

Arrested Development:

The number of Wuhan cases of COVID-19 with onset in 2019

author: Gilles Demaneuf, engineer (ECP) and Data Scientist, email: gilles@demaneuf.com

acknowledgements: I wish to thank the many people within DRASTIC and outside of it who offered constructive advice, starting with Adrian Jones and Steve Massey.

Main Findings:

This study demonstrates that the official COVID-19 numbers for 2019 as given in the final WHO report (100 confirmed cases and 74 diagnosed ones) are in fact far from the numbers recorded in the official database of cases at the end of February 2020. Thanks to Chinese peer-reviewed papers that draw their data from that official database, we can show that there were 146 confirmed cases with onset up to 29 Dec 2019 for Wuhan and neighbouring cities, identified as of 25 Feb 2020, with an additional 92 clinically diagnosed cases for Dec 2019, known as of 27 Feb 2020. These papers also show that out of the Dec 2019 onset cases, 33 had died by 24 Feb 2020, including 29 confirmed ones.

With the addition of 19 confirmed cases for the last two days of Dec 2019, which we draw from the final WHO report, we show that there were at least 165 confirmed cases for Dec 2019. We further show that there were necessarily between 247 and 260 confirmed plus diagnosed cases in the official database at around 27 Feb 2020, with the expected number towards the higher end. That precisely matches the leak of 257 cases published in the South China Morning Post in early March 2020, which additionally included 9 cases from November 2019 cases (for a total of 266).

The authors of one peer-reviewed paper (Wu-Chun Cao et al.) worked with the Chinese CDC to extract their data, while the main author had previous high-profile experience with retrospective search and database cleaning after SARS. The authors of a second peer-reviewed paper (Chuanhua et al.) happen to be a team well known to have worked extensively with the Chinese CDC to do a retrospective search and clean the official database of COVID-19 cases in February 2020, and the main author of that paper had previously mentioned potential November cases in the Chinese medical press.

Other official sources of data, and an additional peer-reviewed paper, allow us to trace a rather detailed course of the number of cases in the official database. We observe that the progression in the number of cases stopped soon after the gag order imposed on the Chinese CDC on 25 Feb 2020 by the National Health Commission, immediately after the first WHO mission (16-24 Feb) left China. We also conclude that the number of confirmed cases eventually given to the final WHO mission (100) is more consistent with the count in the official database around 15 Feb 2020, instead of what was known and recorded in that database by end February.

The demonstrated exclusion in the final WHO report of a large number of recorded laboratory-confirmed Dec 2019 onset cases is a very serious issue. Additionally, the precise match between our count of Dec 2019 onset cases at end February 2020, and the South China Morning Post leak of government data two weeks later, points to the real possibility that November 2019 cases (either confirmed or diagnosed) existed in that official database in early March 2020.

We therefore urge extreme caution in the interpretation of any paper or analysis based on the 'official' dataset for COVID-19 cases with 2019 onset presented to the WHO.

Content

Findings:

1:	146 lab-confirmed onset cases up to 29 Dec 2019, as of 27 Feb 2020	7
2:	92 clinically-diagnosed Dec 2019 onset cases, as of 24 Feb 2020	7
3:	247 to 260 'confirmed + diagnosed' Dec 2019 onset cases, as of 27 Feb 2020	10
4:	Comparison to SCMP leak and its 257 Dec 2019 cases	12
5:	Comparison to official WHO numbers of 100 confirmed and 74 diagnosed	15
6:	33 deaths from Dec 2019 onset cases by 24 Feb 2020, 29 lab-confirmed	16
7:	Pattern of gradual increase of Dec 2019 onset numbers, from other sources	19
8:	Summary of reported number of Dec 2019 onset cases, from all sources reviewed	22
9:	Issues with the evolution over time of 'confirmed' 2019 onset numbers	24
10:	Analysis of the evolution over time of 'diagnosed' 2019 onset numbers	27
11:	Issues with the count of 'suspected' 2019 onset numbers	28
12:	Issues with the retrospective testing of some 'suspected' cases in January 2020	30
13:	Official and final WHO number of Dec 2019 onset cases is as of around 15 Feb 2020	33
14:	Retrospective search for 2019 cases in February 2020	37
15:	Consequences for our understanding of the outbreak origins	42

Boxes:

1:	Key Definitions	5
2:	Examining the jumps in the times series	9
3:	Wuhan vs. Hubei for 2019 onset cases	11
4:	The NNDRS database	14
5:	Pathways for 2019 onset cases by reporting date	18
6:	Total December 2019 cases as per final WHO report	26
7:	The unreported 'suspected' cases on 11 Jan 2020	32
8:	The gag orders	35

Finding 1:

A peer-reviewed paper shows 146 lab-confirmed cases with onset date from 8 Dec to 29 Dec 2019 in Wuhan, using the official CDC data to 27 Feb 2020.

We refer to a peer-reviewed paper titled 'Epidemiology of and Risk Factors for COVID-19 Infection among Health Care Workers: A Multi-Centre Comparative Study'¹ published on 29 September 2020 in the 'International Journal of Environmental Research and Public Health', by a team of scientists from the Institute of EcoHealth in Shandong and from the State Key Laboratory of Pathogen and Biosecurity, Beijing.

The paper (thereafter 'Wu-Chun Cao *et al.*') explains that the authors 'collected data of **confirmed** COVID-19 cases in Wuhan from the National Notifiable Infectious Disease Information System² until 27 February 2020', and also '[...] acknowledge the China CDC for their valuable assistance in coordinating data collection'. The paper also uses the standard definition of confirmed cases as per the 5th edition of the official guidelines: 'related epidemiological history and clinical manifestations with one of the following etiological evidences: SARS-CoV-2 nucleic acid detected by specific real-time PCR assay or viral gene sequence homologous to SARS-CoV-2'.

Figure 1. Time interval from symptom onset to diagnosis against calendar time among healthcare workers (HCWs) and non-HCW cases of COVID-19 in Wuhan. The case number in each group was shown over each box in the main figure. From 6 January 2020, the box plot was shown with notches in smaller scale, which represent the 95% confidence intervals of medians of time from symptom onset to diagnosis.

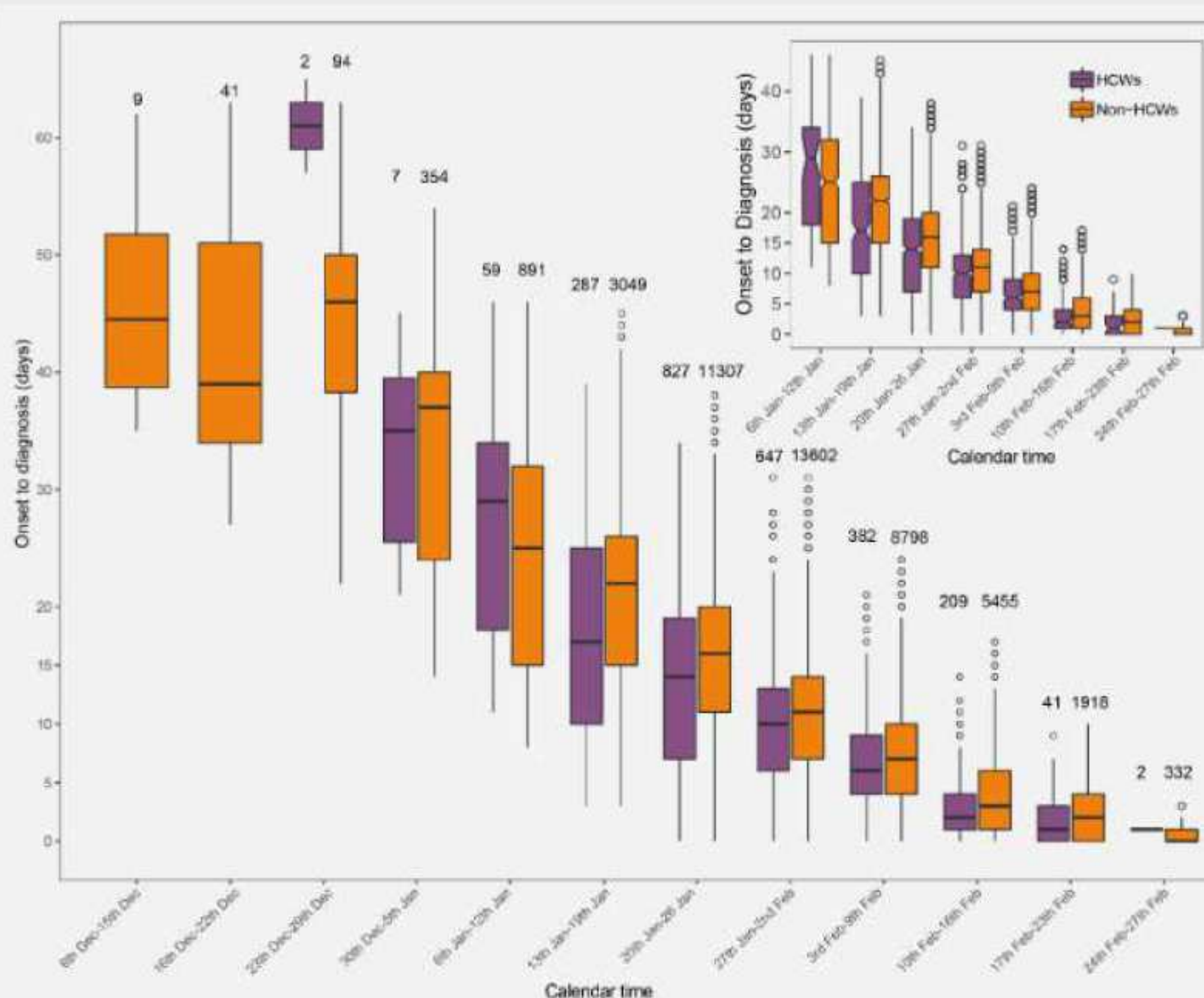


Fig. 1: Figure 1 of 'Epidemiology of and Risk Factors for COVID-19 Infection among Health Care Workers: A Multi-Centre Comparative Study'

¹ Accessible at <https://www.mdpi.com/1660-4601/17/19/7149>.

² We shall use the acronym NNDRS (National Epidemic Network Direct Reporting System) in the paper for that system, in line with the term used in the final WHO report.

The paper has a key figure (Figure 1) which allows us to get a count of laboratory-confirmed³ cases in Wuhan with onset dates from 8 Dec 2012 to 29 Dec 2019. The total is easy to read from the first 4 candles: 9 + 41 + 2 + 94 = 146 lab-confirmed cases up to 29 Dec 2019.

In the same way, we easily extracted from Figure 1 the running totals of confirmed cases⁴ over December 2019. See Table 1:

Onset latest	Running Total Confirmed	HCW	Non-HCW
15 Dec 2019	9		9
22 Dec 2019	50		50
29 Dec 2019	146	2	144
5 Jan 2020	505	7	498
12 Jan 2020	1,448	59	1,389
—	—	—	—
27 Feb 2020	45,852	2	45,850

Table 1: running totals of lab-confirmed Wuhan cases with onset date to 29 Dec 2019.

The main corresponding author, Wu-Chun Cao is a very prominent PLA scientist with the rank of colonel⁵ and is a member of the Scientific Advisory Committee of the Centre for Emerging Infectious Diseases at the Wuhan Institute of Virology.⁶

Amongst his notable papers and many achievements:

- He was the last author of a 2009 paper that retrospectively searched for, and classified, SARS-1 cases from the CDC database and the parallel but unconnected military reporting system, to produce what is now the official and final count for SARS-1 cases in China⁷.
- He is the leading Chinese expert on tick viruses, for the study of which he has received many grants from the National Science Foundation of China.
- He was a speaker with Ralph Baric, Shi Zhengli, David Franz and Yuan Zhiming and other top names at some US-China biosafety conference in 2015.⁸
- He more recently co-authored a well-known Nature paper on SARS-CoV-2-related virus in Malayan pangolins with Edward Holmes.⁹

In other words Wu-Chun Cao is a top Chinese scientist, with excellent national and international connections and a previous experience with gathering and cleaning epidemiological data, working in the process with the Chinese CDC to do so.

³ As noted in 'Box 1, Case Definitions', 'laboratory-confirmed' and 'confirmed' mean the same thing and are used interchangeably in the key papers cited and the WHO reports. 'clinically-diagnosed' and 'diagnosed' are also commonly used interchangeably.

⁴ See tab 'WuChunCao_paper' of the [supplementary spreadsheet](#) for more data extraction from that graph.

⁵ See <https://archive.ph/KrkGu>.

⁶ See <https://thenationalpulse.com/wp-content/uploads/Screen-Shot-2021-05-17-at-11.46.47-AM-800x537.png>

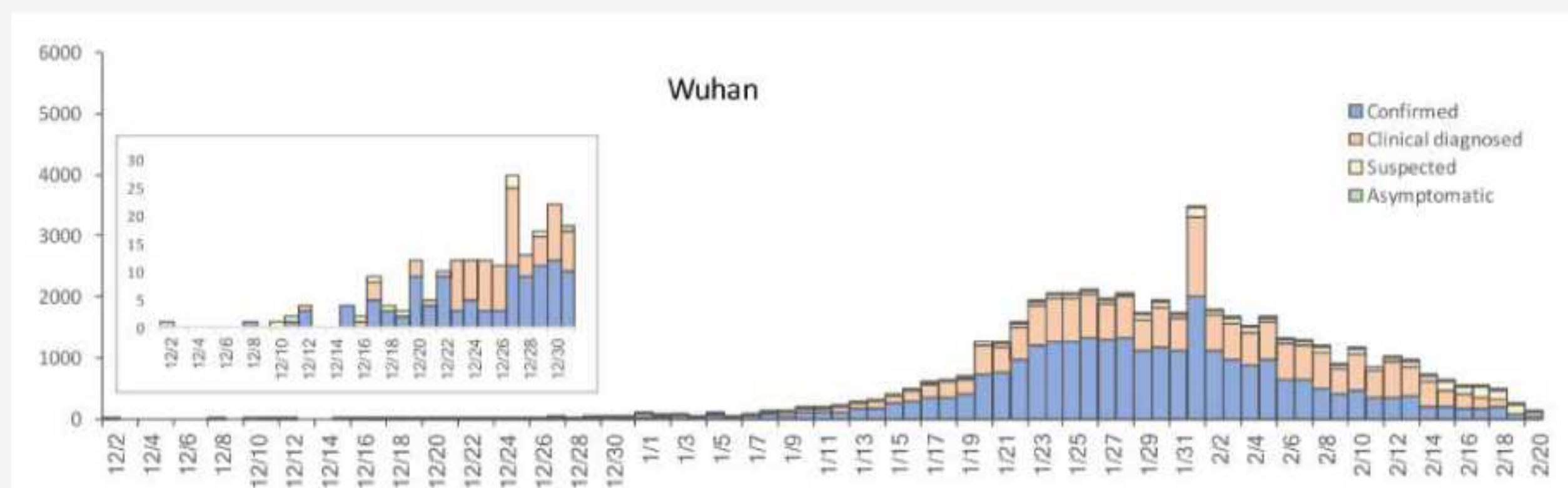
⁷ 'The SARS epidemic in mainland China: bringing together all epidemiological data', available at <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7169858/>.

⁸ See <https://archive.ph/YJOWx>.

⁹ 'Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins', Nature, July 2020, available at <https://pubmed.ncbi.nlm.nih.gov/32218527/>, also mentioned by Edward Holmes in a NY Times interview: <https://archive.ph/RP0Fg>

Box 1: Key Definitions¹⁰

Taking a typical reporting example, let's clarify some essential definitions.



Wuhan panel of Figure 2 of the report of the first WHO mission (Feb 2020), data as of 20 Feb¹¹

What are 'confirmed' cases?

The SARS-CoV-2 genome was published only on 10 Jan 2020, so RT-PCR testing started progressively on 11 Jan. Before that cases were either 'suspected' or, for a few cases, "confirmed" via Next Generation Sequencing showing a SARS-like match. Once RT-PCR tests were available, 'suspected' cases that returned a positive test were considered 'confirmed'. Cases may also be confirmed via full genome sequencing with a match against the newly published SARS-CoV-2 genome; this however was more time consuming and thus less common once RT-PCR tests were available.

What are 'diagnosed' cases?

'Diagnosed' relied on the rather reliable identification of COVID-19 cases via lung imagery, especially CT-scans, by experienced medical practitioners.

Officially because there were still too few RT-PCR test kits available for a while in Hubei, 'diagnosed' cases were introduced in Hubei on 5 Feb to 2020 be able to handle infected patients properly (isolation, etc), until the category was closed for new entries on 17 Feb, once hopefully systematic testing of patients had become easier.

Beyond the official explanations, it is also important to note that initially RT-PCR tests resulted in quite a few false negatives, to the point where Dr. Zhang, a radiologist at Zhongnan Hospital (a top tier research hospital in Wuhan) who had served on the frontlines of SARS in 2003, took to her social messaging account on 3 Feb 2020 to call for the additional use of lung imaging in diagnosing COVID-19, beyond RT-PCR tests.¹²

Also on 6 Feb 2020, the day after the publication of the revised case definition introducing the 'diagnosed' case definition, doctors from the radiology department of Wuhan Jinyintan hospital published a preprint about the importance of using CT-scans for COVID-19 diagnostic¹³, while mentioning the false-negative limitations of RT-PCR tests.¹⁴ That study was further confirmed by a paper published online on 26 Feb, from the Tongji

¹⁰ For more details, see for instance this [Lancet paper](#) (especially its [Supplementary Material](#)) and the Annex E3 of the [final WHO report](#). Tab 'CaseDef' of our [supplementary spreadsheet](#) has some additional references.

¹¹ Report of first WHO mission, published Feb 2020 and available at [https://www.who.int/publications/i/item/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-\(covid-19\)](https://www.who.int/publications/i/item/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19))

¹² The episode is well described in the following WSJ article:

https://www.wsj.com/articles/if-we-fail-what-happens-to-you-all-one-doctors-life-on-the-coronavirus-front-lines-11583344415?mod=article_inline&tesla=y

¹³ 'Radiologic Features of Patients with 2019-nCoV Infection', Chuansheng Zheng et al., 6 Feb 2020, available at <http://www.mlpla.mil.cn/dzfw/yhij/xgzl/202002/P020200207581581112243.pdf>.

¹⁴ Incidentally Jinyintan hospital had seen one of the earliest applications of RT-PCR testing once the genome was published, resulting in the Wuhan Health Commission dropping 18 suspected cases on 11 Jan from the 59 patients regrouped in Jinyintan hospital.

Hospital and Tongji Medical College (Wuhan), which reviewed 1,014 cases and concluded that detecting cases based on RT-PCR testing on its own was problematic, due to its much lower sensitivity (~60%) compared to CT-scans, so that 'Chest CT may be considered as a primary tool for the current COVID-19 detection in epidemic areas'.¹⁵

What are 'suspected' cases?

'Suspected' cases were essentially a holding pen for further identification, either by lung imagery between 5 and 17 Feb 2020 (on which basis they could move to 'diagnosed'), or directly by RT-PCR or full genome sequencing at any time after 10 Jan 2020 (on which basis they could move to 'confirmed').

During the main retrospective search for 2019 cases in hospital records (from 5 Feb to end Feb 2020) it seems that 'diagnosed' (till 17 Feb) and 'confirmed' (till end Feb) were the main focus, with few (if any) new 'suspected' cases added, while the backlog of 'suspected' cases from before the introduction of the 'diagnosed' category on 5 Feb was being processed.

What are 'asymptomatic' cases?

This is a little used category for lab-confirmed asymptomatic cases, kept separate from 'confirmed' which requires some clinical manifestation. Asymptomatic cases were mentioned in the 'CDC Weekly' update (17 Feb) and the first WHO report (28 Feb) and hardly reported after that.

Asymptomatic cases were most often detected as part of some contact-tracing, or in hospital settings (including testing of Healthcare Workers). In effect very few asymptomatic cases were caught as they were never the focus of any systematic detection, especially during the main retrospective search which focussed on hospital cases. Hence the lack of further reporting of asymptomatic cases does not hurt much since the few asymptomatic cases reported were never representative of their true number.

Why are 'diagnosed' cases still to be found in the counts after the category stopped being applicable, even as far as the final WHO report (March 2021)?

After 17 Feb 2020, the 'diagnosed' category stopped being available for new case reports.

However, once reported between 5 and 17 Feb 200, existing 'diagnosed' cases would have to remain in the statistics if (i) *original* samples from the time of hospitalisation were available but had not been tested yet, (ii) *original* samples from the time of hospitalisation were not available and *retrospective* sampling and testing had not been done, and (presumably) (iii) *retrospective* sampling and testing was eventually done, but returned a positive (IgG+), which unfortunately cannot distinguish an infection at the time of hospitalisation from a later infection (possibly asymptomatic).

Why are there no 'suspected' cases mentioned in the later statistics published?

There is a potential issue here, especially given the very different behaviour over time of the 'suspected' case counts compared to 'diagnosed'. See Finding 11.

Important notes:

- 'laboratory-confirmed' and 'confirmed' mean the same thing: RT-PCR or full genome sequencing.
- 'clinically-diagnosed' and 'diagnosed' mean the same thing.
- 'confirmed', 'suspected' and 'diagnosed' all have clinical manifestations, as per the applicable case definitions. In particular no 'asymptomatic' case was ever counted in 'confirmed'.
- The main retrospective search (5 Feb to end Feb 2020) focussed on hospital cases, and would typically not include many mild and moderate cases (not to talk about asymptomatic cases) that never went to hospital.

¹⁵ See 'Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases', Tao Ai *et al.*, online 26 Feb 2020. See <https://archive.ph/oNQb1> and the PDF archived at <https://bit.ly/3eVXJpN>.

Finding 2:

A peer-reviewed paper, from the Chinese team in charge of cleaning and back-populating the national database of cases with the CDC, shows 135 confirmed and 92 diagnosed cases (total 227) for Dec 2019 onset in Wuhan, using the official CDC data to 24 Feb 2020.

We refer to a peer-reviewed paper titled 'The comparison of epidemiological characteristics between confirmed and clinically diagnosed cases with COVID-19 during the early epidemic in Wuhan, China'¹⁶ published on 28 May 2021 in 'Global Health Research and Policy'.

That paper is from a Wuhan University team under the direction of Chuanhua Yu, of the Department of Epidemiology and Biostatistics. Chuanhua Yu was mentioned in late February 2020 in various Chinese media reports¹⁷ as working with his team on a retrospective review of cases data, in order to back-populate the official infectious disease reporting information system (the NNDRS database) which is under the management of the Chinese CDC.

In another of his peer-reviewed papers¹⁸, the role of his team in that retrospective search is clearly laid out:

Our team was among those who participated in obtaining the original uncleaned data from the designated hospitals during the epidemic; cleaning, shortening and analyzing the original data; sending the data to the CDC and the infectious-disease-reporting information system; and working with CDC staff in ensuring the quality of the data. We were given official approvals for monitoring the data of registered patients with severe and critical COVID-19 illness in Wuhan until 25 February 2020 (i.e. the final date of follow-up in this study).

The paper of May 2021 (thereafter 'Chuanhua *et al.*') reveals that, according to the database extract from 24 Feb 2020:

- there were 227 COVID-19 cases in Wuhan with December 2019 onset, as 135 lab-confirmed ones plus 92 clinically-diagnosed ones,
- 33 of these 227 Dec 2019 patients had died by 24 Feb 2020: 29 confirmed ones and 4 diagnosed ones.

These numbers can be read in Tables 1 and 2 of that May 2021 paper¹⁹. Here we shall just show an extract:

Characteristics	Confirmed Cases			Clinically Diagnosed Cases			P
	Death (%)	Not Dead (%)	Case Fatality Rate (%)	Death (%)	Not Dead (%)	Case Fatality Rate (%)	
Total	1564	28322	5.2	257	21703	1.2	<0.001
Date of onset							
Before Dec 31, 2019	29(1.9)	106(0.4)	21.5	4(1.6)	88(0.4)	4.3	<0.001

Table 2: Extract from Table 2 of [The comparison of epidemiological characteristics between confirmed and clinically diagnosed cases with COVID-19 during the early epidemic in Wuhan, China](#)

¹⁶ Accessible at <https://ghrp.biomedcentral.com/articles/10.1186/s41256-021-00200-8>

¹⁷ See for instance the description of his work in

<https://gillesdemaneuf.medium.com/early-cases-of-suspected-covid-19-in-wuhan-feb-20-data-collection-b7740ed1436f> and the interview he gave to the Health Times at end February 2020:

https://docs.google.com/document/d/e/2PACX-1vTQxG822DtqP7IZSjLj751Mrm8Ev7leKsXfjBLsA9KJ0_tbGV6YJAAjujPnwz_YmUQGY1PZUI5LcCI/pub

¹⁸ 'Epidemiological characteristics of patients with severe COVID-19 infection in Wuhan, China: evidence from a retrospective observational study', published online Nov 2020 in International Journal of Epidemiology, available at <https://academic.oup.com/ije/article/49/6/1940/5956328?login=false#authorNotesSectionTitle>

¹⁹ Table S3 breaks these cases by severity, for those cases where the severity was known. In doing so only 157 of the 227 Dec 2019 cases have a known severity and can thus be reported in Table S3.

The count of 33 deaths from December 2019 onset cases is also available in two more peer-reviewed papers by Chuanhua Yu and his team, given as papers A and C in the Table 3 below:

Chuanhua Yu's Papers	Paper A	Paper B	Paper C
Title	The epidemiological characteristics of deaths with COVID-19 in the early stage of epidemic in Wuhan, China	The comparison of epidemiological characteristics between confirmed and clinically diagnosed cases with COVID-19 during the early epidemic in Wuhan, China	The Association of Hypertension With the Severity of and Mortality From the COVID-19 in the Early Stage of the Epidemic in Wuhan, China: A Multicenter Retrospective Cohort Study
Up to	24 Feb 2020	24 Feb 2020	24 Feb 2020
Extracted	<16 Oct 2020	<17 Dec 2020	<30 Oct 2020
Sent to journal	16 Oct 2020	17 Dec 2020	30 Oct 2020
Disclosed	21 Dec 2020	28 May 2021	12 May 2021
Focus	All of Wuhan	All of Wuhan	All of Wuhan
Total Dec 2019 onset cases - Wuhan		227	
lab-confirmed		135	
diagnosed		92	
Deaths for Dec 2019 onset - Wuhan, by 24 Feb 2020	33	33	33
from lab-confirmed		29	
from diagnosed		4	

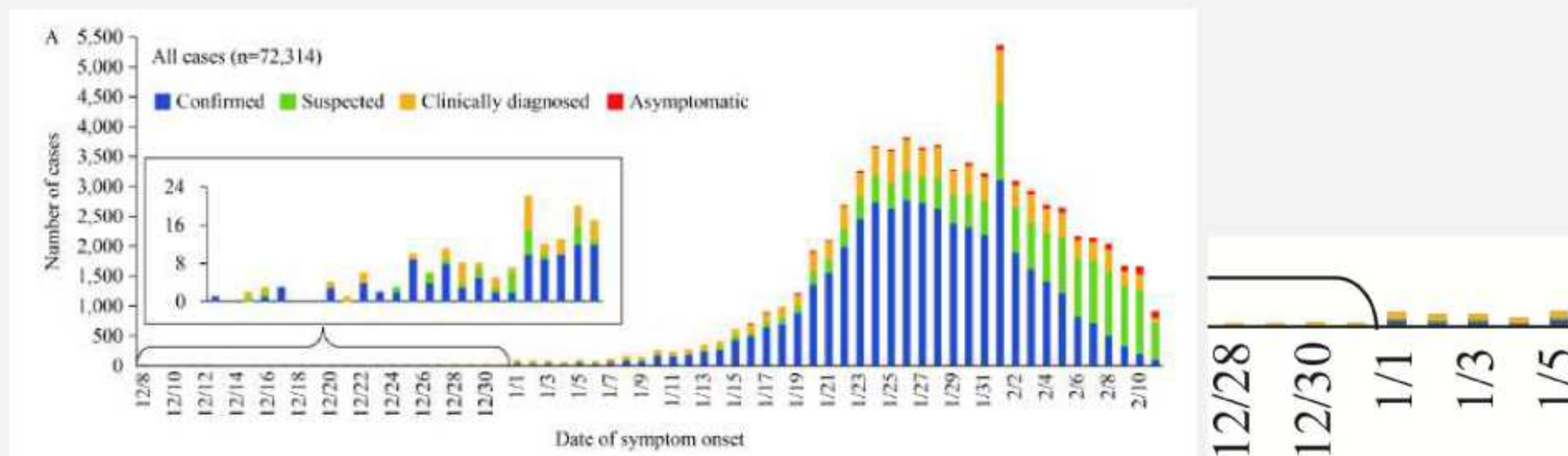
Table 3: Summary of main Dec 2019 counts of three papers by Chuanhua and his team.

Box 2: Examining the jumps in the 'confirmed' cases

Why are 'confirmed' cases jumping on 1 Jan 2020 and 1 Feb 2020 (by onset date)?

As shown in the example below extracted from the 'CDC Weekly' update of 17 Feb 2020 (data as of 11 Feb), the Wuhan cases counts have a major spike in cases with 1 Feb onset, and also a level jump on 1 Jan 2020.

We may postulate an indirect 'default date' effect as many cases were processed and added to the NNDRS database specifically between 9 and 12 Feb 2020 (see Finding 14), possibly without very precise details about their onset date. This could have resulted in some end of January onset cases being reported as 1 Feb, and some beginning of February cases reported as of 1 Feb. Something similar, possibly more focussed on 'delayed' late Dec 2019 onset cases, may explain the jump on 1 Jan (by onset date) which we can also see.



Part of Figure 3 of 'CDC Weekly' Update (17 Feb), with enlarged section on right

Why are 'confirmed' cases also slightly jumping on 9 Feb 2020 (by reporting date)?

The graph below, extracted from the report of the first WHO mission (28 Feb 2020), shows the 1 Jan and 1 Feb jumps by onset, but also a jump by reporting date on 9 Feb.

Following the addition on 5 Feb of a new case category ('clinically diagnosed' or simply 'diagnosed'), a retrospective search started and went on for most of the month. That search returned not only some 'diagnosed' cases but also a sizable number of retrospectively 'confirmed' cases identified as part of the same search effort.

As a result, there was a sudden increase in the number of 'confirmed' cases reported into the NNDRS database on 9 Feb 2020, the day a large bunch of cases were entered into the database based on the search started only a few days before (see also Finding 14 for more details about the retrospective search process). Practically Chinese authorities still delayed until 12 Feb 2020 to publicly disclose a large one-off spike of 'clinically diagnosed' cases, with 13,332 announced on that day, as well as a lesser peak in 'confirmed' cases.²⁰



Part of Figure 3 of the report of the First WHO mission (Feb 2020)

²⁰ See WHO Situation Report of 13 Feb 2020, available at <https://bit.ly/3Sjw6Fg>.

Finding 3:

The number of Dec 2019 onset cases in Wuhan, recorded in the NNDRS database as of 27 Feb 2020, was necessarily between 247 and 260 cases with at least 165 confirmed, and most likely towards the top of that range.

1. First minimal count:

We can first get a straightforward minimal count of Dec 2019 onset cases in the CDC official database (NNDRS) as of 27 Feb 2020, by combining the 146 lab-confirmed cases of Finding 1 (database as of 27 Feb) with the 19 confirmed Wuhan cases for 30 and 31 Dec 2019 onset available in the final WHO report (see Finding 5), and adding to it the 74 diagnosed cases of the final WHO report.

This gives us a strict **minimum** of 239 December cases (165 confirmed + 74 diagnosed) as of 27 Feb 2020. This is a minimum since (i) all the 74 diagnosed of the WHO report were likely pulled from the CDC database around February 2021, and would all have to be already present in that database on 27 Feb 2020 as no new diagnosed cases could be added after 17 Feb 2020, (ii) confirmed cases must remain confirmed.

Minimum count of 2019 cases in CDC DB (NNDRS) as of 27 Feb 2020	Cases	DB extract	Source (see Table 6)
Confirmed Dec 19 to 29 Dec from Wu-Chun Cao et al	146	27 Feb 2020	#7
Confirmed for 30 & 31 Dec 2019 from final WHO report	19	~ Feb 2021	#12
Total Confirmed for Dec 2019 as of 27 Feb 2020	165		
Diagnosed for Dec 19 from final WHO report	74	~ Feb 2021	#12
Total Confirmed + Diagnosed for Dec 2019 as of 27 Feb 2020	239		

Table 4a: First minimum count of confirmed and diagnosed cases in the CDC NNDRS as of 27 Feb 2020.

Instead of using the final WHO report (Mar 2021) for the Wuhan confirmed cases on onset dates 30 and 31 Dec 2019 (19 cases), we may have been tempted to use either (see Finding 7):

- the first WHO report published 28 Feb 2020 using data as of 20 Feb 2020: 22 confirmed cases for 30/31 Dec 2019 onsets, or
- the 'CDC Weekly' update published 17 Feb 2020 using data as of 11 Feb 2020: 24 confirmed cases for 30/31 Dec 2019 onsets.

All these choices are defensible since in theory confirmed cases cannot be retracted (at most the onset date might be adjusted). However, as we are calculating a minimum, we should automatically use the lowest count (19) which is the one given by the final WHO report.

2. Tighter minimal count:

We can then get a tighter minimal count of Dec 2019 onset cases in the CDC official database (NNDRS) as of 27 Feb 2020, by using the 82 diagnosed from a peer-reviewed paper which extracted the data from the NNDRS on 8 Mar 2020 (see Finding 7.c). That paper, Pan et al., is from scientists from Harvard, Huazhong University (Wuhan) and Fudan University (Shanghai).

This gives us a stricter **minimum** of 247 December cases (165 confirmed + 82 diagnosed) as of 27 Feb 2020, since all the 82 diagnosed had to be already in the database on 27 Feb and all the confirmed must remain so.

Minimum count of 2019 cases in CDC DB (NNDRS) as of 27 Feb 2020	Cases	DB extract	Source (see Table 6)
Confirmed Dec 19 to 29 Dec from Wu-Chun Cao et al	146	27 Feb 2020	#7
Confirmed for 30 & 31 Dec 2019 from final WHO report	19	~ Feb 2021	#12
Total Confirmed for Dec 2019 as of 27 Feb 2020	165		
Diagnosed for Dec 19 from Pan et al	82	8 Mar 2020	#8
Total Confirmed + Diagnosed for Dec 2019 as of 27 Feb 2020	247		

Table 4b: Second minimum count of confirmed and diagnosed cases in the CDC NNDRS as of 27 Feb 2020.

3. Maximal count:

We can now get a maximum count of Dec 2019 onset cases in the CDC official database (NNDRS) as of 27 Feb 2020, by using a number of diagnosed cases extracted just a few days before. For that we turn to the 92 'diagnosed' from Chuanhua *et al.* which were extracted on 24 Feb 2020.

Note that in theory some of these diagnosed could have been requalified as confirmed between the 24 and 27 Feb, which would result in double-counting. Also some diagnosed could have dropped altogether after testing negative. So there could be some double-counting of a few confirmed and some counting of few rejected diagnosed by 27 Feb, which is precisely why this is a maximum. In all fairness, since the number of diagnosed only dropped by 10 between 24 Feb and 8 Mar 2020 (13 days), there would at most be a few cases in excess (likely no more than 3 or 4 out of the 10) for the 3 days from 24 Feb to 27 Feb.

Also, as this is a maximum, we should not use the lower count of confirmed for 30/31 Dec 2019, but the higher bound closer to 27 Feb 2020. We thus take the count of 22 confirmed as of 20 Feb 2020 (based on the first WHO report). This is 3 more than if we had taken instead the final WHO report count of 19.

Maximum count of 2019 cases in CDC DB (NNDRS) as of 27 Feb 2020	Cases	DB extract	Source (see Table 6)
Confirmed Dec 19 to 29 Dec from Wu-Chun Cao <i>et al</i>	146	27 Feb 2020	#7
Confirmed for 30 & 31 Dec 2019 from final WHO report	22	20 Feb 2020	#5
Total Confirmed for Dec 2019	168		
Diagnosed for Dec 19 from Chuanhua <i>et al</i>	92	24 Feb 2020	#6
Max Confirmed + Diagnosed for Dec 2019	260		
Comparison: SCMP Leak for Dec 2019	257	< 13 Mar 2020	#10

Table 4c: Maximum count of confirmed and diagnosed cases in the CDC NNDRS as of 27 Feb 2020.

We thus obtain a **maximum** number of 260 Dec 2019 cases as of the 27 Feb 2020, given as 168 confirmed and 92 diagnosed, which may overshoot by a few diagnosed and at most 3 confirmed as argued above.

Box 3: Wuhan vs. Hubei vs. China for 2019 onset cases

As far as 2019 onset cases are concerned, the available data shows that they were mostly in the city of Wuhan with a few cases (10 in the final WHO report) sprinkled in neighbouring cities.

The geographical classification of 2019 onset cases in the final WHO report is based on declared home addresses (see WHO report on page 43), while the entering of the data into the main database (NNDRS) would have been done by the hospitals in Wuhan where the cases would have eventually ended up and been asserted (if necessary after transfer from local clinics or hospitals).

A material distinction between Hubei and Wuhan cases comes to the fore only for onset dates starting in Jan 2020, when cases spread out of Wuhan. Before that, even the few cases pinned to neighbouring cities were still often reported in Wuhan hospitals. Accordingly many papers explicitly refer to pulling Dec 2019 onset cases from the NNDRS database only over Wuhan hospitals (typically not all of them). However there are some exceptions: the Terms of Reference of the WHO mission mention five confirmed Dec 2019 onset cases from neighbouring cities 'and other provinces', all with travel links to Wuhan, which by definition means some Dec 2019 onset cases out of Wuhan and even out of Hubei. Also the first WHO report, which definitely pulled data at the national level and not just data contributed by Wuhan hospitals, shows at least four Dec 2019 onset cases in Hubei out of Wuhan, and at least six out of Hubei (either diagnosed or confirmed).

Finding 4:

The 257 Dec 2019 onset cases from the leak published in the South China Morning Post of Hong Kong on 13 Mar 2020 are perfectly consistent with this necessary range of 247 to 260.

The South China Morning Post published on 13 Mar 2020 a detailed article²¹ about cases with onset dates in November and December 2019. The article gives unreleased 'government data' as its source, meaning most likely data from the Chinese CDC.

The article mentions that:

Chinese authorities have so far identified at least 266 people who were infected last year, all of whom came under medical surveillance at some point.

Some of the cases were likely backdated after health authorities had tested specimens taken from suspected patients.

It then proceeds to give more details about these cases:

According to the government data seen by the Post, a 55 year-old from Hubei province could have been the first person to have contracted COVID-19 on November 17.

From that date onwards, one to five new cases were reported each day. By December 15, the total number of infections stood at 27 – the first double-digit daily rise was reported on December 17 – and by December 20, the total number of confirmed cases had reached 60.

By the final day of 2019, the number of confirmed cases had risen to **266**. On the first day of 2020 it stood at 381.

Of the first **nine** cases to be reported in November – four men and five women – none has been confirmed as being "patient zero". They were all aged between 39 and 79, but it is unknown how many were residents of Wuhan, the capital of Hubei and the epicentre of the outbreak.

The wording of the article lets us understand that these are Hubei numbers, just as with the final WHO report. Note that despite the liberal use of 'confirmed' in the SCMP article, the count is most likely of 'confirmed' plus 'diagnosed' cases. Indeed 'confirmed' on its own would seem excessive, while 'suspected' cases don't show up in either Chuanhua et al. (Finding 2) or Wu-Chun Cao et al. (Finding 1) and actually quickly lost centre-stage once 'diagnosed' was introduced as a category.

The main numbers are summarised below, while reinterpreting 'confirmed' as "confirmed plus diagnosed", which was the focus of contemporary reporting:

Onset Latest	Running Total 'Conf.' + 'Diagn.'	Note
17-Nov-19	1	55 year old
30-Nov-19	9	4 men, 5 women, from 39 to 79 year-old
15-Dec-19	27	
20-Dec-19	60	First double-digit increase on 17 Dec
27-Dec-19	> 180	more than 180
31-Dec-19	266	
1-Jan-20	381	jump

Table 5: Case count (as running total) from the SCMP leak of 13 March 2020.

Removing the nine cases with November 2019 onsets from the total of 266 to 31st Dec 2019, we get 257 cases for December 2019. This falls precisely within our necessary range of 247 to 260 'confirmed' + 'diagnosed'

²¹ Available at <https://archive.ph/wip/FeOyq>

cases in the NNDRS database as of 27 Feb 2020, calculated from peer-reviewed papers and the WHO reports. It also lands towards the top of that range as we expected.

This represents an excellent and simple match, even after allowing for possible still unaccounted marginal effects, such as some possible mismatch between 'Wuhan and neighbouring cities' vs. Hubei (see Box 3).

We further note that the SCMP leak reported a first double-digit jump on 17 Dec 2019:

By December 15, the total number of infections stood at 27 – the first double-digit daily rise was reported on December 17.

This is actually fully consistent with Wu-Chun Cao et al. which shows 41 confirmed cases for the week of 16 to 22 Dec 2019. With an average of 6 confirmed per day during that week, to which one needs to add diagnosed cases, it seems very likely that one day in that week would get into double digit. Inversely the count of confirmed cases for the previous week (9 in all) seems much too low to reach at least 10 'confirmed + diagnosed' cases on a day during that previous period.

For all purposes the SCMP leak is fully credible. A random creation of the SCMP numbers in early March 2020, for the purpose of a fake, would be very unlikely to deliver that excellent quality of match to our inferred necessary range of 247 to 260 cases as of 27 Feb 2020, all based on peer-reviewed papers and WHO reports, especially before these peer-reviewed papers were even published.

Box 4: The NNDRS database

The NNDRS is the National Epidemic Network Direct Reporting system set up by the CDC after SARS-1. The system was supposed to reflect the exact count of reportable cases of various contagious diseases at the end of every day, with a required alert if five cases or more of a new pneumonia appeared together in one place.

The system failed to detect the outbreak in late 2019 due to a variety of issues, including low awareness of its existence and functioning among many front line doctors, confusion as to the reporting obligation for none-identified respiratory disease, cumbersome manual entry, the unwillingness to take the risk of making a wrong report and getting blamed, and - most importantly - multi level interference by local authorities. It is also notable that the system is [under the control of the party secretary of the Chinese CDC](#), and not under the control of his director (Gao Fu), which adds another avenue for interference.

One key design fault was that the system just translated the previous paper process into an electronic one without rethinking it. The main reason for this is that the paper process enshrined the subordination of the CDC to the local, provincial and national health commissions and to the political body. The CDC (which is under the National Health Commission) has no administrative power. Under the relevant laws the CDC is not allowed to publish epidemic numbers on its own or to declare an epidemic. It only acts as a consulting entity to the health commissions, which are all political entities and are the ones with administrative power.

This resulted in a slow multi-layered process, largely under the influence of the health commissions, all political entities. These could filter reporting at many levels. There were for instance [deletions of cases entered into the NNDRS by front-line doctors](#) at a top research hospital in Wuhan (Zhongnan hospital, south side of the river) on 9 Jan 2020, right when there was no new case officially reported for about 13 days (5 to 17 Jan 2020).

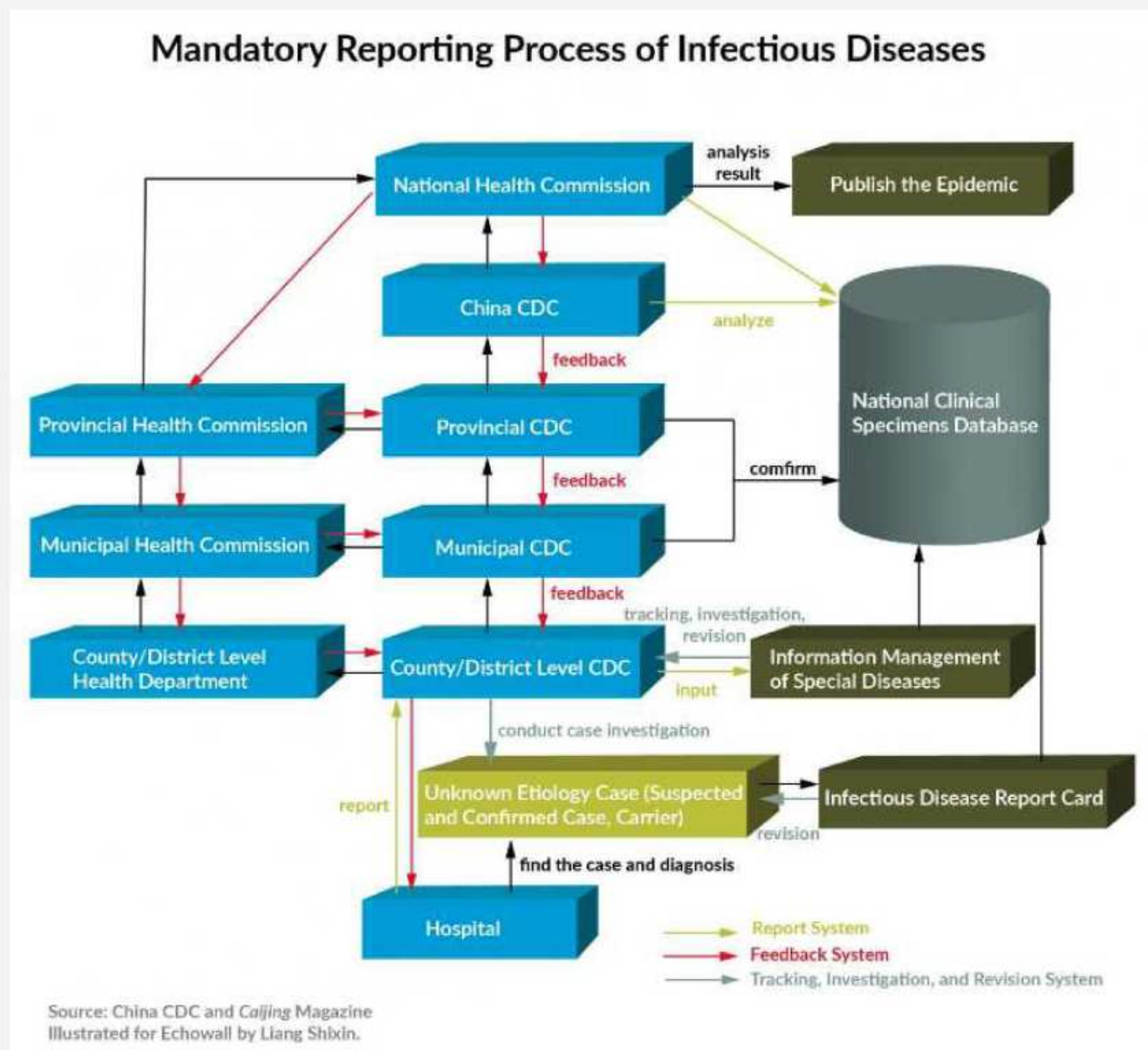


image from ['For Whom the Warning System Tolls'](#) by Liu Cong, Jan 2021

Sources: ['Why did the infectious disease network direct reporting system built with an investment of 730 million fail for 28 days?'](#) (Caixin), ['Technical governance and emergency decision-making in a risk society: Focusing on the early treatment of the new coronavirus pneumonia epidemic'](#) (China Law Review), ['Zhongnan Hospital physicians in the Wuhan Epidemic: There were sick people everywhere'](#) (Caixin), ['Design and Application defects of the National Epidemic Direct Reporting System. Part 1 — Why the system is virtually useless'](#) (Jeff Zhao), ['For Whom the Warning System Tolls'](#) (Liu Cong). Hua Sheng (Watson)'s [blog entries](#).

Finding 5:

This necessary range of 247 to 260, with a minimum of 165 confirmed, is much more than the 174 Dec 2019 cases (100 confirmed + 74 diagnosed) for Wuhan and neighbouring cities of the final WHO report.

The final WHO report²² explains that the Wuhan CDC extracted 174 cases with Dec 2019 onset dates from its NNDRS database, as 100 confirmed cases plus 74 diagnosed ones, 164 being in the city of Wuhan and 10 in neighbouring cities in Hubei (using home addresses, as per Box 3).

We note that, in particular, the minimum 165 confirmed cases in the NNDRS database at end Feb 2020 are much more than the 100 confirmed cases of the WHO report. This may partly reflect the fact that the retrospective search was able to focus on 'confirmed' cases after the 17 Feb (when 'diagnosed' was not an applicable case definition any more) till the end of Feb and the ensuing gag orders. That retrospective search would effectively have had some time to more systematically review cases through Wuhan hospitals' records, not just the hospitals linked to the initial case reports or those in their neighbourhood, and organise some testing of samples.

Precise daily counts for December 2019 can be extracted from a graph on page 146 of the [Annexes](#) of the final WHO Report, as per Fig. 2 below.²³

Onset Date	Lab-Confirmed	Clinically Diagnosed	Total
8-Dec-19	1		1
9-Dec-19			
10-Dec-19			
11-Dec-19	1		1
12-Dec-19	3		3
13-Dec-19	1		1
14-Dec-19			
15-Dec-19	3		3
16-Dec-19		2	2
17-Dec-19	4	2	6
18-Dec-19	4	1	5
19-Dec-19	3		3
20-Dec-19	8	3	11
21-Dec-19	3	1	4
22-Dec-19	4	1	5
23-Dec-19	6	8	14
24-Dec-19	2	7	9
25-Dec-19	7	10	17
26-Dec-19	2	7	9
27-Dec-19	11	5	16
28-Dec-19	8	6	14
29-Dec-19	10	4	14
30-Dec-19	12	9	21
31-Dec-19	7	8	15
From graph	100	74	174

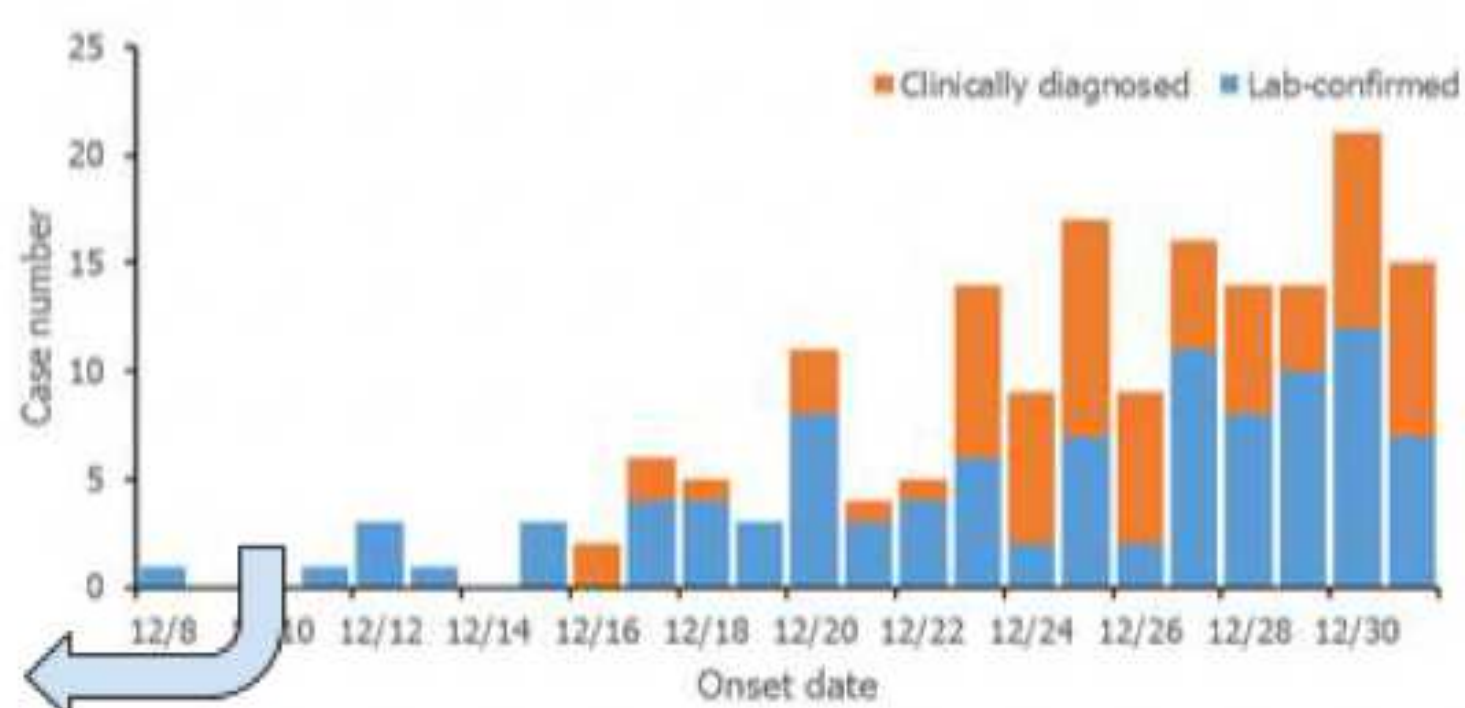


Fig. 2: Extracted onset dates of December 2019 cases from the final WHO report (March 2021)

²² Available at

<https://www.who.int/publications/i/item/who-convened-global-study-of-origins-of-sars-cov-2-china-part>

²³ Extraction of the data is relatively easy with a dedicated software, such as <https://apps.automeris.io/wpd/>.

Finding 6:

Three peer-reviewed papers from Chuanhua Yu's team show 33 deaths by 24 Feb 2020 (last day of follow-up) from cases with onset date in Dec 2019, including 29 from lab-confirmed cases. The only other official count for Dec 2019 onset deaths is the 'CDC Weekly' update published on 17 Feb 2020, which showed 15 deaths.

All three papers with Chuanhua Yu as corresponding author (listed in Table 3) give 33 deaths from Dec 2019 onset cases, with the deaths being recorded up to 24 February 2020 (last date for which the authors were able to pull patients' records). Paper B ([The comparison of epidemiological characteristics between confirmed and clinically diagnosed cases with COVID-19 during the early epidemic in Wuhan, China](#)) further gives the breakdown of these deaths by 25 Feb 2020 as 29 lab-confirmed cases and 4 diagnosed cases.

We extracted the exact Dec 2019 onset dates from the high-resolution Fig. 2 of Paper A ([The epidemiological characteristics of deaths with COVID-19 in the early stage of epidemic in Wuhan, China](#))²⁴:

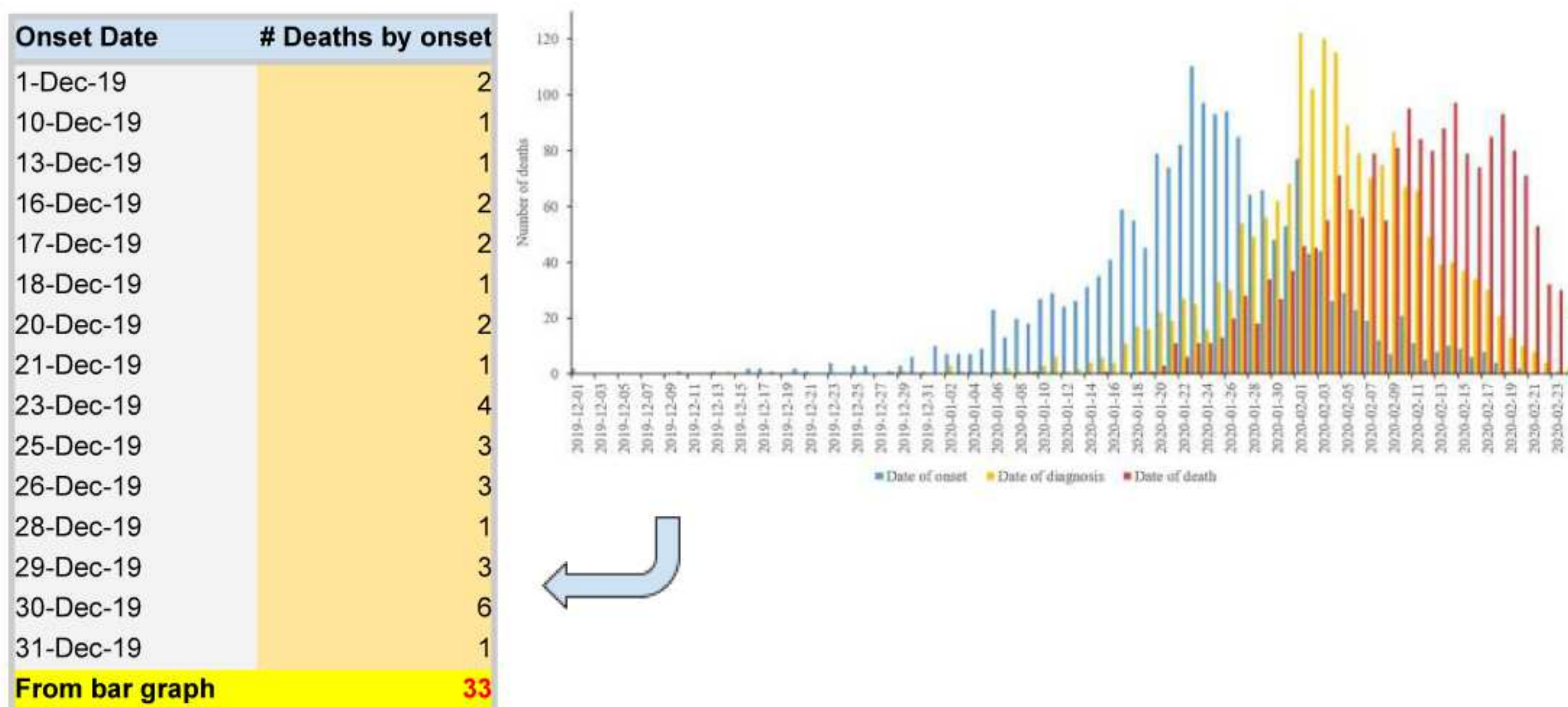


Fig. 3: Extracted onset dates of December 2019 cases that ended in death by 25 Feb 2020, from Fig. 2 (right) of [The epidemiological characteristics of deaths with COVID-19 in the early stage of epidemic in Wuhan, China](#).

In contrast, the [Chinese 'CDC Weekly' update published 17 Feb 2020](#) (using data as of 11 Feb) gave 15 deaths from 104 confirmed cases with Dec 2019 onset.

TABLE 1. Patients, deaths, and case fatality rates, as well as observed time and mortality for n=44,672 confirmed COVID-19 cases in Mainland China as of **February 11, 2020**.

Baseline characteristics	Confirmed cases, N (%)	Deaths, N (%)	Case fatality rate, %	Observed time, PD	Mortality, per 10 PD
Period (by date of onset)					
Before Dec 31, 2019	104 (0.2)	15 (1.5)	14.4	5,142	0.029
Jan 1–10, 2020	653 (1.5)	102 (10.0)	15.6	21,687	0.047
Jan 11–20, 2020	5,417 (12.1)	310 (30.3)	5.7	130,972	0.024
Jan 21–31, 2020	26,468 (59.2)	494 (48.3)	1.9	416,009	0.012
After Feb 1, 2020	12,030 (26.9)	102 (10.0)	0.8	87,799	0.012

Fig. 4: Extract of Table 1 of 'CDC Weekly' update published 17 Feb 2020.

²⁴ See tab 'Chuanhua-DecDeaths' of the [supplementary spreadsheet](#) for the details of the data extraction.

We shall note that the data in the Chuanhua papers includes two deaths for cases with an onset date of 1st Dec 2019. It could be that the onset dates for at least one of these two cases may be wrong, since the final WHO report mentions one 1st Dec onset case that was eventually changed to a later Dec 2019 onset date (see extract below). This in any case would not change the overall Dec 2019 count of 33.

Final WHO Report, page 46:

Other initially suspected cases in December 2019

Three possible cases with disease onset on 1, 2 and 7 December 2019, respectively, were initially identified as potential cases in the retrospective case search and have been included in some published papers. Clinical review of these three cases by the Chinese expert team led to their exclusion as possible cases on the basis of the clinical features of their illness.

In the case with onset on 1 December, a 62-year-old man with past history of cerebrovascular disease was judged to have had a minor respiratory illness in early December, which responded to antibiotics. He developed a further illness with onset on 26 December 2019, which was later laboratory-confirmed to be COVID-19. This patient had no reported contact to the Huanan market, whereas his wife, who was admitted on 26 December with a COVID-19 compatible illness, reported close contact with the Huanan market. She was also later laboratory-confirmed to have COVID-19. This couple, together with their son, became part of the first recognized family cluster of COVID-19.

Against this interpretation is the fact that the 'wrong' 1st Dec case mentioned in the final WHO report did not die, according to the information that was given about it. So it clearly cannot be one of these 2 deaths listed with onset on 1st Dec 2019, whatever the true onset date of that case was.²⁵

Alternatively, as we can also see for the number of cases with onset date 1st Jan and 1st Feb 2020 (see Box 2), the 1st Dec 2019 onset cases may be examples of onsets at 'end of November' or 'beginning of December' being defaulted to the first day of December, before being more precisely pinned to the actual November or December onset dates later.

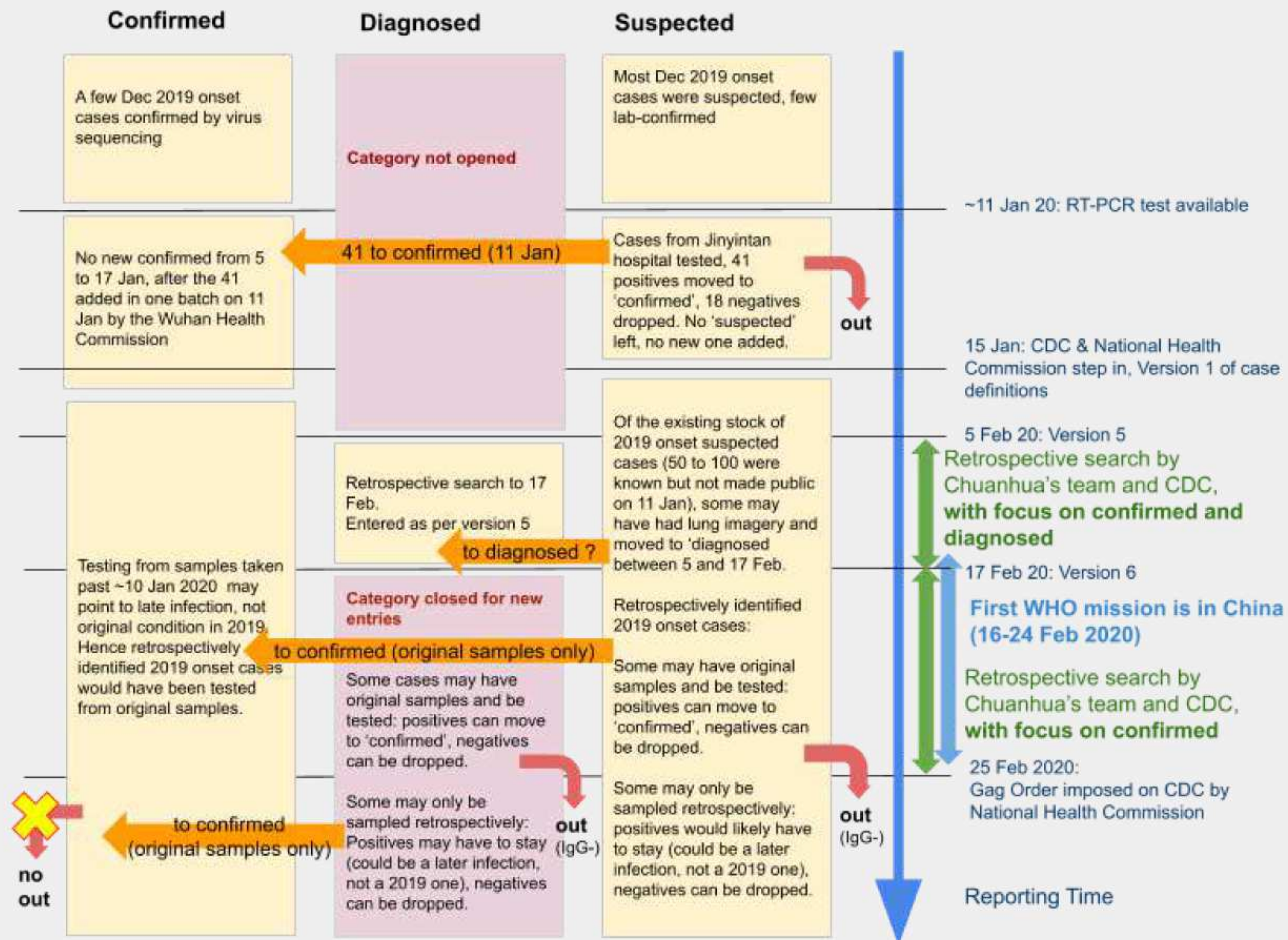
Incidentally, the 'CDC Weekly' update has its own incoherence related to December data: the graph in the update gives 102 confirmed cases (see Fig. 5 below) while the Table 1 of the update gives 104 cases (see Fig. 4 above). These two issues could be related, which would also mean that the two 1st Dec cases that ended in deaths are 'confirmed' ones (since the 104 count is a count of confirmed cases), something that should not be surprising at all since 29 of the 33 deaths were known to be from confirmed cases. If so, the issue with the graph showing 102 cases would simply be that the horizontal axis starts on 8 Dec 2019, with the two 'missing' confirmed cases having onsets falling before that date and thus not being included in the graph.

However confused the picture about early Dec 2019 onset cases may already seem by this stage, this is still not the end of the inconsistencies related to early December cases; following the publication of its final report in March 2021, the WHO soon had to correct some mis-reporting in the identity of the earliest December cases, after some questioning by the Washington Post²⁶.

²⁵ For confirmation of survival, see Extended Table 1 from 'A pneumonia outbreak associated with a new coronavirus of probable bat origin' published in Nature on 3 Feb 2020:
<https://www.nature.com/articles/s41586-020-2012-7/tables/1>.

²⁶ See 'WHO clarifies details of early covid patients in Wuhan after errors in virus report', 15 July 2021. Available at
https://www.washingtonpost.com/world/asia_pacific/covid-wuhan-outbreak-who/2021/07/15/51e7e8a6-e2c6-11eb-88c5-4fd6382c47cb_story.html

Box 5: Pathways for 2019 onset cases by reporting date



Finding 7:

Before the final WHO report with its 174 cases (104 confirmed, 70 diagnosed) for Dec 2019 onset, other sources constantly showed higher numbers.

We will show in this section that the following numbers are available from official sources and/or peer-review literature:

- 104 confirmed cases + 62 (diagnosed + suspected) ones, as of 11 Feb 2020, from the CDC
- 108 confirmed cases + 83 diagnosed ones, as of 20 Feb 2020, in the first WHO report
- 115 confirmed cases + 82 diagnosed ones, as of 8 Mar 2020, in a peer-reviewed paper
- 124 confirmed cases in the Aug 2020 Terms of Reference (ToRs) of the WHO mission.

a. The [Chinese 'CDC Weekly' update](#) published on 17 Feb 2020 was the one and only detailed CDC publication about the COVID-19 case count for December 2019, and it was soon followed by a gag order imposed on the CDC by the National Health Commission (25 Feb). The 'CDC Weekly' update explains that the data was extracted from the CDC NNDRS system on 11 Feb.

By inverting the high definition Figure 3 of the 'CDC Weekly' update, we can tabulate the cases for December 2019 onset date²⁷. In doing so we note that we only get 102 confirmed cases against the 104 in the associated 'CDC Weekly' table (see Fig. 4).

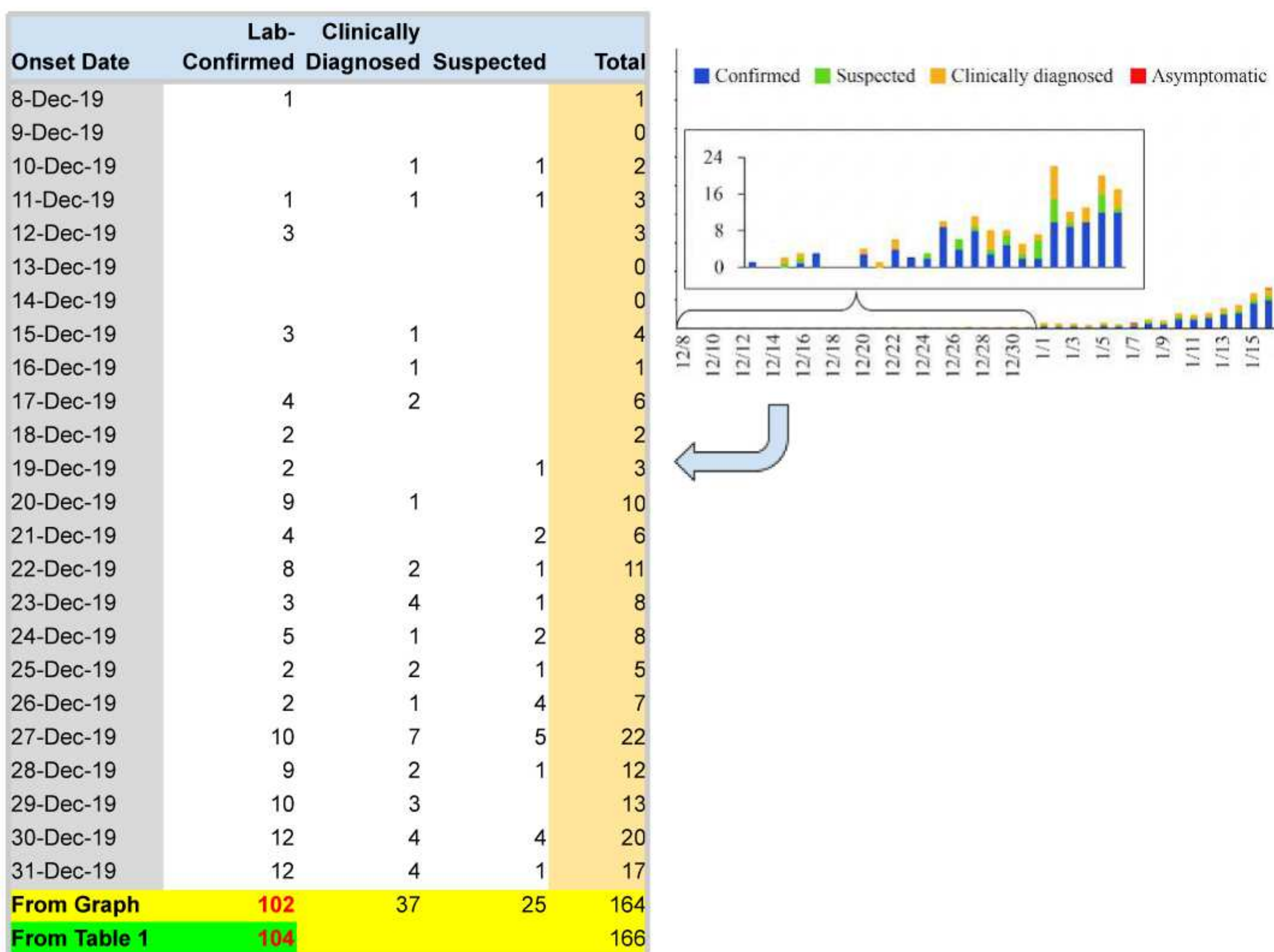


Fig. 5: Extracted onset dates of Dec 2019, from Fig. 3 (right) of the [Chinese 'CDC Weekly' update](#)

²⁷ In doing so we observe that the graph (starting 8 Dec) has only 102 confirmed cases while Table 1 has 104.

b. The report of the first WHO mission, published on 28 Feb 2020 and using data as of 20 Feb, offers another snapshot only 9 days after the ‘CDC Weekly’ update (20 Feb vs. 11 Feb). The report explains that the data was extracted from the CDC NNDRS system. Again we obtain the tabulated data by inverting the high resolution image in the report²⁸.

We note a first diagnosed case on 2nd Dec 2020.

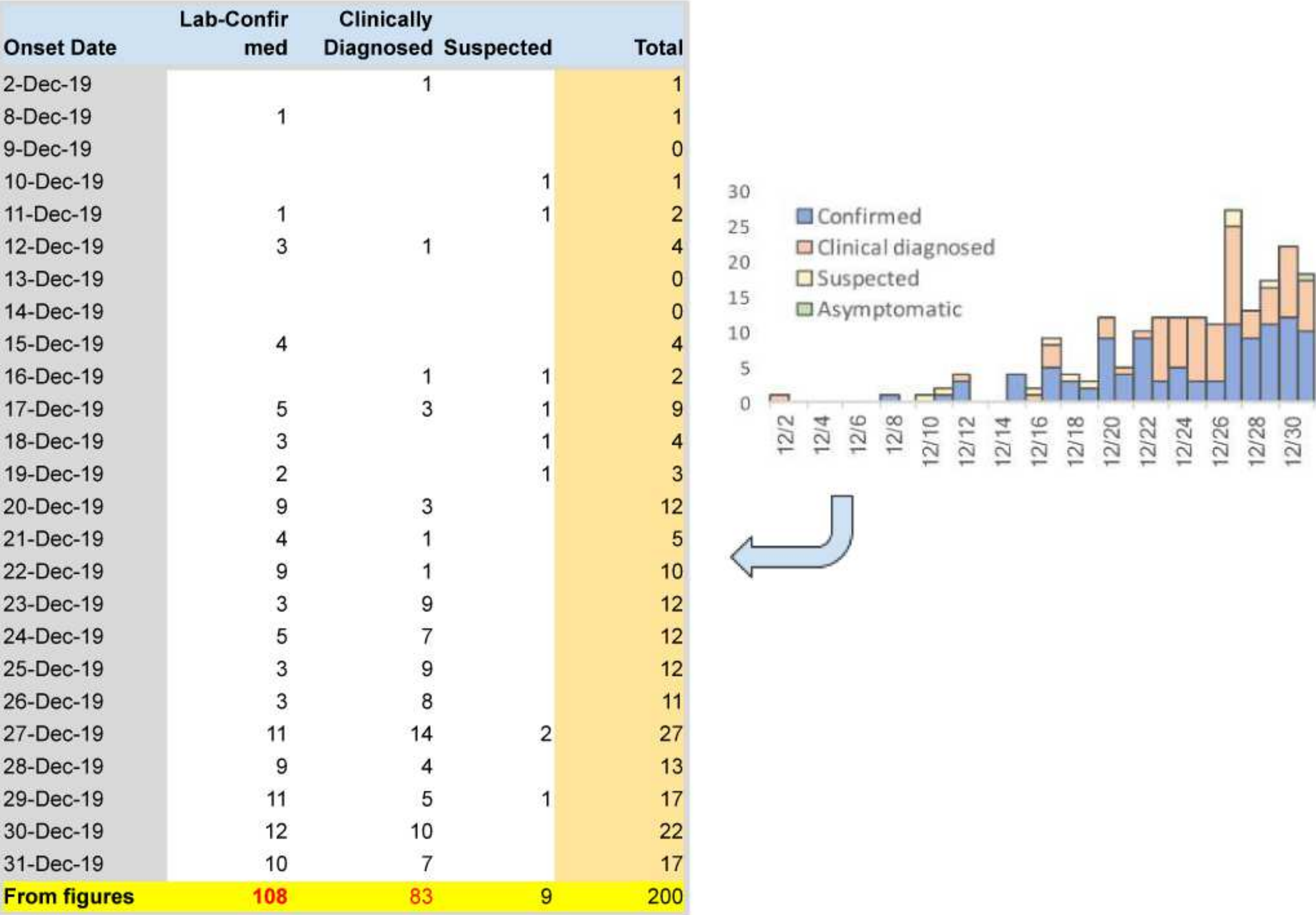


Fig. 6: Extracted onset dates of Dec 2019 for Wuhan, from Fig. 2 (right) of [the first WHO report \(Feb 2020\)](#)

²⁸ See tab ‘WHO_20Feb2020’ of the [supplementary spreadsheet](#) for the details of the data extraction.

c. A paper by scientists from Harvard, Huazhong University (Wuhan) and Fudan University (Shanghai), [‘Association of Public Health Interventions With the Epidemiology of the COVID-19 Outbreak in Wuhan, China’](#) published in April 2020 in the Journal of the American Medical Association (JAMA), gives us a count of 115 ‘confirmed’ cases and 82 ‘diagnosed ones’ for Dec 2019.

The paper (hereafter Pan et al.) explains that the data was pulled from the CDC NNDRS database and is as of 8 Mar 2020. We again extracted the data from a high resolution graph given in Figure 1 of that paper, as per Fig. 5 below.²⁹

Onset Date	Lab-Confirmed	Clinically Diagnosed	Total
2-Dec-19			
8-Dec-19	1		1
9-Dec-19			
10-Dec-19			
11-Dec-19	1		1
12-Dec-19	3	1	4
13-Dec-19			
14-Dec-19			
15-Dec-19	4		4
16-Dec-19		1	1
17-Dec-19	5	3	8
18-Dec-19	3		3
19-Dec-19	2		2
20-Dec-19	9	3	12
21-Dec-19	4	1	5
22-Dec-19	10	1	11
23-Dec-19	5	9	14
24-Dec-19	5	7	12
25-Dec-19	4	9	13
26-Dec-19	3	8	11
27-Dec-19	13	13	26
28-Dec-19	9	4	13
29-Dec-19	11	5	16
30-Dec-19	13	10	23
31-Dec-19	10	7	17
From figure	115	82	197

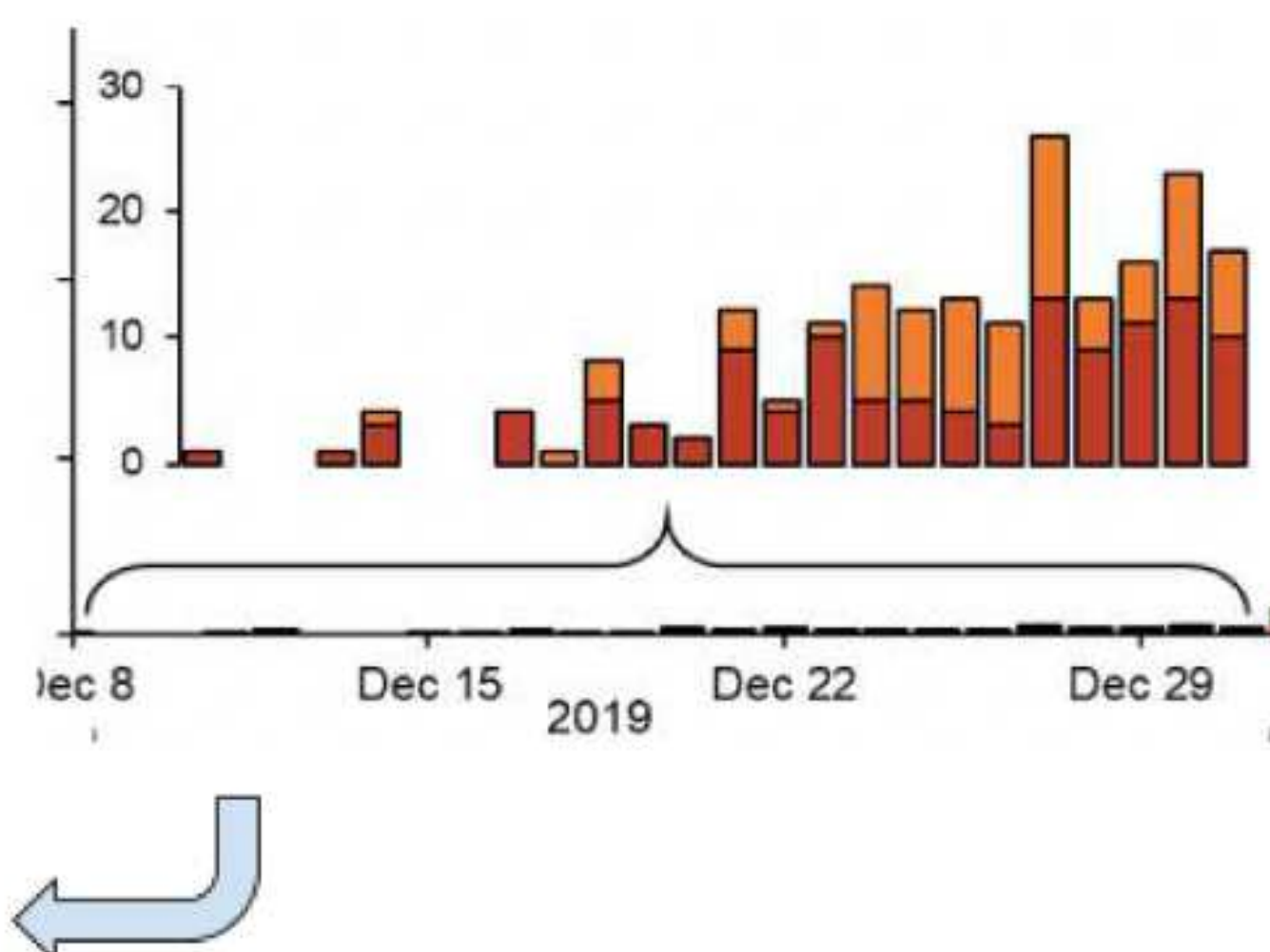


Fig. 8: Extracted Dec 2019 onset dates, using eFigure 1 in Supplement (right) of [‘Association of Public Health Interventions With the Epidemiology of the COVID-19 Outbreak in Wuhan, China’](#).

We were further able to check our inversion of the graph against the data in the GitHub repository of a Nature paper, [‘Reconstruction of the full transmission dynamics of COVID-19 in Wuhan’](#), published in July 2020 and written by the same authors from Huazhong University (Wuhan). The [data file](#) included in that repo fully confirms our ‘lab-confirmed’ count extracted from the JAMA paper, with exactly the same cases count per day for December 2019. The [Readme.md](#) file of that GitHub repo further explains that the data used is the same as the one for their JAMA paper:

²⁹ See tab ‘CDC_8Mar20’ of the [supplementary spreadsheet](#) for the details of the data extraction.

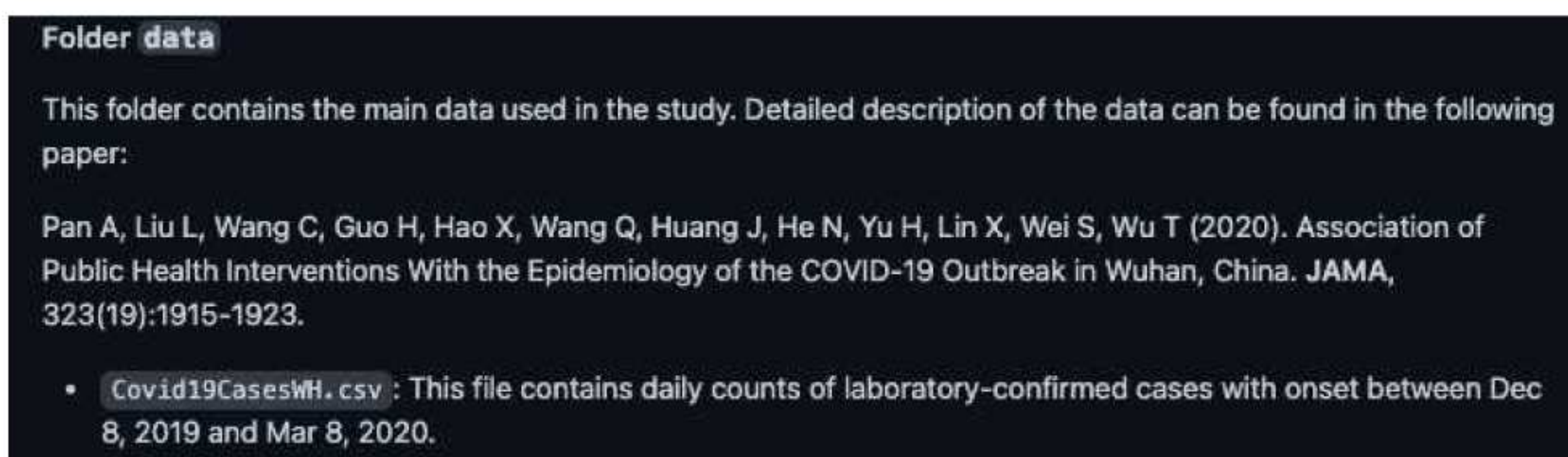


Fig. 9: Extract of Readme.md file of GitHub repo for '[Reconstruction of the full transmission dynamics of COVID-19 in Wuhan](#)'.

d. The [Terms of Reference](#) (ToRs) of the first WHO mission³⁰, written in August 2020, give 124 'confirmed' cases with an onset date in December 2019, including 119 for Wuhan and 5 that had travel links to Wuhan.

On the subject of these 5 with travel links, the ToRs mention that these are either from Hubei or neighbouring provinces. So effectively that count of 124 cases contains a few (4 at most) cases that are out of the scope of 174 of the final WHO report (which focussed on Wuhan and Hubei).

Terms of References, page 5:

Retrospective review of cases identified a total of **124** confirmed cases with onset date in December 2019, 119 of whom were from Wuhan and 5 others from Hubei or other provinces, but all with travel links to Wuhan during the period of exposure.

Finding 8:

Summary of reported number of Dec 2019 onset cases, from all sources reviewed.

We summarised all these different data points in Table 6 and Fig. 10 below. As we can see, the number of 'confirmed' December 2019 cases in the official NNDRS database increased regularly in February 2020, starting with 104 cases on 11 Feb, to 109 on 20 Feb, then 135 on 24 Feb and 165 on 27 Feb (using 146 up to 29 Dec + 19 cases from the final WHO report for the last two days of December).

That steady increase was relatively fast and happened mostly during the two weeks or so of intense retrospective search to which Chuanhua Yu and his team took a very active part. That search largely overlapped with the presence of the first WHO mission in China (16-24 Feb 2020). Eventually the peak on 27 Feb 2020 came two days after the National Health Commission imposed a gag order on the CDC (25 Feb 2020), which itself happened on the day following the departure of the first WHO mission from China (24 Feb 2020).

³⁰ Available at <https://bit.ly/3V5Je38>. For more details about the ToRs see <https://bit.ly/34fC5l2>.

December 2019 onset cases

Source	Data up to	Data extracted	Disclosure	Confirmed only	Confirmed + Diagnosed	Confirmed + Diagnosed + Suspected	Type	Source
#1	2 Jan 2020	2 Jan 2020	3 Jan 2020	??	41	41	official	China official
#2	5 Jan 2020	5 Jan 2020	5 Jan 2020	??	59	59	semi-official, key paper	Lancet
#3	10 Jan 2020	10 Jan 2020	11 Jan 2020	41	41	>90	confirmed from major paper, leaks for suspected	China official
#4	11 Feb 2020	< 17 Feb 2020	17 Feb 2020	104	141	166	official - from CDC report	CDC weekly
#5	20 Feb 2020	< 28 Feb 2020	28 Feb 2020	109	192	201	semi-official - from WHO report graph	1st WHO report
#6	24 Feb 2020	< 16 Oct 2020	21 Dec 2020	135	227	??	reliable - paper from national cases data team	Chuanhua et al.
#7	27 Feb 2020	27 Feb 2020	29 Sep 2020	146	??	??	146 to 29 Dec 2019, paper using official DB with CDC help	Wu-Chun Cao et al.
min	27 Feb 2020		mixed	165	247	??	based on #7, #8 and #12	mixed
max	27 Feb 2020		mixed	168	260	??	based on #5, #7 and #6	mixed
#8	8 Mar 2020	9 Mar 2020	10 Apr 2020	115	197	??	paper, from graph	Pan et al.
#10	< 13 Mar 2020	< 13 Mar 2020	13 Mar 2020	??	257	??	leaked 'government data', likely CDC	CDC
#11	< 31 Jul 2020	< 31 Jul 2020	5 Nov 2020	124	??	??	semi-official (WHO TORs)	WHO Terms of Ref.
#12	< 20 Feb 2021	< 20 Feb 2021	30 Mar 2021	100	174	??	Official - Final	2nd WHO report

Table 6: Evolution of 2019 onset case counts.

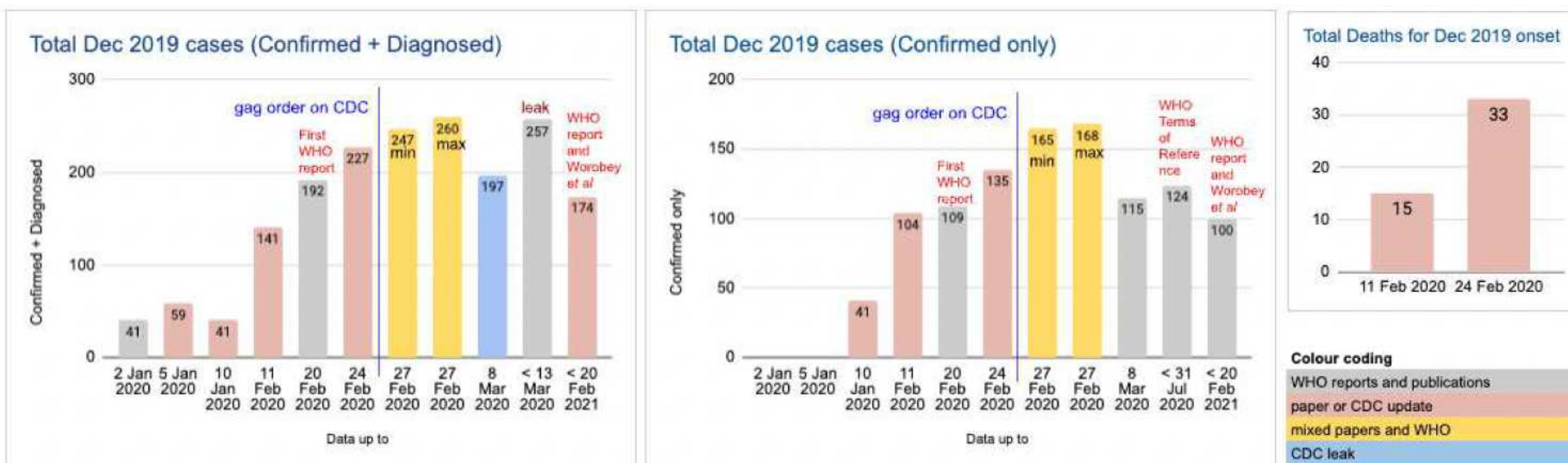


Fig 10: Evolution of 2019 onset case counts by extraction date.

Finding 9:

The evolution of the count of 'confirmed' cases from its peak at the end of February 2020 is very difficult to understand.

Over 16 key days, from the 104 confirmed cases for Dec 2019 onset of the 'CDC Weekly' update (data extracted on 11 Feb 2020) to the 146 confirmed cases with onset from 8 to 29 Dec 2019 of Wu-Chun Cao et al. (extracted on 27 Feb 2020), we can see a gradual and logical progression in the number of confirmed cases, concomitant to the focussed effort at the time to retrospectively identify 'confirmed' and 'diagnosed' cases (see Table 6).

In particular we note that after 17 Feb 2020, as no new 'diagnosed' cases could be added after that case definition was dropped, the retrospective search involving Chuanhua Yu and the CDC would have likely turned its focus to 'confirmed' cases. This is indeed what we can see in Table 6, with a strong acceleration in the number of 'confirmed' cases between 20 Feb and 27 Feb.

As all these cases were laboratory-confirmed, there is no logical explanation at all to the eventual reversion to 100 confirmed cases with Dec 2019 onset in the final WHO report. One could at most imagine a few double-counted cases, but certainly not enough to explain the disappearance of 65 laboratory-confirmed cases that were in the NNDRS database on 27 Feb 2020 (146 from Wu-Chun Cao et al. up to 29 Dec minus the 81 from the final WHO report to 29 Dec 2019)

In that context it is worth noting the wording in the main WHO report: **'including some that were included in early publications'**.

Other "cases" were identified as part of the search for other potential cases with onset in December 2019 (including some that were included in early publications). After clinical review by the Chinese team, none of the other cases were considered to be compatible with COVID-19 disease, leaving only the 174 notified cases.

Final WHO report, page 42.

This seems to indicate that the supernumerous cases previously identified in papers using the CDC NNDRS database, and also the supernumerous cases in all the sources listed in Table 6 that pulled data out of that official database over these key 16 days, were at some later stage found to be somehow 'wrong'.

This is rather difficult to imagine.

Starting with Chuanhua et al., how could that team, which was specifically in charge of cleaning and back-populating the official CDC NNDRS database, have either mis-tested its 35 supernumerous lab-confirmed cases for Dec 2019 (false positives) or somehow double-counted them³¹? Especially since the team was given 'official approvals for monitoring the data of registered patients with severe and critical COVID-19 illness in Wuhan until 25 February 2020'³².

Then how could Wu-Chun Cao et al. get things so wrong that its underlying data somehow contains 65 double-counted or mist-tested 'confirmed' cases for onset between 16 and 29 Dec 2019 (see Table 7), despite working with the help of the CDC to extract the relevant numbers from the NNDRS database, and despite Wu-Chun Cao himself being the very experienced chinese scientist that previously finalised the official count of SARS-1 cases?

We acknowledge the China CDC for their valuable assistance in coordinating data collection.

Acknowledgments, Wu-Chun Cao et al..

³¹ Another paper by Chuanhua and his team gives a daily detailed description of the data cleaning process, including avoidance of duplicates. See Figure S1 in Supplement of 'Epidemiological characteristics of patients with severe COVID-19 infection in Wuhan, China: evidence from a retrospective observational study' (Nov 2020), available at

<https://academic.oup.com/ije/article/49/6/1940/5956328?login=false#authorNotesSectionTitle>.

³² Note that 25 Feb 2020 happens to be the exact day a gag order was imposed on the CDC by the National Health Commission (and also the day immediately. The gag order is available at <https://documentcloud.org/documents/7340336-China-CDC-Sup-Regs.html>

Presumably the official version of the events alluded to in the WHO report is that the CDC and the teams working on the retrospective search, data cleaning and entering these cases in the official NNDRS database (such as Chuanhua's team), were grossly incompetent, resulting in a succession of papers and updates with wrong data sent for publication over 2020, including peer-reviewed papers by one of the best specialists in China (Wu-Chun Cao).

All this would have gone on for a while, until the realisation of these damaging errors suddenly hit someone. Whoever then took over the task of cleaning and back-populating the data over from Chuanhua's team and the CDC, may have taken some time and was clearly not finished by August 2020, when 24 supernumerous confirmed cases were still written in the Terms of Reference of the WHO mission (124 confirmed cases, 119 for Wuhan, as per Finding 7).

Confirmed cases		New Confirmed cases			Running Total Confirmed		
From	To	Wu-Chun Cao <i>et al</i>	Total Confirmed (final WHO report)	Diff	Wu-Chun Cao <i>et al</i>	Total Confirmed (final WHO report)	Diff
8 Dec 2019	15 Dec 2019	9	9	0	9	9	0
16 Dec 2019	22 Dec 2019	41	26	15	50	35	15
23 Dec 2019	29 Dec 2019	96	46	50	146	81	65

Table 7: Comparison of running total of Confirmed cases between Wu-Chun Cao *et al.* and final WHO report.

Then eventually the rectification of these numbers would have been so thorough, and a retrospective search over Wuhan so detailed, that when the WHO mission in turn asked in February 2021 for a search of 'all the fever patients, influenza-like illness patients, acute respiratory tract infection patients and unspecified pneumonia patients in all medical institutions in Wuhan' across 233 medical institutions and going back to October 2019, that search did not return a single new 2019 onset case from 76,253 files with possible compatible symptoms³³.

If that version of events is true, we shall logically thank the prescient gag order imposed on the CDC on 25 February 2020 for limiting further damaging spread of incorrect Wuhan COVID-19 numbers for 2019, and most enthusiastically applaud the remarkable skills of the rectification team who did such a thorough job of somehow bringing the number of confirmed cases back from 165 to 100 (a tad under the 104 of the 'CDC Weekly' using data of 11 Feb 2020), and keeping them firmly there.³⁴

³³ As per page 48 of the final WHO report, 76,253 cases of fever, influenza-like illness (ILI), acute respiratory illness (ARI) or unspecified pneumonia across 233 health institutions in Wuhan were identified then reviewed within only a few weeks in Jan-Feb 2021. A shortlist of 92 cases was selected, out of which 67 had serology tests, all returning negative.

³⁴ Interestingly, while he was negotiating the Terms of Reference of the WHO mission in Beijing in July 2020, in a confidential memo that was later leaked, Peter Embarek complained that the Chinese side did not seem to have done much of an epidemiological investigation since January 2020. See "China did 'little' to hunt for Covid origins in early months, says WHO document", Guardian, 23 Feb 2021, available at <https://bit.ly/3TF1SNW>. "Following extensive discussions with and presentation from Chinese counterparts, it appears that little had been done in terms of epidemiological investigations around Wuhan since January 2020. The data presented orally gave a few more details than what was presented at the emergency committee meetings in January 2020. No PowerPoint presentations were made and no documents were shared."

Box 6: Total December 2019 cases as per final WHO report

Final WHO Report:

page 42:

Clinical review of early cases conducted as part of Phase 1 studies

As part of the Phase 1 studies, a review was carried out of all cases reported as potential cases of COVID-19 with onset in December 2019, including all cases that were accepted as formally notified cases in the NNDRS system and other cases that were re-interviewed in December 2020 or January 2021.

Results

A total of **174** cases of COVID-19 were reported to the NNDRS with onset in December 2019: **100** were retrospectively laboratory-confirmed (by sequencing, NAT or serology) cases and a further **74** were clinically diagnosed cases (see Fig. 22). A detailed description of the cases is provided in Annex E2. Other "cases" were identified as part of the search for other potential cases with onset in December 2019 (including some that were included in early publications). After clinical review by the Chinese team, none of the other cases were considered to be compatible with COVID-19 disease, leaving only the **174** notified cases.

page 43:

Cases were scattered by place of residence across the city of Wuhan (164) with a further 10 in seven neighbouring cities.

Annex E2:

page 144:

2.2.1 Wuhan Center for Disease Control and Prevention (Wuhan CDC) searched the NNDRS for COVID-19 cases with the onset date in 2019, and clinical experts were organized to review and discuss the medical records of all cases. Moreover, Wuhan CDC conducted an in-depth epidemiological study of all the cases, mainly focusing on the travel history, occupational exposure, market exposure, animal contact and social contact and other exposure history. Taking into consideration the epidemiological study results and the opinions of clinical experts, **174** cases of COVID-19 were identified with onset date in 2019, including **100** laboratory-confirmed cases and **74** clinically diagnosed cases.

page 146:

3.1 Overview

A total of 174 COVID-19 cases with onset date in 2019 in Wuhan were found in the NNDRS, including 100 laboratory-confirmed cases and 74 clinically diagnosed cases.

A search was conducted on all the fever patients, influenza-like illness patients, acute respiratory tract infection patients and unspecified pneumonia patients in all medical institutions in Wuhan. No suspected COVID-19 cases were found.

Finding 10:

The evolution of the count of 'diagnosed' cases from its peak at the end of February 2020 is not difficult to explain.

'Diagnosed' cases were all reported between the 5 and 17 Feb 2020, which is when Version 5 of the official case definitions applied. Once reported at that time, existing 'diagnosed' cases would still remain in the statistics if (i) *original* samples from the time of hospitalisation were available but had not been tested yet, (ii) *original* samples from the time of hospitalisation were not available and *retrospective* sampling and testing had not been done, and (presumably) (iii) *retrospective* sampling and testing was eventually done, but returned a positive (IgG+), which unfortunately cannot distinguish an infection at the time of hospitalisation from a later infection (possibly asymptomatic).

Since 'diagnosed' became a closed category on 18 Feb 2020, Chuanhua et al. with its 92 diagnosed cases at 24 Feb 2020 is effectively the peak of the number of 'diagnosed' cases. By the time of the final WHO mission (February 2021), a proportion of these 'diagnosed' patients, entered by then a year earlier, would have had no original sample to test and could not be retrospectively contacted for new samples to be taken, and thus had to remain as 'diagnosed'. Or they could be contacted and were retrospectively tested, and came back positive, which meant they had to stay as 'diagnosed' (being indistinguishable from a later infection). And since CT-scans are known to be fairly good diagnostic tools of COVID-19 in the hands of experienced practitioners, a 'diagnosed' case retrospectively tested would indeed have a very good chance of returning a positive.

In other words, most of the 92 'diagnosed' at the 24 Feb 2020 peak should be expected to remain as 'diagnosed' a year later. Hence the official count of 74 diagnosed cases in the final WHO report is consistent with the 92 at the peak. We should note that this consistency largely draws on the fact that 'diagnosed' was a short-lived category closed by 18 Feb 2020, and that this is not applicable to 'confirmed' and 'suspected' cases.

Beyond the evolution of the overall count, we should also mention the 2nd Dec onset diagnosed case that appeared in the first WHO report (see Fig. 6). That case disappears from further counts and is obliquely referred to in the final WHO report:

Final WHO Report, page 46:

Other initially suspected cases in December 2019

Three possible cases with disease onset on 1, 2 and 7 December 2019, respectively, were initially identified as potential cases in the retrospective case search and have been included in some published papers. Clinical review of these three cases by the Chinese expert team led to their exclusion as possible cases on the basis of the clinical features of their illness.

Finding 11:

The disappearance of the 'suspected' Dec 2019 onset cases that were given in the first WHO report (Feb 2020) is not entirely consistent, but more importantly the published count of 'suspected' cases does not tell us much about the actual number of cases that would meet the 'suspected' definition criteria in a proper retrospective search.

The evolution of Dec 2019 onset 'suspected' Dec 2019 cases over time is somehow problematic. But before we delve more into the subject it is important to first understand a few points about 'suspected' 2019 onset cases.

As far as identifying 2019 onset cases is concerned (if necessarily retrospectively), 'suspected' cases played a very important role before RT-PCR tests became available around 10 Jan 2020. Until then 'suspected' was the only category officially available (even if it included a few cases that had been for all purposes already laboratory-confirmed by full sequencing or by a positive test for a SARS-like virus).

From 11 Jan 2020, official reports introduced 'confirmed' cases, but did not immediately reflect the real number of backlog 'suspected' cases waiting to be tested, including most likely quite a few 2019 onset cases (see Box 7). The 'suspected' category nevertheless maintained an important role until the 'diagnosed' category was created on 5 Feb 2020, due to the shortage of RT-PCR tests and the high level of false negatives at the time.

After that, identifying new 'suspected' 2019 cases was actually not the focus of the main retrospective search that took place between 5 Feb and end Feb 2020. The focus was instead very much at the time on the new 'diagnosed' category (until 17 Feb) and on 'confirmed' cases in hospital records.

Step 1: 11 Feb 2020 to 20 Feb 2020				
CDC/WHO 20Feb2020 minus CDC 11Feb2020				
	Confirmed	Diagnosed	Suspected	Total
2-Dec-19		1		1
8-Dec-19				
9-Dec-19				
10-Dec-19		-1		-1
11-Dec-19		-1		-1
12-Dec-19		1		1
13-Dec-19				
14-Dec-19				
15-Dec-19	1	-1		
16-Dec-19			1	1
17-Dec-19	1	1	1	3
18-Dec-19	1		1	2
19-Dec-19				
20-Dec-19		2		2
21-Dec-19		1	-2	-1
22-Dec-19	1	-1	-1	-1
23-Dec-19		5	-1	4
24-Dec-19		6	-2	4
25-Dec-19	1	7	-1	7
26-Dec-19	1	7	-4	4
27-Dec-19	1	7	-3	5
28-Dec-19		2	-1	1
29-Dec-19	1	2	1	4
30-Dec-19		6	-4	2
31-Dec-19	-2	3	-1	
Total	6	47	-16	37

Fig. 11: Differences in Dec 2019 onset cases: CDC/WHO First Report using data as of 20 Feb 2020 (Fig. 2 above) minus CDC report using data as of 11 Feb 2020 (Fig. 2 above).

Accordingly, we do not see the count of 'suspected' 2019 onset cases growing during the February retrospective search. On the contrary we see an attrition in the number of 'suspected' 2019 cases between the 'CDC Weekly' update (data as of 11 Feb) and the first WHO report (data as of 20 Feb 2020) as per Fig 11.

A main source of attrition of 'suspected' 2019 cases in Feb 2020 would be through testing of original or later samples: 'suspected' cases that were tested from original 2019 samples (extending probably to the first week of January) and returned a positive could be upgraded to 'confirmed' and those returning a negative could be dropped altogether, while 'suspected' Dec 2019 that were tested from later samples and returned a negative could be dropped altogether.

Another source of attrition available between 5 Feb and 17 Feb 2020 (when the 'diagnosed' category was available) is the conversion of 'suspected' cases to 'diagnosed' ones, for those who had consistent lung imaging symptoms.

We note that as of 20 Feb 2020 there are still 9 suspected 2019 onset cases as per the first WHO report (Fig. 6). However there is soon after no mention at all of any 'suspected' 2019 cases, in papers or in the final WHO report of March 2021.

To a certain extent this is an issue, as it is difficult to imagine that these 9 suspected cases of 20 Feb 2020 were later all converted to 'confirmed' (this requires original samples, somehow identified rather late, and all being positive), or all dropped for having retrospectively tested negative from original or more likely later samples, a result (0 positive out of 9 'suspected') that would contrast strongly with the 41 out of 59 'suspected' that returned positive test on 11 Jan 2020.

So overall we can expect a large decrease over time in retrospective suspected cases, but not necessarily a full disappearance. The final WHO report does not offer any explanation for that disappearance, but while this may seem puzzling, for all purposes we are talking of 9 cases at most, which in the greater scheme of things is not essential.

It is much more important to acknowledge that the retrospective search between 5 Feb and end Feb 2020 never focussed on identifying new 'suspected' 2019 cases, and that that category is not able to tell us much in that context. Its absence from further counts after the first WHO report is therefore not surprising.

Consequently, the insistence of the WHO team in January 2021 for their Chinese counterparts to retrospectively do a broad search of 'suspected' 2019 cases makes a lot of sense, and is worth looking into in detail.

Finding 12:

The total absence of positive tests amongst 67 retrospectively shortlisted patient files in January 2021, tested at that time, makes little sense.

At the request of the WHO team, which met 3 or 4 times on video conference calls with its Chinese counterpart before heading to China, in January 2021 the Wuhan Health Commission retrospectively tested the sera of 67 suspected COVID-19 cases, with symptoms onset from mid Sep 2019 to the 8 Dec 2019.³⁵ These 67 patients were from a short list of 92 which Chinese authorities had put together from 76,253 episodes with a illness compatible with COVID-19 in the last quarter of 2019 across 233 medical institutions in Wuhan³⁶. 25 of these 92 were unreachable, were unwilling to contribute samples or had died, leaving the 67 cases that could be retrospectively tested.

All 67 tests came back negative, which raises at least two issues:

1. Faced with the unavailability of original samples, which according to Chinese authorities were not kept³⁷, only retrospective testing one year later was possible. This obviously raises the possibility of antibodies being undetectable by that time. In which case the whole elimination of these 'suspected' cases based on negative results is totally unsound.

This was duly mentioned as a limitation in the WHO report and also mentioned by one of the mission members in an interview:

page 49 of WHO report:

Following review by the health institutions, only 92 cases of the 76 253 episodes were considered to have an illness clinically compatible with SARS-CoV-2 infection. [...]. **Following further review by the external multidisciplinary clinical team, all these cases were assessed not to be cases of SARS-CoV-2 infection.**

The 92 cases were followed up in January 2021 and **blood for SARS-CoV-2 serology collected from 67 of them** (the remainder either having died, refused or were unobtainable). **All 67 sera were reported to be SARS-CoV-2-specific antibody negative.**

Conclusions and limitations

The retrospective search for cases compatible with COVID-19 illness identified 76 253 episodes with one of four indicator conditions. A rise in one of these conditions, ARI (as well as ILI and fever), was seen in this group of individuals in the over-60-year age group in early December. The clinical assessment of the 76 253 individuals revealed 92 cases clinically compatible with COVID-19. It is possible that the application of stringent clinical criteria, resulting in the identification of only 92 clinically compatible cases, may have decreased the possibility of identifying a group or groups of cases with milder illness.

All the 92 cases were rejected as cases of SARS-CoV-2 infection on further clinical review. None of these cases (where blood could be obtained) was positive on SARS-CoV-2 serological testing carried out more than 12 months later. **The use of retrospective serological testing so long after the illness cannot be relied on to exclude the possibility of SARS-CoV-2 infection at the time of the presenting illness, given the possible drop in SARS-CoV-2-specific antibody over time and the associated reduced sensitivity of commercial assays. The possibility that earlier transmission of SARS-CoV-2 infection was occurring in this community cannot be excluded on the basis of this evidence.**

³⁵ See 'Annex A2 - Schedule of Work' of the WHO report for the detail of that work in January 2021 and the questions that the WHO mission had about the methodology, ahead of its arrival in Wuhan,

³⁶ For the mention of 233 medical institutions, see page 7 of the final WHO report, or page 145 of the Annexes.

³⁷ As per WSJ article of 21 Feb 2021, ['China Refuses to Give WHO Raw Data on Early COVID-19 Cases'](#), "Dr. Dwyer said Chinese authorities had provided influenza surveillance data from before December 2019 but only from one children's hospital and one general hospital. The authorities told the WHO team that its hospitals generally didn't store physical samples from patients with respiratory diseases.

"They said they were destroyed and so on. You know, I guess one accepts that on face value," Dr. Dwyer said.'

NY Times, [‘On W.H.O. Trip, China Refused to Hand Over Important Data’ 12 Feb 2021](#):

The W.H.O. scientists were frustrated by the Chinese government’s reluctance to explain how they had gathered the data, according to interviews with team members.

Dr. Fischer said she would have expected to find many more cases of individuals who were hospitalized with such symptoms in a city the size of Wuhan.

In heated discussions, Dr. Fischer recounted, the W.H.O. experts urged the Chinese scientists to conduct a more thorough search. **The team also expressed concerns about the reliability of antibody tests administered so long after the infections. Testing any original nose or throat swabs would be useful, but Dr. Dwyer said there were none.**

2. If we suppose that the tests are reliable, then given a background Wuhan urban infection rate of around 4% over April-May 2020³⁸, the chance of returning no positive at all over 67 tests in January 2021 is roughly 1 in 15 (around 6.5%), which is rather low³⁹.

Incidentally, had the Chinese team agreed with the WHO demand and tested a larger group of ‘suspected’ cases⁴⁰, then the probability of returning zero positive (supposing meaningful testing one year later) would have quickly converged towards extremely low values. For instance, using the same 4.0% background infection rate for April-May 2020, the probability of returning all negatives out of 120 cases tested is less than 1%, and is less than 0.1% for 170 cases tested - both totally reasonable, if not still low, populations to test.

The fact that the Wuhan Health Commission is the entity that performed these tests⁴¹ is also intriguing. The Wuhan Health Commission is well documented as having fought the CDC in the early days of the outbreak to avoid having to report cases, going as far as deleting cases entered in the NNDRS by doctors from Zhongnan hospital⁴² on 9 Jan 2020. Also on 11 Jan 2020, the Wuhan Health Commission ignored the strongly voiced advice of the CDC and did not include 50 to 100 ‘suspected’ cases in its announcement on the outbreak situation that day (see Box 7), when it instead focussed on a reduction in the number of COVID-19 from 59 ‘suspected’ to 41 ‘confirmed’ after 18 patients returned negative RT-PCR tests. These never previously disclosed 50-100 suspected cases simply remained undisclosed, creating a false impression of an outbreak coming under control.

In conclusion, either the retrospective all-negative result for these 67 ‘suspected’ cases is rather implausible (only around 6.5% chance), or the tests are unreliable.

³⁸ We refer to a cross-sectional seropositivity study in April-May 2020 of around 35,000 Wuhan individuals, which found a 95% confidence interval of [4.0%, 4.8%] for the Wuhan urban seropositivity rate: [‘Association of Public Health Interventions With the Epidemiology of the COVID-19 Outbreak in Wuhan, China’](#)

³⁹ Using a binomial distribution which is reasonable for a well spread selection of 67 cases across so many medical institutions in Wuhan.

⁴⁰ See NY Times, [‘On W.H.O. Trip, China Refused to Hand Over Important Data’ 12 Feb 2021](#) quoted earlier.

⁴¹ For the mention of the Wuhan Health Commission being responsible for the testing, see section 2.2.2, page 144 of the Annexes of the WHO report.

⁴² The Wuhan Health Commission likely deleted two cases entered on the 9 Jan 2020, by doctors at Zhongnan hospital, a top tier research hospital linked to Wuhan University, given a confirmation of a CDC insider that the CDC did not delete them. These doctors had also faced much pressure not to report. [‘Zhongnan Hospital physicians in the Wuhan Epidemic: There were sick people everywhere’](#) (Caixin).

Box 7: The unreported 'suspected' cases on 11 Jan 2020

Extract from '[For Whom the Warning System Tolls](#)' (Jan 2021):

A Key Meeting

During the initial investigations, a crucial meeting took place at the Wuhan Municipal Health Commission on January 10. The meeting, organized by the provincial health commission, included the second batch of experts, as well as provincial and municipal health officials, members of local CDC departments, and doctors and virologists from Wuhan.

Dr. Zhang Qi [pseudonym] participated in this meeting. That night, he explained, there was controversy regarding which cases to announce. Wuhan provided case data: 41 cases confirmed by laboratory test results and over 100 suspected cases that had not yet undergone testing.

The expert group demanded that both suspected and confirmed cases be reported and also asked that suspected patients be treated in isolation. Yet the next day, the local government announced only the 41 confirmed cases.

For Dr. Zhang Qi, this meeting was a missed opportunity, one that he will never forget. It showed that experts had lost against administrative power. Public health workers make decisions based on science, but science is only one consideration for government officials, and definitely not their primary concern, in general, as well as at that time..

Extract from '[Interview with the second batch of experts sent to Wuhan by the Health and Welfare Commission: why no human-to-human transmission was found?](#)' (Caixin, 28 Feb 2020, [automated translation](#)):

Caixin: Did Wuhan hear your suggestions and opinions?

Expert: After the pathogen was found, before releasing the news, the members of the expert group and the local government held a meeting. What we actually discussed is, how many cases are there? Of the case data provided by Wuhan, 41 cases were confirmed by laboratory test results. In addition to this batch of cases, there is also a group of suspected cases that have not been laboratory tested.

What kind of cases were published was controversial at the time. Our expert group agreed that all suspected and confirmed should be reported, and we agreed before leaving. But it wasn't like that when I saw the newspaper the next day. When the news came out, 41 cases were reported by the local government, which were only a group of people confirmed by laboratory methods. I don't understand what's going on behind the scenes.

(Editor's Note: The Wuhan Municipal Health Commission issued a notice on January 11 stating that after the initial determination of the "unknown cause of viral pneumonia" pathogen as a new coronary virus, the Wuhan Health Commission organized⁴³ the testing of existing patient specimens. At 2400 on January 10, there were 41 cases of pneumonia initially confirmed with new coronavirus infection, including 7 cases with severe illness and 1 case, and stable condition.)

Caixin: How many patients were suspected in the cases you saw at the time?

Expert: I can't remember the specifics. To be sure, the number of suspected cases I saw at the time was greater than the number of confirmed cases.

⁴³ These samples were sent to the "National Health Commission Key Laboratory of Systems Biology of Pathogens and Christophe Mérieux Laboratory" (Beijing) for testing.

Finding 13:

A careful examination of the count of 'confirmed' + 'diagnosed' cases over time shows that the data that was communicated to the WHO for its final report (the 174 'confirmed' + 'diagnosed' cases for Wuhan) is consistent with the cases population in the CDC NNDRS at around 15 Feb 2020, so fairly early in the period of active retrospective identification of 2019 onset cases.

Putting all the numbers together from the various sources cited so far, plus a few more minor ones, and ordering the data by date of extraction from the official database (NNDRS), we obtain Table 8 below.

Research by DRASTIC - Contact: gilles@demaneuf.com

	Source	Data up to (DB allowing)	Extracted	Disclosure	Status	Source	Deaths	Total C+D 2019	Confirmed 2019	Diagnosed 2019	Suspected 2019	Count type
Official	#1	2 Jan 2020	2 Jan 2020	3 Jan 2020	official	WHO / China		41		41		Pneumonia of Unknown Etiology A few confirmed by SARS testing, some by NGS sequencing pointing to new virus
	#2	5 Jan 20	5 Jan 20	5 Jan 20	semi-official, key paper	WHO / China		59		59		Pneumonia of Unknown Etiology A few confirmed by SARS testing, some by NGS sequencing pointing to new virus
	Genome released (10 Jan)											
	#3	10 Jan 2020	10 Jan 2020	11 Jan 2020	confirmed from major paper, leaks for suspected	Lancet / WHO / China		41	41		[50-100] not disclosed	Confirmed via just released new test kit for SARS-CoV-2. CDC wanted to include over 100 suspected too, but Wuhan Health Commission refused
	#4	11 Feb 2020	< 17 Feb 2020	17 Feb 2020	official - from CDC report	CDC	15	141	104	37	25	Confirmed, suspected and diagnosed cases with onset in Dec 19
	#5	20 Feb 2020	< 28 Feb 2020	28 Feb 2020	semi-official - from WHO report graph	WHO report / China		192	109	83	9	Confirmed, Suspected, diagnosed onset 2019
Leaks	#6	24-Feb-20	< 16 Oct 2020	21 Dec 2020	reliable - paper from national cases data team	Papers (Chuanhua Yu / CDC data)	33	227	135	92		33 deaths, 29 from confirmed, 4 from clinically diagnosed
	25 Feb: Gag Order imposed on CDC by the Wuhan Health Commission											
	#7	27-Feb-20	27 Feb 2020	29 Sep 2020	paper using official DB with CDC help	Paper (CDC data)		~257 Dec only	146 to 29 Dec	[92]?	?	2 Health Care Workers, 144 patients, all from 8 to 29 Dec 2019. Add 19 for 30/31 Dec (from final WHO report) for total 165 lab-confirmed Dec 2019 onset
Official	27 Feb: Prof Yu Chuanhua retracts Nov cases he mentioned in the Health Times one day earlier, citing the NNDRS as of 25 Feb											
	#8	8-Mar-20	9 Mar 2020	10 Apr 2020	paper, from graph	Paper (CDC data)		197	115	82	?	laboratory-confirmed
	#9	not clear, but likely end Feb 2020	< 1 Dec 2020	1 Dec 2020	leak of official CDC data, checked via digital forensic	CNN / Hubei CDC		> 200	~200		?	"In the bottom left hand column of the graph marked 2019 the number of "confirmed cases" and "clinically diagnosed" cases appears to reach around 200 altogether"
	#10	< 13 Mar 2020	< 13 Mar 2020	13 Mar 2020	leaked 'government data', likely CDC	SCMP/ gov. source		266	266		?	Total 'confirmed' cases as of 31st Dec 2019, 9 Nov 19 cases, so 257 Dec 19 cases. + 115 on 1 Jan for a total of 381.
Official	#11	< 31 Jul 2020	< 31 Jul 2020	5 Nov 2020	semi-official (WHO TORs)	WHO / China		124	124	?	?	Confirmed cases with onset in Dec 19, + 5 For Hubei and other provinces (124)
	#12	< 20 Feb 2021	< 20 Feb 2021	30 Mar 2021	official and final	WHO report / China		174	100	74	0	100 Confirmed + 74 diagnosed cases with onset in Dec 19, 164 in Wuhan city, 10 in neighbouring cities

Table 8: Full breakdown of 'confirmed', 'diagnosed' and 'suspected' numbers over the main sources, by **data extraction** date.

By focussing first on 'confirmed' cases, which cannot be retracted unless there is some double-counting or mis-testing and should thus only show a gradual increase, we can see that the count of 100 'confirmed' cases for Wuhan and neighbouring cities in the final WHO report is most consistent with the count in the NNDRS database on 11 Feb 2020, given by the 104 'confirmed' of the 'CDC Weekly' update extracted on that day.

For the reasons mentioned in Finding 10, the count of 'diagnosed' cases is subject to both entries (including transfers from 'suspected' between the 5 and 17 Feb) and exits, and is thus much less stable over time. However the total of 'confirmed', 'diagnosed' and 'suspected' should be more stable as it neutralises all internal transfers. On that basis, the total of 174 cases in the final WHO report seems to be halfway between the 'CDC Weekly' update with its data as of 11 Feb, and the data as of 20 Feb 2020 given by the first WHO report.

Overall, one can therefore summarise this as saying that the final WHO count of 100 confirmed and 74 diagnosed Dec 2019 onset cases for Wuhan and neighbouring cities seems to reflect the state of the CDC NNDRS database at around 15 Feb 2020.

We can also appreciate this by looking at cumulative curves per Dec 2019 onset dates and the individual counts on each day.

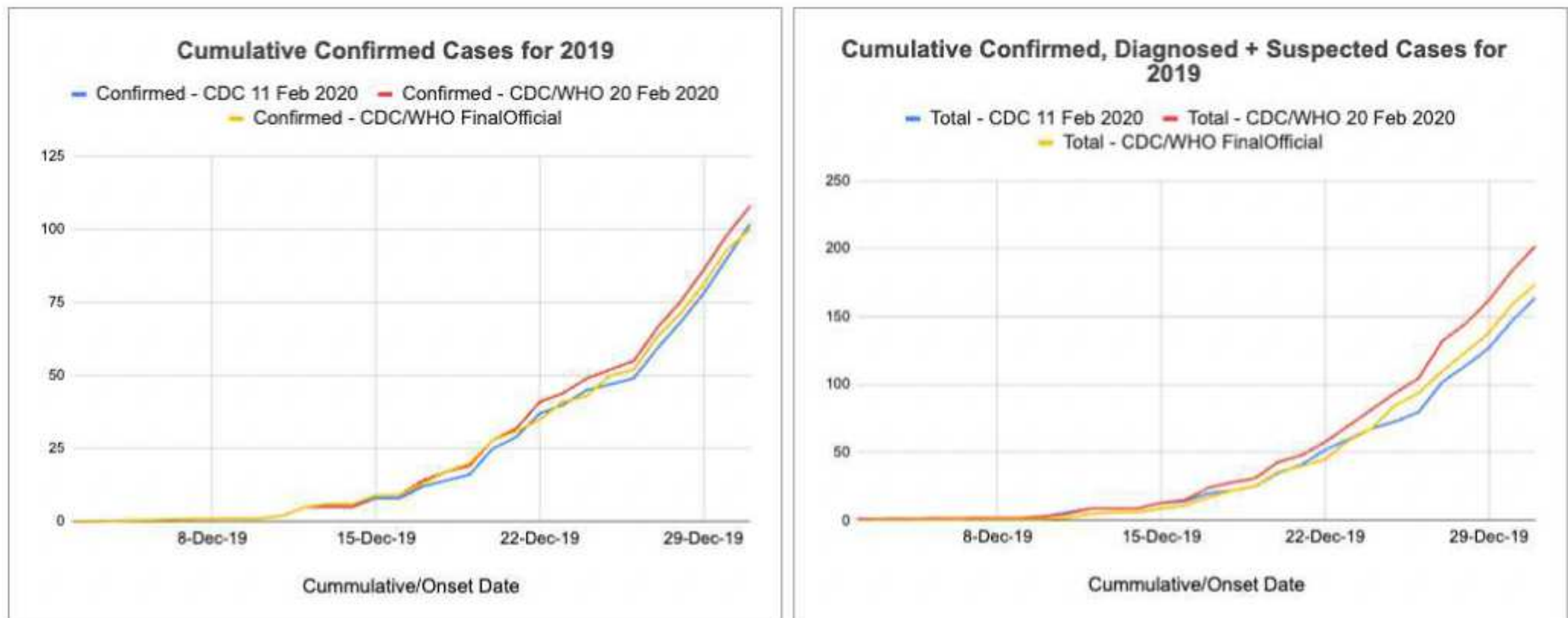


Fig. 12: Cumulative counts of 2019 onset cases, Confirmed on left, Confirmed + Diagnosed + Suspected on right.

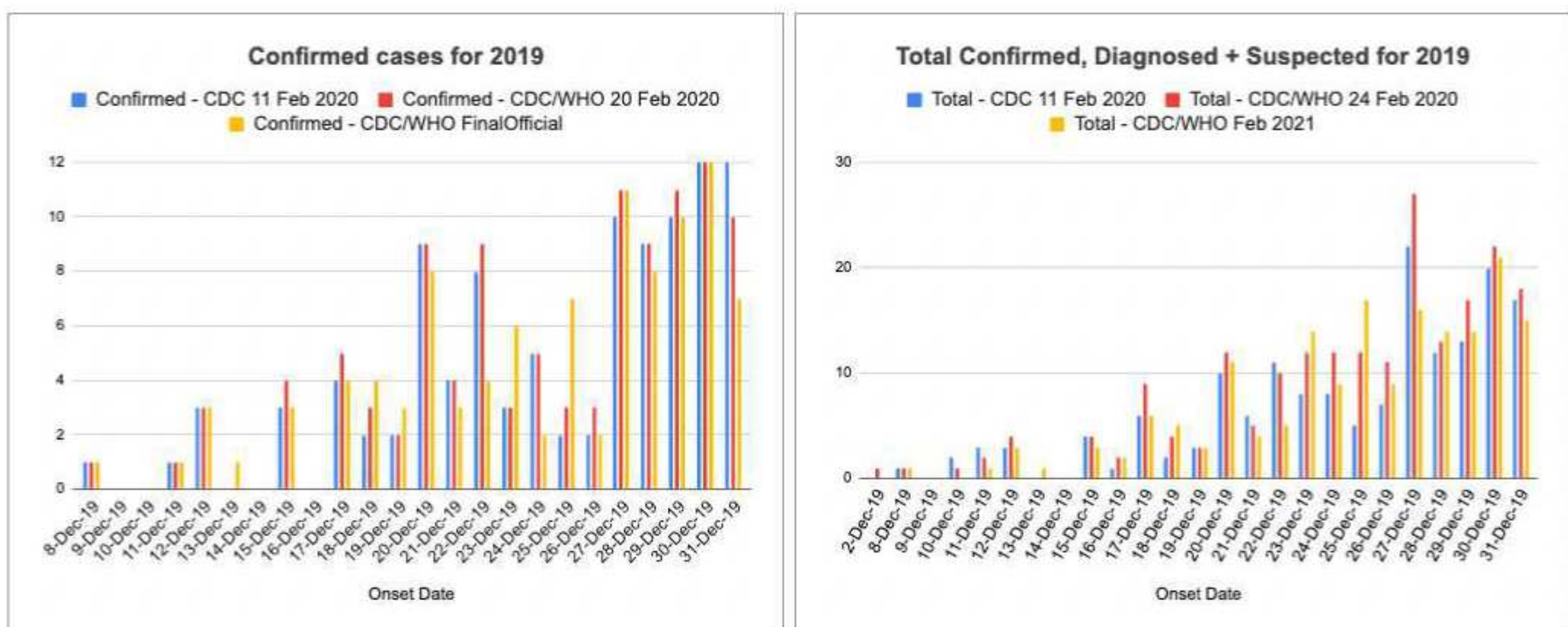


Fig. 13: Counts of 2019 onset cases, Confirmed on left, Confirmed + Diagnosed + Suspected on right.

Table 8 also shows us how the count of confirmed and diagnosed cases kept growing quickly during the retrospective search for diagnosed cases from 5 to 17 Feb 2020, and then beyond until 27 Feb 2020 (confirmed only) when it suddenly stopped, just after the National Health Commission issued a gag order on the CDC.

After 27 Feb 2020, the few official numbers published are quickly retracing that growth, even in terms of normally unalterable confirmed cases which peaked at 165 around 27 Feb 2020, to go all the way back to 100 in the final WHO report.

The only apparent exception to that sudden retracing remains the CDC leak published on 13 Mar 2020 by the South China Morning Post, which perfectly matches our necessary range of 247 to 260 Dec 2019 cases at 27 Feb 2020 based on peer-reviewed papers and the final WHO report. We do not have the extraction date for that leak, but it seems reasonable that it was just after the 27 Feb 2020, so that it could eventually be sent to the

SCMP where it would have taken a few days for it to be published on 13 Mar 2020. In other words it is most likely a confirmation, and a rather motivated one, instead of an exception.

Box 8: The gag orders

1. Notice on ‘Strengthening Management of Scientific and Technological Achievements During the Prevention and Control Period of the Novel Coronavirus Pneumonia Epidemic’

On 24 Feb 2020, the day the first WHO mission presented its conclusions in a press conference in Beijing⁴⁴, the Chinese CDC issued an internal notice⁴⁵ that in particular regulated any CDC publication, under “important instructions” from Chinese President Xi Jinping”. That notice on that day may read as a well-meaning attempt to strengthen the quality of publications.

2. Supplementary Regulations on ‘Strengthening Management of Scientific and Technological Achievements During the Prevention and Control Period of the Novel Coronavirus Pneumonia’

The next day (25 Feb 2020), when the WHO mission was on its way back home, a much more strongly worded supplement to the previous day notice was published⁴⁶, with explicit threats of punishment for non compliance. That supplement articulates additional regulations.

In particular it states that:

- The CDC priority should not be ‘on publishing papers until the epidemic is under control’.
- Going forward, any new research project must first be validated by the National Health Commission.
- Nobody within the CDC can share information or data on the epidemic, or samples.
- The publication of any research paper or results by the CDC must first be approved by the National Health Commission.
- Papers that were already submitted for publication in journals must be urgently reviewed, modified if required as per guidelines above and then seek approval by the National Health Commission before being published.
- ‘Anyone who violates the above regulations shall be dealt with severely in accordance with discipline, law and regulations’.

3. Notice on the Standardisation of the Management of Publication of Novel Coronavirus Pneumonia Scientific Research

Finally a few days later, on 3 Mar 2020, a much wider and all-compassing confidential notice⁴⁷ was issued by the State Council. That notice stipulated similar regulations as the supplementary notice of 25 Feb 2020, but this time applied to all research and data in China relative to COVID-19, whatever the institution:

[A]ll publication work on epidemic prevention research and information related to COVID-19, including medication, vaccines, virus origins, virus transmission routes, testing reagents, etc. will be taken over by the Joint Prevention and Control Mechanism of the State Council’s scientific research group (hereinafter referred to as “the scientific research group”) for coordinated deployment.

Most importantly, it added a determined communication strategy dimension:

⁴⁴ See WHO briefing at <https://www.who.int/docs/default-source/coronaviruse/transcripts/joint-mission-press-conference-script-english-final.pdf> and the full report at <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>

⁴⁵ Notice available at <https://www.documentcloud.org/documents/7340335-China-CDC-Sci-Tech.html> and leaked to Associated Press, which wrote about it in this important article: <https://apnews.com/article/pandemics-china-bats-coronavirus-pandemic-fe870dd24e2b33468ba409bd806df2fe>

⁴⁶ Also leaked via Associated Press and available at <https://www.documentcloud.org/documents/7340336-China-CDC-Sup-Regs.html>

⁴⁷ Also leaked via Associated Press and available at <https://www.documentcloud.org/documents/7340337-State-Research-regulations.html>

During the epidemic prevention and control period, all localities and units must conscientiously implement the spirit of a series of important instructions on epidemic prevention and control by General Secretary Xi Jinping, further strengthen overall awareness and sense of responsibility, strengthen reviews and checks, and actively communicate and coordinate the publication of scientific research on COVID-19 in a pattern like “**moves in a game of chess**” across the country. Important and sensitive scientific research results must be repeatedly reviewed and checked, and if they are not certain, procedures should be followed to request instructions from the scientific research team and related departments in a timely manner.

Last, it again reiterated that non-compliance would be punishable.

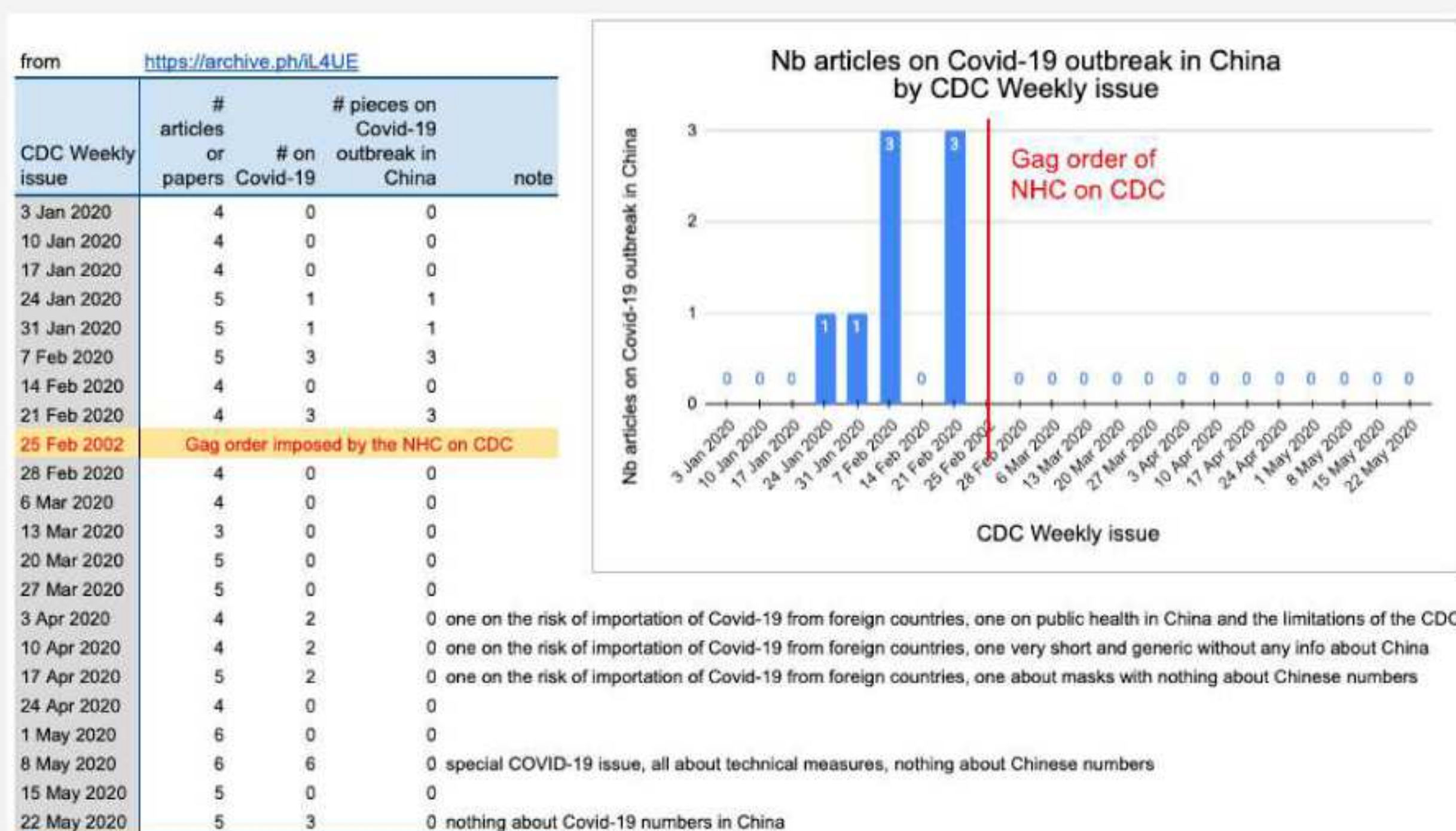
4. Concrete example

As a very concrete example of the impact of the gag order on the Chinese CDC, we checked the number of articles in the ‘CDC Weekly’ journal that gave valuable information about the COVID-19 situation in China, such as actual case numbers or any information that may offer a clue as to the origins (beyond purported foreign imports).

Each ‘CDC Weekly’ issue compiles communications published by the Chinese CDC during the week, with typically 3 to 5 articles in all per issue. For instance, the key update on the case numbers published on the CDC website on 17 Feb (using data as of 11 Feb) was included in the ‘CDC Weekly’ issue of 21 Feb.

We found 8 relevant articles prior to 25 Feb 2020, spread over 4 issues⁴⁸, with the ‘CDC Weekly’ issues of the 7 Feb and 21 Feb being particularly informative.

Following the gag order of 25 Feb 2020, for many weeks not a single relevant article was published. All articles touching to COVID-19 after that date either focussed on the risk of import via foreign travellers or were describing technical measures that provided no relevant insight.



Count of articles about COVID-19 situation in China in ‘CDC Weekly’

5. A natural order

The gag order that the National Health Commission imposed on the CDC on 25 Feb 2020 can be seen as only restoring the natural order between the two: the NHC administratively oversees the CDC which is effectively just a decision support agency under the ‘Law on Prevention and Control of Infectious Diseases’.

That law in particular means that the CDC has no power of decision, being just a decision support agency. That power instead lays with the responsible administrative authority/ies, which may include the city, provincial or national health commission, all with a political dimension. In particular the decision to announce

⁴⁸ See <https://archive.ph/iL4UE> for all the issues and their content.

an outbreak, and the publication of outbreak numbers, were always meant to be the responsibility of such authorities only⁴⁹.

In that sense, even the key 'CDC Weekly' update published online on 17 Feb 2020, that gave a first detailed count of cases (including amounts Healthcare Workers), was an exception to the rule.⁵⁰

In the same way, in the early days of the outbreak and until around 15 January 2020, when the first version of the case definitions was published at the instigation of the National Health Commission and some expert groups of frontline medical professionals from top Wuhan universities and hospitals, the definition and reporting of cases had been a battlefield between the Wuhan Health Commission and the Chinese CDC. Acting only as a decision support agency to the Wuhan Health Commission, the Chinese CDC could only argue with, but not stop the decisions of the Wuhan Health Commission⁵¹.

Finding 14:

The retrospective search for 2019 onset cases was at its peak between the 9 and 25 February 2020 and involved Prof. Chuanhua's team at Wuhan University working with the CDC. That search and the resulting identification of 2019 cases seems to have totally stopped soon after the gag order imposed on the CDC.

On 17 Jan 2020 the Chinese authorities started again announcing new cases after a total reporting lull since 5 Jan 2020 (punctuated only by a reduction in reported cases on 11 Jan). Shortly after, on 20 Jan, Chinese authorities admitted that human-to-human infection may be an issue, leading to the decision to lockdown Wuhan on 23 Jan. It was only then that it became clear that precise tracking of cases would be necessary to fight what was by now officially a true epidemic.

For instance on 29 Jan, the Hubei Provincial Committee of the China Democratic League submitted a straight-through proposition "to establish a database of confirmed cases of new coronary pneumonia to support precise prevention and control of the epidemic" to Ying Yong, the Hubei Provincial Party Secretary, who approved it. The suggestion came from Prof. Feng Dan, dean of the School of Computer Science of Huazhong University of Science and Technology (Wuhan)⁵².

Soon after, an effort involving the Chinese CDC and external expert teams was under way. One of these expert teams, and possibly the main one as we cannot find explicit mentions of any other, was the team led by Prof. Yu Chuanhua (宇传华), vice president of the Hubei Health Statistics and Information Society and professor of epidemiology and health statistics at Wuhan University.

In one of the resulting papers that Prof. Chuanhua's team published, the contribution of that team is clearly described:

Our team was among those who participated in obtaining the original uncleared data from the designated hospitals during the epidemic; cleaning, shortening and analyzing the original data; sending the data to the CDC and the infectious-disease-reporting information system; and working with CDC staff in ensuring the quality of the data. We were given official approvals for monitoring the data of registered patients with

⁴⁹ See for instance Zhong Nan Shan's analysis on 27 Feb 2020:

<https://m.21jingji.com/article/20200227/herald/e65028dbe95d2974b861cf980a546d3f.html>, with [translation](#).

⁵⁰ 'The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19)'.

Available the same day in English on 'CDC Weekly' website at

<https://weekly.chinacdc.cn/en/article/doi/10.46234/ccdcw2020.032> and in Chinese in the Chinese Journal of Epidemiology at <https://archive.ph/n7xs7>.

⁵¹ One clear example was the refusal of the Wuhan Health Commission on the 11 Jan 2020 to include 50 to 100 suspected cases in the official announcement in which the National Health Commission reported a reduced number of cases from 59 to 41 after laboratory confirmation (see <https://archive.ph/pgHiD> and <https://finance.sina.com.cn/wm/2020-02-26/doc-iimxxstf4577235.shtml>, with [translation](#)). See also Box 7.

⁵² See 'Need to establish a database of confirmed cases of new coronary pneumonia to support precise prevention and control' available at <https://archive.ph/pbeQg>

severe and critical COVID-19 illness in Wuhan until 25 February 2020 (i.e. the final date of follow-up in this study).⁵³

Practically Prof. Chuanhua's team got official access to relevant medical files then started applying the case definitions criteria available at the time (going through at least two changes in the process). In doing so he did not only focus on the main objective behind allowing diagnostics via lung imaging on the 5 February (introduced in version 5 of the cases definitions), which was to get a better handle on the ongoing sanitary emergency at the time by allowing proper isolation at home or in hospital setting; his team and the CDC also focused on 2019 onset cases, thus demonstrating what looks like a genuine attempt to do proper epidemiological work.

Despite the earlier failure of the CDC NNDRS system to facilitate the detection of the outbreak, the NNDRS database was still likely the best place to keep recording new and retrospective cases. Additionally the NNDRS system itself already reflected the approved mode of interaction between the CDC and local and national authorities, from the district to the National Health Commission via the Wuhan Health Commission (See Box 4). The NNDRS database, under the control of the party secretary of the CDC and not of the head of the CDC (George Gao), thus became de facto the official database of cases that the situation required.

Form the papers published by Prof. Chuanhua and his team, one can see that the retrospective review and entering of 2019 onset cases into the NNDRS was at its peak around the 9 to 25 February 2020, with no indications from his papers and other literature of additions of 2019 onset cases after 27 Feb 2020, except possibly for the November 2019 cases mentioned shortly after in the SCMP leak of CDC data on the 13 March 2020.

In particular a figure in the supplement of one of Chuanhua's team papers⁵⁴ provides a way of obtaining the range of reporting dates per onset dates starting in the last few days of 2019:

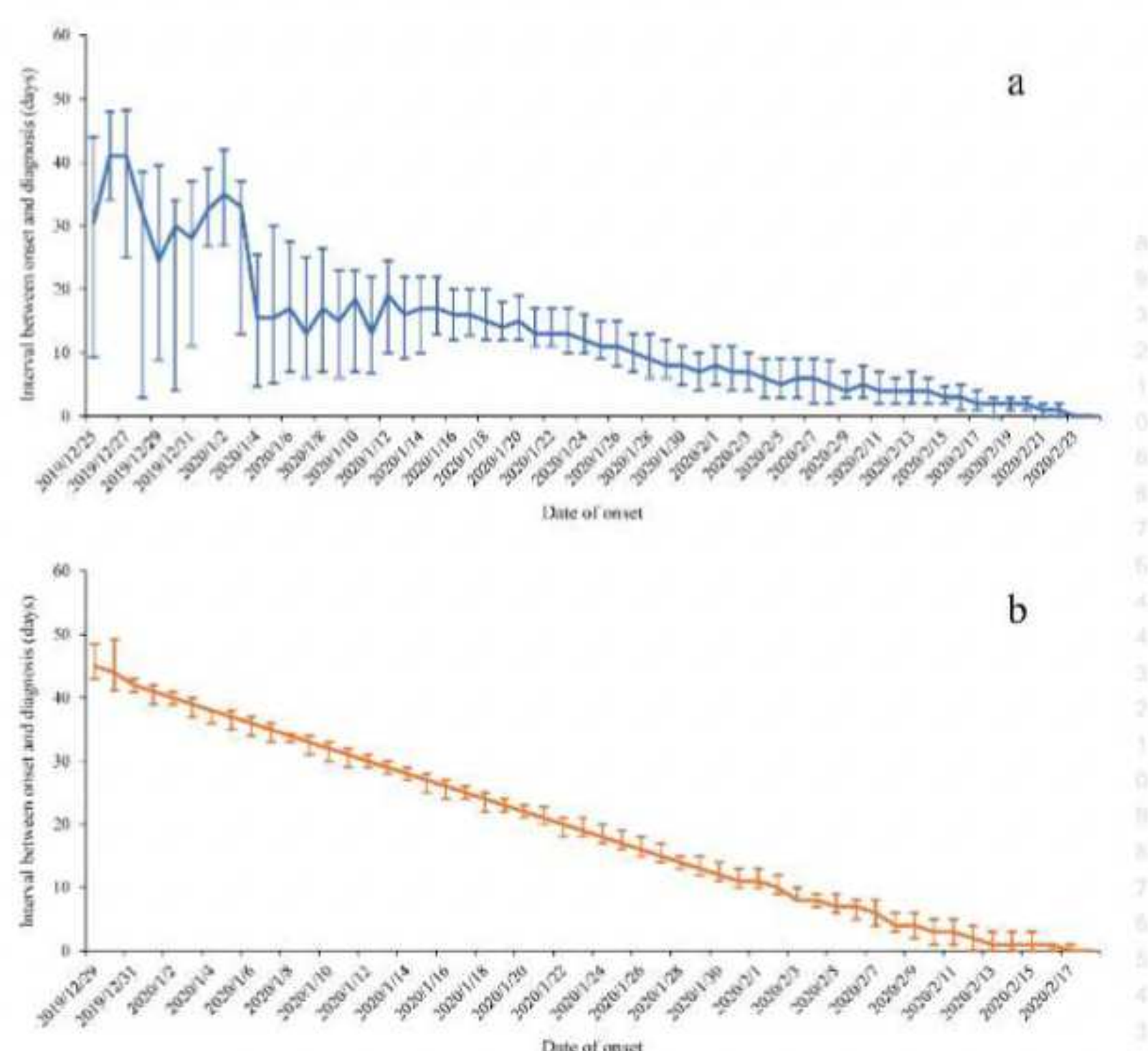


Fig. S2 Time trend of interval between onset and diagnosis. (a) confirmed cases, (b) clinically diagnosed cases

⁵³ 'Epidemiological characteristics of patients with severe COVID-19 infection in Wuhan, China: evidence from a retrospective observational study', published online Nov 2020 in International Journal of Epidemiology, available at <https://academic.oup.com/ije/article/49/6/1940/5956328?login=false#authorNotesSectionTitle>

⁵⁴ 'The comparison of epidemiological characteristics between confirmed and clinically diagnosed cases with COVID-19 during the early epidemic in Wuhan, China', May 2021, available at <https://ghrp.biomedcentral.com/articles/10.1186/s41256-021-00200-8#article-info>, Additional File at https://static-content.springer.com/esm/art%3A10.1186%2Fs41256-021-00200-8/MediaObjects/41256_2021_200_MOESM1_ESM.docx

Fig. 14: Fig. S2 of the Additional File to Chuanhua et al. (May 2021).

Inverting the supplementary Fig. S2 of Chuanhua *et al.* (May 2021), we obtain the following ranges of reporting dates by onset date:⁵⁵

not in CDC Weekly update of 17 Feb (data up to 11 Feb), which had 104 confirmed cases and 37 diagnosed for Dec 19 as above, and more particularly not included in the Dec 19 onset cases of the CDC Weekly update					
first date: 9 Feb 2020 last date: 18 Feb 2020			first date: 31 Dec 2019 last date: 25 Feb 2020		
Clinically Diagnosed			Confirmed		
onset	diag. from	diag. to	onset	conf. from	conf. to
29-Dec-19	10-Feb-20	15-Feb-20	25-Dec-19	3-Jan-20	7-Feb-20
30-Dec-19	9-Feb-20	17-Feb-20	26-Dec-19	29-Jan-20	12-Feb-20
31-Dec-19	10-Feb-20	12-Feb-20	27-Dec-19	21-Jan-20	13-Feb-20
1-Jan-20	9-Feb-20	12-Feb-20	28-Dec-19	31-Dec-19	4-Feb-20
2-Jan-20	10-Feb-20	12-Feb-20	29-Dec-19	7-Jan-20	6-Feb-20
3-Jan-20	9-Feb-20	12-Feb-20	30-Dec-19	3-Jan-20	2-Feb-20
4-Jan-20	9-Feb-20	11-Feb-20	31-Dec-19	11-Jan-20	6-Feb-20
5-Jan-20	9-Feb-20	12-Feb-20	1-Jan-20	28-Jan-20	9-Feb-20
6-Jan-20	9-Feb-20	12-Feb-20	2-Jan-20	29-Jan-20	13-Feb-20
7-Jan-20	9-Feb-20	12-Feb-20	3-Jan-20	16-Jan-20	9-Feb-20
8-Jan-20	10-Feb-20	11-Feb-20	4-Jan-20	9-Jan-20	29-Jan-20
9-Jan-20	9-Feb-20	12-Feb-20	5-Jan-20	10-Jan-20	4-Feb-20
10-Jan-20	9-Feb-20	12-Feb-20	6-Jan-20	13-Jan-20	2-Feb-20
11-Jan-20	9-Feb-20	12-Feb-20	7-Jan-20	13-Jan-20	1-Feb-20
12-Jan-20	10-Feb-20	12-Feb-20	8-Jan-20	15-Jan-20	3-Feb-20
13-Jan-20	10-Feb-20	12-Feb-20	9-Jan-20	15-Jan-20	1-Feb-20
14-Jan-20	10-Feb-20	12-Feb-20	10-Jan-20	17-Jan-20	2-Feb-20
15-Jan-20	9-Feb-20	12-Feb-20	11-Jan-20	18-Jan-20	2-Feb-20
16-Jan-20	9-Feb-20	12-Feb-20	12-Jan-20	22-Jan-20	5-Feb-20
17-Jan-20	10-Feb-20	12-Feb-20	13-Jan-20	22-Jan-20	4-Feb-20
18-Jan-20	9-Feb-20	12-Feb-20	14-Jan-20	24-Jan-20	5-Feb-20
19-Jan-20	10-Feb-20	12-Feb-20	15-Jan-20	28-Jan-20	6-Feb-20
20-Jan-20	10-Feb-20	12-Feb-20	16-Jan-20	28-Jan-20	5-Feb-20
21-Jan-20	10-Feb-20	13-Feb-20	17-Jan-20	30-Jan-20	6-Feb-20
22-Jan-20	9-Feb-20	12-Feb-20	18-Jan-20	30-Jan-20	7-Feb-20
23-Jan-20	10-Feb-20	13-Feb-20	19-Jan-20	31-Jan-20	6-Feb-20
24-Jan-20	10-Feb-20	13-Feb-20	20-Jan-20	1-Feb-20	8-Feb-20
25-Jan-20	10-Feb-20	13-Feb-20	21-Jan-20	1-Feb-20	7-Feb-20
26-Jan-20	10-Feb-20	13-Feb-20	22-Jan-20	2-Feb-20	8-Feb-20
27-Jan-20	10-Feb-20	13-Feb-20	23-Jan-20	2-Feb-20	9-Feb-20
28-Jan-20	10-Feb-20	12-Feb-20	24-Jan-20	3-Feb-20	9-Feb-20
29-Jan-20	10-Feb-20	13-Feb-20	25-Jan-20	3-Feb-20	9-Feb-20
30-Jan-20	10-Feb-20	13-Feb-20	26-Jan-20	3-Feb-20	10-Feb-20
31-Jan-20	10-Feb-20	13-Feb-20	27-Jan-20	3-Feb-20	9-Feb-20
1-Feb-20	11-Feb-20	14-Feb-20	28-Jan-20	3-Feb-20	10-Feb-20
2-Feb-20	11-Feb-20	14-Feb-20	30-Jan-20	5-Feb-20	11-Feb-20
3-Feb-20	11-Feb-20	13-Feb-20	31-Jan-20	5-Feb-20	11-Feb-20
4-Feb-20	11-Feb-20	13-Feb-20	1-Feb-20	5-Feb-20	11-Feb-20
5-Feb-20	11-Feb-20	14-Feb-20	2-Feb-20	7-Feb-20	13-Feb-20
6-Feb-20	11-Feb-20	14-Feb-20	3-Feb-20	7-Feb-20	14-Feb-20
7-Feb-20	11-Feb-20	15-Feb-20	4-Feb-20	8-Feb-20	14-Feb-20
8-Feb-20	11-Feb-20	14-Feb-20	5-Feb-20	8-Feb-20	14-Feb-20
9-Feb-20	11-Feb-20	15-Feb-20	6-Feb-20	9-Feb-20	15-Feb-20
10-Feb-20	11-Feb-20	15-Feb-20	7-Feb-20	10-Feb-20	16-Feb-20
11-Feb-20	12-Feb-20	16-Feb-20	8-Feb-20	10-Feb-20	17-Feb-20
12-Feb-20	12-Feb-20	16-Feb-20	9-Feb-20	11-Feb-20	18-Feb-20
13-Feb-20	13-Feb-20	16-Feb-20	10-Feb-20	13-Feb-20	17-Feb-20
14-Feb-20	14-Feb-20	17-Feb-20	11-Feb-20	14-Feb-20	19-Feb-20
15-Feb-20	15-Feb-20	18-Feb-20	12-Feb-20	14-Feb-20	19-Feb-20
16-Feb-20	16-Feb-20	17-Feb-20	13-Feb-20	15-Feb-20	19-Feb-20
17-Feb-20	17-Feb-20	18-Feb-20	14-Feb-20	16-Feb-20	21-Feb-20
			15-Feb-20	17-Feb-20	21-Feb-20
			16-Feb-20	18-Feb-20	21-Feb-20
			17-Feb-20	18-Feb-20	22-Feb-20
			18-Feb-20	19-Feb-20	22-Feb-20
			19-Feb-20	19-Feb-20	22-Feb-20
			20-Feb-20	21-Feb-20	23-Feb-20
			21-Feb-20	22-Feb-20	24-Feb-20
			22-Feb-20	22-Feb-20	24-Feb-20
			23-Feb-20	23-Feb-20	25-Feb-20

⁵⁵ See tab 'NNDRS-backpopulating' of the [supplementary spreadsheet](#).

Fig. 15: Inversion of Fig S2 showing range of reporting dates per onset dates, for Diagnosed and Confirmed cases.

We can then clearly see that:

- 'diagnosed' cases were added from 9 Feb 2020, shortly after Version 5 of the case definitions published on 5 Feb⁵⁶, up to 18 Feb 2020, which is the day version 6 was published (ending the possibility to enter more 'diagnosed' cases).
- A large batch of retrospective 'diagnosed' cases was reported on 9 Feb 2020, four days after the creation of the 'diagnosed' category, which corresponds to the spike in reporting cases that was discussed in Box 2.⁵⁷
- At least 3 late December 2019 onset diagnosed cases were added after the data cutoff date (11 Feb) of the 'CDC Weekly' (which had 37 Dec 2019 diagnosed cases), as can be seen from the upper ranges for entering 'diagnosed' days.
- 'Confirmed' cases were entered from 31 Dec 2019 to 25 Feb 2020 (date of the gag order on the CDC), with, as should be expected, a mix of retrospective and non-delayed ones.
- At least 2 'confirmed' cases with 2019 onset dates were entered after the data cut-off date of the 'CDC Weekly' update (which had 104 Dec 2019 confirmed cases), as can be seen for onset dates 26 and 27 Dec 2019.
- All 'confirmed' cases entered for onset 1 Jan and 2 Jan were retrospective (at least 4 weeks for onset), including at least one missed in the 'CDC Weekly' update.

While all this is very informative as to the overall retrospective search and back-populating effort, it must be clearly stated that this does not tell us much at all about the reporting of retrospective 2019 onset cases. The problem is that Fig. S2. only gives us information on the reporting dates for onset cases in the last few days of 2019, so that we can only infer a few Dec 2019 onset additions, which is far away from the increment of 61 confirmed Dec 2019 onset cases from the 'CDC Weekly' update as of 11 Feb to the count on 27 Feb (61 = 165 - 104, see Table 6).

Equally unfortunately, the Fig. S3 of the First WHO report, which should in theory help us understand how the Dec 2019 onset cases evolved between the 5 Feb and 20 Feb with a midpoint on 12 Feb, is marred by inconsistencies such as showing no Dec 2019 onset cases at all for the 20 Feb extract, and some visible lower counts for early January 2020 onset from the 12 Feb to the 20 Feb extract, on top of likely issues with the extract dates themselves (see Fig. 16 below).

There is however some media mentions of the work of Chuanhua Yu and his team on possible 2019 cases, in an interview he gave to the Chinese medical journal, the Health Times, on 25 or 26 Feb 2020, at the exact time of the CDC gag order.⁵⁸ That article published on 27 Feb 20 reports that Chuanhua and his team, while working on checking records, found two possible November 2019 onset cases and 'five or six' cases with onset before 8 Dec 2019. Nevertheless these cases were still being checked and had not yet been promoted to either 'confirmed' or 'diagnosed'.

"The amount of data is too large. It is difficult to verify the data, distinguish between true and false, and sort and analyze, and the team is working hard."

Yu Chuanhua told reporters that many of these data are incomplete or incorrect, and some have no contact history. One needs to further contact the patients or families to verify.

⁵⁶ A revised Trial 5 version was published on the [8 February](#), with no material difference to the case definitions.

⁵⁷ Practically Chinese authorities waited until 12 Feb 2020 to announce a large one-day spike in reported numbers due to that new 'diagnosed' category. See WHO Situation Report of 13 Feb 2020, available at <https://bit.ly/3Sjw6Fg>.

⁵⁸ 'Experts judge the source of the new coronavirus: December 8 last year may not be the earliest time of onset in Wuhan', Health Times (健康时报), 27 Feb 2020. Original text and translation available at <https://bit.ly/303Ehwo>.

Just before publication on 27 Feb, Prof. Chuanhua called back the journalist to let him know that the two specific November 2019 cases he had just mentioned were incorrect inputs, and to add that the other cases with possible onset before 8 Dec he had mentioned earlier may just be further errors, and should be taken very carefully.

Before the deadline for publication [Note: 27 Feb 20], Professor Yu Chuanhua told the Health Times reporter that the data of two patients with new coronary pneumonia in November were entered incorrectly, and other data have yet to be verified one by one. When is the earliest time of onset? Yu Chuanhua said,

“We need to verify one by one and investigate the relevant contact history before we can analyze it.”

Whatever the status of these possible cases discussed by Prof. Chunhua in the interview, the nine November cases that were included in the SCMP leak of 13 Mar 2020 would have to be ‘confirmed’ or ‘diagnosed’ by the time the data of the leak was extracted, which most likely would be within a few days of the 27 Feb 2020, meaning early March, so as to be able to reach the SCMP and be processed and eventually published there on 13 Mar 2020.

Figure 3 presents epidemic curves of laboratory-confirmed cases, by symptom onset and separately by date of report, at 5, 12, and 20 February 2020. Figures 2 and 3 illustrate that the epidemic rapidly grew from 10-22 January, reported cases peaked and plateaued between 23 January and 27 January, and have been steadily declining since then, apart from the spike that was reported on 1 February (note: at a major hospital in Wuhan, fever clinic patients fell from a peak of 500/day in late January to average 50/day since mid-February).

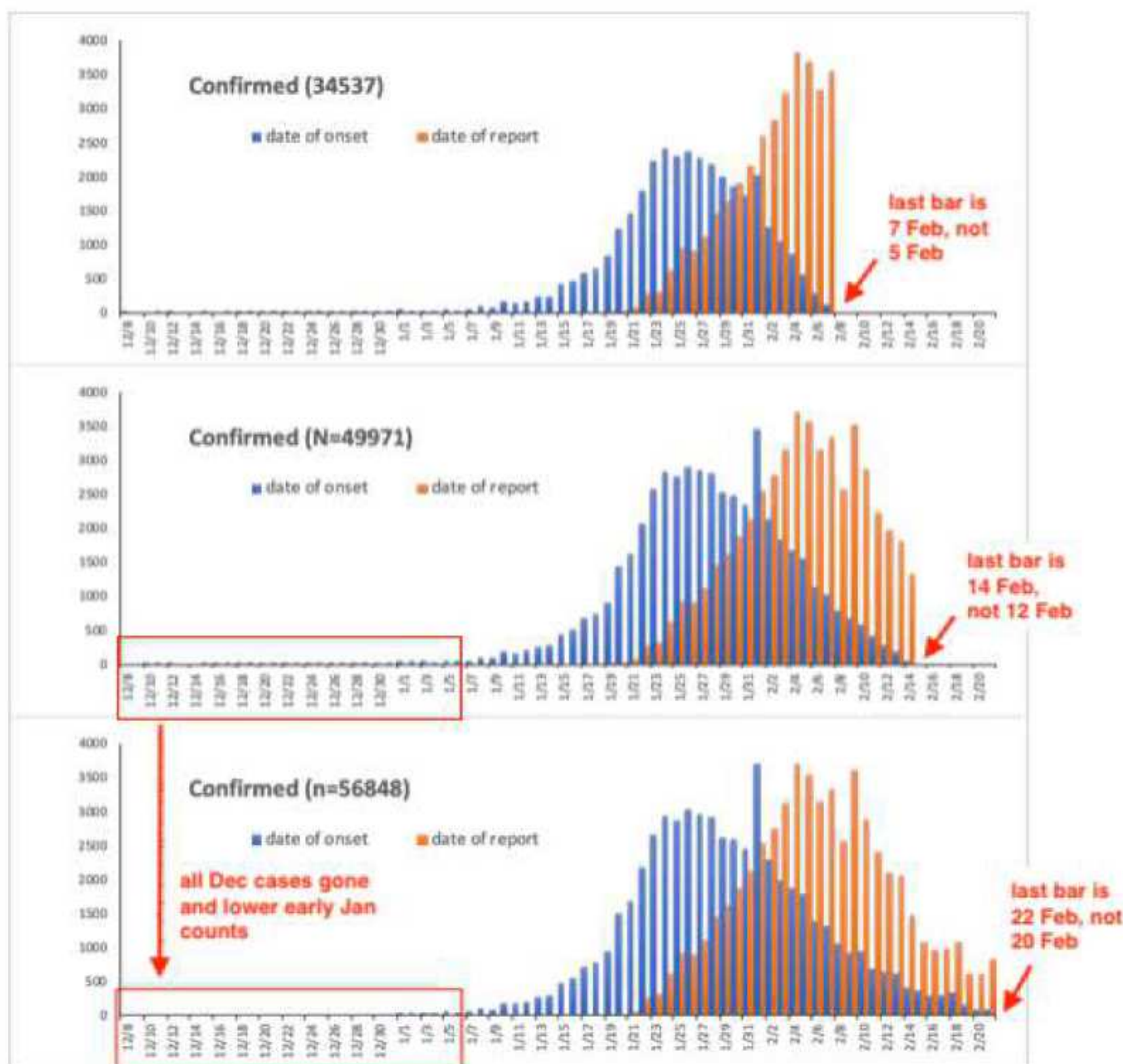


Figure 3. Epidemic curves by symptom onset and date of report as of 5 February (top panel), 12 February (middle panel) and 20 February 2020 (lower panel) for laboratory confirmed COVID-19 cases for all of China

Fig. 16: Figure 3 of first WHO report (published Feb 2020).

Finding 15:

Consequences for our understanding of the outbreak origins.

The official WHO count of 174 cases with onset in Dec 2019 (100 confirmed + 74 diagnosed) seems to be representative of what was in the official CDC NNDRS database as of 15 Feb 2020, half way through a detailed retrospective search effort coordinated by the CDC, which resulted in a quick progression of numbers in that database.

We can further show that by 27 Feb 2020 the NNDRS database already had between 247 and 260 cases with onset in December 2019 (with at least 165 confirmed).

Unfortunately that retrospective search seems to have been interrupted precisely around 27 Feb 2020, following a gag order imposed on the CDC on the 25 Feb. Soon after, a final count of 257 Dec 2019 cases, plus 9 November 2019 onset cases, leaked to the South China Morning Post of Hong Kong.

All this has some important consequences:

- With necessarily between 247 and 260 known cases for December 2019 (with at least 165 confirmed), all based on peer-reviewed papers and WHO reports, any analysis or conclusion based on the somehow retained 174 cases of the final WHO report (100 of which confirmed) is seriously compromised
- With 33 deaths of December 2019 cases (recorded by 24 Feb 2020), including 29 from confirmed ones, the situation in December 2019 was likely much worse than has ever been reported.
- With some very likely November 2019 cases, given the very good validation of the number of December 2019 cases in the SCMP leak of early March 2020, any conclusion that the outbreak started at the Huanan market in December 2019 is very likely wrong.
- The well-reported high hurdle facing the first correct diagnosis and reporting of COVID-19 cases⁵⁹, especially in the middle of a flu season, could easily have led to the first identification and disclosure of COVID-19 cases in the particularly sensitive setting of the Huanan wet market (older demographics of wet markets visitors combined with a semi-clandestine enclosed mah-jong room right in the market's south-west corner, all very close to a major transport hub).
- That detection may have also been facilitated by a few conscientious medical professionals in nearby top-tier research hospitals, some part of the CDC pneumonia surveillance network, and more generally thanks to the higher density of top-tier hospitals in Wuhan city centre. The original outbreak itself could have started somewhere else entirely in November 2019 or even earlier.⁶⁰
- Had the retrospective search been allowed to be pursued after 27 Feb 2020, the number of 2019 onset cases would likely have kept growing, revealing a much more detailed picture of the early outbreak. While very telling already, the count of 247 to 260 December 2019 cases as of 27 Feb 2020 (with at least 165 confirmed) is nevertheless a form of 'arrested development' (via a gag order on the CDC), which was eventually compounded with a case of 'reverse ageing' when the available data was effectively rolled back to around 15 Feb 2020 for the benefit of the final WHO report.

'Maybe someone is not at all interested in finding out what the truth is.'

Peter Embarek, head of the WHO expert team in Wuhan,
about the lack of contact-tracing of the earliest known cases by Feb 2021, [interview with TV2 Denmark](#)

⁵⁹ For some good analysis of the problem in Chinese media: '[People outside the statistics: Did they die of "ordinary pneumonia"](#)' (Caixin), '[Lost opportunity — the case of a small hospital that was neglected early in the Wuhan coronavirus epidemic](#)' (8 o'clock News), '[White and Green Manuals: Changes in New Coronavirus Pneumonia Diagnostic Standards](#)' (Freezing Point).

⁶⁰ This point about the distribution of tertiary hospitals in Wuhan and its possible impact on detection is notably made in Chuanhua et al. (<https://www.mdpi.com/1660-4601/17/19/7149/htm#B6-ijerph-17-07149>).