shown by comparing predicted and prospectively collected data. Additional limitations were the ignorance of misclassification in diagnosis and changes of the diagnostic criteria. These inherent limitations should be addressed in future studies with individual patient data, which however are not yet available.

In summary, our application of Farr's law on daily incidence of 2019-nCoV in China, Hubei Province and Wuhan City appears successful, which was confirmed by the good concordance of the predicted daily incidence and prospectively collected data. Our model, with reasonable confidence, suggests that the 2019-nCoV daily incidence in China, Hubei Province and Wuhan City has peaked in early February 2020, will likely continue decreasing in the coming weeks, and may reach single digit by mid-March, 2020 if additional outbreaks or changes in practice occur.

Acknowledgements

We salute and thank our teachers, classmates and colleagues, fellow residents and public servants, who fiercely and tirelessly fought against the novel coronavirus outbreak in Wuhan and other parts of China. The data will be regularly updated at https://github.com/thezhanglab/coronavirus, as new prospectively collected incidence data become available.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Lanjing Zhang, Department of Pathology, Princeton Medical Center, 1 Plainsboro Road, Plainsboro, NJ 08563, USA. E-mail: lanjing.zhang@rutgers.edu; Dr. Jie Xu, Department of Infectious Disease, Shanghai Ninth People's Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China. E-mail: dr.xu@aliyun.com

References

[1] Zhang RQ, Liu H, Li FY, Zhang B, Liu QL, Li XW and Luo LM. Transmission and epidemiological characteristics of Novel Coronavirus (2019-nCoV) Pneumonia (NCP): preliminary evidence obtained in comparison with 2003-SARS. MedRxiv 2020.

[2] Wu JT, Leung K and Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet 2020; 395: 689-697.

[3] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Li M, Tu W, Chen C, Jin L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TTY, Wu JT, Gao GF, Cowling BJ, Yang B, Leung GM and Feng Z. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med 2020; 382: 1199-1207.

[4] Wuhan Municipal Health Commission N (Feb. 23, 2020). "Briefings on the viral disease [Chinese]". Retrieved Feb. 23, 2020, from http://wjw.wuhan.gov.cn/front/web/list3rd/no/802.
[5] National Health Commission of China A (2020). "Briefings on the viral outbreak [Chinese]". Retrieved Feb. 24, 2020, from http://www.nhc.gov.cn/xcs/yqtb/list_gzbd.shtml.
[6] Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol 2020; 5: 536-544.
[7] WHO (2020). "Coronavirus disease (COVID-19) outbreak". Retrieved February 23, 2020, from https://web.archive.org/web/20200223043-035/https://www.who.int/emergencies/diseases/novel-coronavirus-2019.
[8] News B (2020, Jan. 23). "Coronavirus: Wuhan shuts public transport over outbreak". Retrieved Feb. 9, 2020, from https://web.archive.org/save/https://www.bbc.com/news/world-asia-china-51215348.
[9] Bregman DJ and Langmuir AD. Farr's law applied to AIDS projections. JAMA 1990; 263: 1522-1525.
[10] Darakjy S, Brady JE, DiMaggio CJ and Li G. Applying Farr's Law to project the drug overdose mortality epidemic in the United States. Inj Epidemiol 2014; 1: 31.
[11] Lilienfeld DE. Celebration: William Farr (1807-1883)-an appreciation on the 200th anniversary of his birth. Int J Epidemiol 2007; 36: 985-987.
[12] Goudsmit J. Alternative view on AIDS. Lancet 1992; 339: 1289-1290.
[13] Brownlee J. Studies in the meaning and relationships of birth and death rates: II. density of population and death rate (Farr's Law). J Hyg (Lond) 1915; 15: 11-16.
[14] Santillana M, Tuite A, Nasserie T, Fine P, Champredon D, Chindelevitch L, Dushoff J and Fisman D. Relatedness of the incidence decay with exponential adjustment (IDEA) model, "Farr's law" and SIR compartmental difference equation models. Infect Dis Model 2018; 3: 1-12.