

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/351711216>

An analysis of the results of routine employee testing for SARS-like infections within the WIV and other Wuhan labs raises serious issues about their validity

Preprint · May 2021

DOI: 10.13140/RG.2.2.17411.02082/2

CITATION

1

READS

7,295

2 authors:



Steven Quay

Atossa Therapeutics, Inc.

133 PUBLICATIONS 4,205 CITATIONS

SEE PROFILE



Gilles Demaneuf

Engineer and Data Scientist

14 PUBLICATIONS 3 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Pandemics [View project](#)



DRASTIC [View project](#)

An analysis of the results of routine employee testing for SARS-like infections within the WIV and other Wuhan labs raises serious issues about their validity

Authors: Steven Quay¹ (MD, PhD), [ORCID](#); Gilles Demaneuf² (MS Eng., Data Scientist), [ORCID](#)

¹Email: Steven@DrQuay.com

²correspondence to: gilles@demaneuf.com

Abstract:

Various [representations](#) were offered during the China-WHO joint-team visit about the routine testing of employees at some Wuhan labs. The final [China-WHO joint-team](#) report used these representations as an argument to dismiss the possibility of the COVID-19 outbreak being caused by an early infection in a Wuhan institution involved in BatCov research.

A review of the test results mentioned in the China-WHO report and in interviews shows multiple issues both with the relevance of the tests, in terms of the exact institutions covered and the populations within these institutions, but also with the purported results of these routine tests which at best seem to point to very limited number of employees being tested.

Main Findings:

1. [The tests did not cover all relevant institutions](#)
Some key institutions that are attested to have worked on bat coronaviruses (BatCoV) at BSL2 or BSL3 are not mentioned at all in the statements.
2. [The tests performed in one of the institutions mentioned are irrelevant](#)
The tests done at the Jiangnan CDC cannot establish if some laboratory staff got infected in late 2019, as these tests were performed in June 2020 and were [PCR tests](#) which can only detect a positive infection within a few weeks of the first symptoms. In any case the Jiangnan CDC itself is most likely irrelevant anyway as not being known for any BatCoV research. The mention of the Jiangnan CDC in the Annex D5 of the China-WHO joint-team report can thus only create confusion.
3. [The tests did not cover all relevant people](#)
For a proper detection of possible research-related infections in 2019 that could have sparked the pandemic, the correct approach would have been to test (1) all field sampling staff, (2) all staff, students and temporary workers at institutions handling coronaviruses, (3) a sample of people living in proximity (especially downwind) from all these institutions.
4. [A proper testing at the WIV would have returned positive cases](#)
Given the prevalence of COVID-19 in Wuhan during the first half of 2020, with near absolute certainty some of the WIV 590 staff and students would have tested positive (igG+) in March 2020, following an infection with SARS-CoV-2.
Hence the statement that nobody at the WIV tested positive for SARS-CoV-2 is either (a) misleading as being based on an unreliably small fraction of the entire WIV population that should be tested, or (b) it is simply untrue.
The situation is made worse when considering the four labs [mentioned](#) by Peter Embarek.
5. **Clarification of the test populations and results is urgently required**
The WIV and other [Wuhan labs of interest](#) should either (a) clarify that they only did limited testings and detail exactly who was tested and when, or (b) provide case histories of the employees/students that necessarily did test positive if they tested large enough populations.

1. Background:

Since the early days of the COVID-19 pandemic, the possibility of the COVID-19 pandemic being the result of a lab-related accident has been regularly mentioned and has triggered various polemics. On their side, Chinese officials [tried to address these concerns](#) by pointing to the results of routine health checks of labs personnel, including antibodies testing for SARS-Like infections - which they claimed were all negatives.

In January 2021 the departing US Secretary of State nevertheless issued a factsheet, which while stating clearly that the the US Government had reached no specific conclusion, pointed to possible early infections involving lab personnel, without disclosing the source:

"We have not determined whether the outbreak began through contact with infected animals or was the result of an accident at a laboratory in Wuhan, China.

[--]

The United States government has reason to believe that several researchers inside the WIV became sick in autumn 2019, before the first identified case of the outbreak, with symptoms consistent with both COVID-19 and common seasonal illnesses. This raises questions about the credibility of WIV senior researcher Shi Zhengli's public claim that there was "zero infection" among the WIV's staff and students of SARS-CoV-2 or SARS-related viruses." (source: [fact-sheet](#))

In April 2021 the China-WHO joint-team published its final report, which considered as extremely unlikely the possibility of COVID-19 being the result of a lab-related accident. The negative routine testings in some Wuhan labs (including the WIV) were a key element in the joint-team assessment

"The three laboratories in Wuhan working with either CoVs diagnostics and/or CoVs isolation and vaccine development all had high quality biosafety level (BSL3 or 4) facilities that were well-managed, with a staff health monitoring programme with no reporting of COVID-19 compatible respiratory illness during the weeks/months prior to December 2019, and no serological evidence of infection in workers through SARS-CoV-2-specific serology-screening." (source: [China-WHO joint-team report](#), page 119).

Given the importance of these routine testings in the origins debate, a detailed and careful evaluation of the significance of these testings and of their purported results is required.

2. Available statements about tests:

We searched the China-WHO joint-team report and various interviews with the media for the available joint-team statements about the serological tests for SARS-CoV-2 and found four key statements listed below.

As can be seen these statements are not particularly specific on the exact population tested. In particular we could have reasonably expected the Annexes D5 and D7 of the China-WHO report to be much more precise on this subject, given its crucial importance:

S1: Shi Zhengli's [Reply to Science Magazine](#), 27th July 2020:

Q: Is it possible that someone associated with the institute became infected in some other way, for instance while collecting, sampling, or handling bats?

A: Such a possibility did not exist. Recently we tested the sera from all staff and students in the lab and nobody is infected by either bat SARS-CoV or SARS-CoV-2. To date, there is "zero infection" of all staff and students in our institute.

S2: Annex D7 of the [China-WHO report](#) on Covid-19 origins:

- The reserved sera in April 2019 and **March 2020** from all the workers and students in the research group led by Professor Shi Zhengli were seronegative for SARS-CoV-2 antibodies.

S3: Annex D5 of the [China-WHO report](#) on Covid-19 origins:

- Hubei [region] CDC. All its laboratory staff have been tested for SARS-CoV-2-specific antibodies: all had negative IgM and IgG results.
- Wuhan [city] CDC. one of its staff was confirmed SARS-CoV-2 seropositive after infection due to family cluster transmission. All other staff have tested negative. A health check is mandatory for all BSL-2 laboratory workers, but no serum is preserved.
- Jiangnan [district] CDC. All PCR tests for SARS-CoV-2 of all laboratory workers in June 2020 were negative.

S4: Peter Embarek quoted in [Science Magazine](#), 14th Feb 2021:

Q: But my question is whether you learned anything new in China. Now that you've been there, do you have more reason to say it's "extremely unlikely" than before?

A: Yes. We had long meetings with the staff of the Wuhan Institute of Virology and three other laboratories in Wuhan. They talked about these claims openly. We discussed: [--] 'Did you test your staff? [--]. They had retrospectively tested serum from their staff. They tested samples from early 2019 and from 2020.

3. Who should have been tested?

The population to test for the purpose of detecting a lab-related infection depends on the lab-related accident scenario considered. The table below shows the candidate populations for the 3 main scenarios: field sampling accident, Lab Acquired Infection (LAI) and Lab escape without LAI:

Lab-Related Accident Scenario:	Field sampling accident LS1	Lab acquired infection (LAI) of Wuhan lab personnel LS2	Lab escape without LAI LS3
Description	Infection during field sampling by or on behalf of a Wuhan laboratory	Infection inside a Wuhan institution with laboratories	Infection outside a Wuhan institution with laboratories
Index case	Personnel present at field sampling site, went back to Wuhan or infected someone who went back to Wuhan ¹	Can be lab personnel, staff, student, or anybody present in the institution (including temporary worker or visitor).	Someone out of the lab, in proximity typically, or in relation to lab activities (such as waste processing)
Actual Biosafety Level	Very limited to equivalent of Biosafety Level 2 (BSL-2) laboratory, full PPE commonly not worn ²	Bat SARS-related coronaviruses research officially performed at BSL-2 & BSL-3 levels	Bat SARS-related coronaviruses research officially performed at BSL-2 & BSL-3 levels
Incident	Either via contact with animal hosts or animal waste on site	<ul style="list-style-type: none"> infection in lab suite handling virus infection in common facilities shared with lab suite handling virus infection in institution precinct via aerosols, wastes or stray lab animal 	<ul style="list-style-type: none"> infection via aerosol outside lab precinct infection via incompletely neutralized liquid or solid lab wastes outside lab infection via stray lab animal
First person infected in a lab/institution?	Possibly - could also be employee, contractor, contract worker, collaborator, associate, or visitor, university student	Probably - could be employee, contractor, contract worker, collaborator, associate, or visitor	First person is infected outside of lab/institution
Entity of Interest	Any field sampling team	Any institution in Wuhan working on BatCoVs (may include Wuhan Institute of Biological Products (WIBP))	Any institution in Wuhan working on BatCoVs (may include WIBP)
Who should be tested?	Anybody present during field sampling	<ul style="list-style-type: none"> Direct Infection in lab suites: All people with access to the lab suites and lab wastes. Indirect Infection via contact: Anybody sharing space/facilities with people working directly in lab suites or handling the lab wastes inside the institution. Indirect infection via aerosols: Anybody within the lab institution 	<ul style="list-style-type: none"> Indirect Infection via contact: Anybody in proximity to any liquid wastes from the institution or handling the lab wastes outside of the institution. Indirect infection via aerosols: Anybody in downwind proximity of the institution.

Table 1: Lab-related accident scenario and corresponding candidate population for an index-case

¹ This distinct possibility was for instance [mentioned by Lin Fa Wang](#), a collaborator of Shi Zheng Li on many bat coronavirus research papers and a regular visitor at the WIV.

² It has been amply proven that PPEs are often not worn during field specimen collection trips. For instance Shi Zhengli herself explained that [most of the time only ordinary precautions are taken](#), and many photos and videos confirm it. This is in full contradiction with the assurances of the Annex D7 of the [WHO joint-team study report](#) where she is instead on record saying that 'all fieldwork is done with full PPE'.

As shown in the last row of the above table, the candidate population for an index case may go well beyond the people directly working in a laboratory suite.

Actually, in the case of a field sampling infection (scenario LS1 of the table), it is the field sampling personnel (strictly speaking any personnel at the sampling site) who need to be tested. Also because of the high risk of transmission of SARS-CoV-2 via aerosols, not just the laboratory personnel and students should be tested to detect a potential index case, but potentially anybody within the lab institution or downwind from it, and anybody potentially exposed to lab wastes, be it solid or liquid, inside or outside of the lab institution - as covered by the scenarios LS2 and LS3 of the table.

4. Which testing methods are relevant?

Active-infection (symptomatic) tests, such as PCR tests and IgM tests, are used to detect an infection within a few weeks from the initial symptoms, but will not reliably return a positive test beyond this initial period.

- PCR tests for SARS-like viruses (also called viral tests) determine the actual viral load present in samples. They will generally return a positive test up to 3 or 4 weeks after initial symptoms, and in some cases for longer. They are the most accurate test for active infection.
- IgM tests are antibody tests that can detect an infection 2 to 6 weeks (roughly) from initial onsets.

Thus PCR or IgM tests would not be of any great use to detect an infection prior to the official first outbreak date (8th Dec) if performed past mid January.

PCR tests and IgM tests would thus only be relevant to the detection of a possible infection before the official start of the outbreak (8th Dec 19) if they were performed before approximately mid January 20. In that case a positive test could relate to a possible lab-related accident prior to Dec 8th.

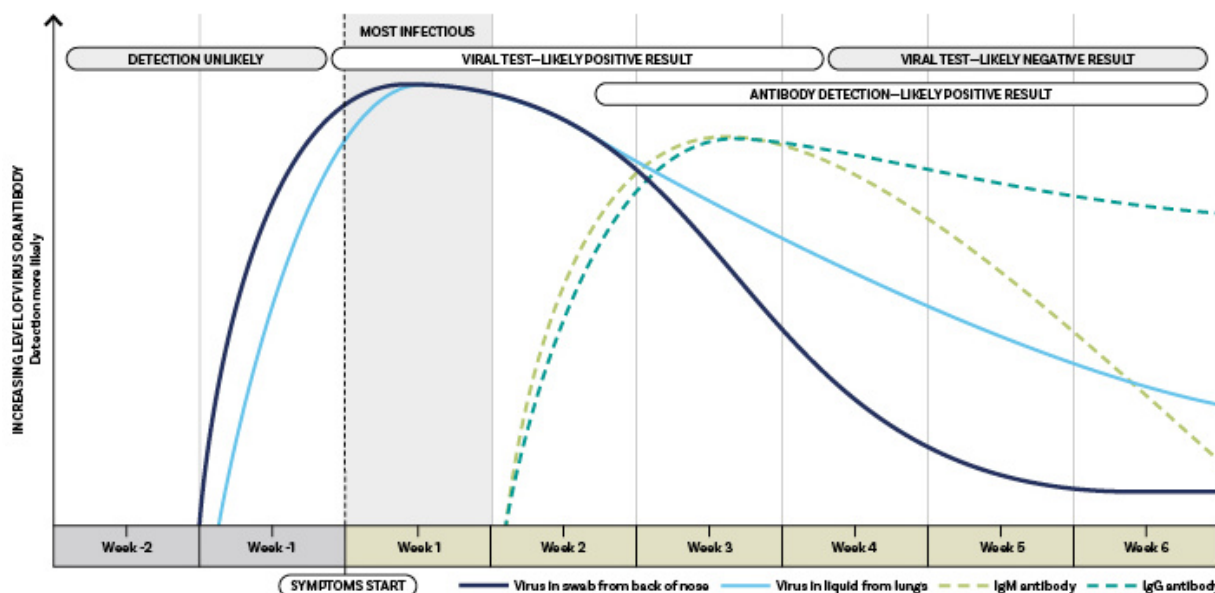


Fig 1: Levels of virus and antibodies after catching COVID-19 and the likelihood they will be detected during testing [source: [health.govt.nz](https://www.health.govt.nz)]

Post-infection (recovery) tests, such as IgG antibody tests, allow one to detect cases amongst laboratory staff or students weeks or months after a possible infection. If these tests are done after the official start of the outbreak (Dec 19) there is a strong possibility that such infection may not be due to a lab-related accident but to a community-acquired infection during the later outbreak.

Standard IgG tests targeting SARS-CoV-like viruses are broad enough to be able to detect SARS-CoV-2, even back in 2019. Some studies show that standard IgG tests are still relatively reliable many months after infection, returning reliable positives 6 to 12 months after infection (good p).

In conclusion, in order to detect infections in a Wuhan lab personnel or students, or in some field sampling team, before the official breakout in Dec 2019, the tests of interests could be:

- SARS-CoV-Like IgM antibodies tests performed in 2019 or the first 2 weeks of 2020 (approximately).
- Post-infection SARS-CoV-Like IgG antibodies tests performed in late 2019 or anytime in 2020, followed by necessary tracing to understand whether they were pre-breakout or post-breakout infections.

Box 1: How long does a positive IgG test persist?

A few studies demonstrate conclusively a sustained IgG positive outcome for over six months after the start of the recovery phase.

- A [paper in Lancet](#) (Seroprevalence and humoral immune durability of anti-SARS-CoV-2 antibodies in Wuhan, China: a longitudinal, population-level, cross-sectional study), based on patients in Wuhan, concluded that:

Although titres of IgG decreased over time, the proportion of individuals who had IgG antibodies did not decrease substantially (from 30 [100%] of 30 at baseline [April 14-15, 2020] to 26 [89.7%] of 29 at second follow-up [Oct9-Dec5] among confirmed cases, 65 [100%] of 65 at baseline to 58 [92.1%] of 63 at second follow-up among symptomatic individuals, and 437 [100%] of 437 at baseline to 329 [90.9%] of 362 at second follow-up among asymptomatic individuals).

In other words 90% of individuals had sustained IgG antibodies 6 months at least after the first test in April 20.

- A [paper in Science](#) ('Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection'), based on COVI-19 patients in the US, reached very similar conclusions that:

'The percentage of subjects sero-positive for RBD IgG at 6 to 8 months PSO was 88% (35 out of 40) [--]The percentage of subjects seropositive for spike IgG at 6 to 8 months PSO (≥ 178 days) was 90% (36 out of 40).

The [actual length of immune answer detectability](#) was also found to depend on the individual.

- A [British study](#) found very similar results, with sustained Spike, RBD and nucleocapsid (N) IgG positive rates for months.

5. Which institutions should have been included in these tests?

BatCov studies were done routinely in P2 and P3 laboratories, not in the much more constraining P4 environment. For instance the National Virus Resource Center (NVRC, 国家病毒资源库), which is affiliated to the WIV, used to list the relevant biosafety level for different pathogens before it was taken down (with all the WIV databases it used to give access to), as per [archived page](#). The list includes under the heading 'Animal Viruses':

- 蝙蝠冠状病毒 BSL-2 [Bat Coronavirus BSL-2]
- 大鼠冠状病毒 BSL-2 [Rodent coronavirus BSL-2]

Practically this means that the China-WHO report is rather misleading as to the relevant BSL levels in its evaluation of the possibility of a lab-related accident, never mentioning BSL-2 where some of the BatCoV research took place, but mentioning instead BSL-4 where no such research was known to take place:

“The three laboratories in Wuhan working with either CoVs diagnostics and/or CoVs isolation and vaccine development all had high quality biosafety level (BSL3 or 4) [sic] facilities that were well-managed,
(‘Argument Against’ p 119)”

As per the following [study](#), these institutions are of particular interest for their demonstrated involvements with BatCovs research at P2 or P3 level:

- WIV at its two sites: Xiaohongshan Park and Zhengdian.
- Wuhan [city] CDC
- Hubei [provincial] animal CDC
- Huazhong Agricultural University (HZAU)
- Wuhan University
- Wuhan Institute of Biological Products³



Fig 2: Biosafety laboratories in Wuhan [source: [rdemaistre.medium.com](#)]

³ The WIBP is part of an integrated vaccine development platform which includes the WIV at Zhengdian. For a detailed review of the WIBP activities, past and present, and the reasons for its inclusion see [this study](#).

6. Which laboratories are mentioned in the China-WHO report?

6.a Laboratories mentioned in report:

The annexes of the China-WHO joint team report mentions the following laboratories. In blue are the ones which were asked about their testing of staff and students.

Annex D5	Annex D6	Annex D7
<ul style="list-style-type: none"> Hubei [provincial] CDC Wuhan [city] CDC Jiangnan [district] CDC 	<ul style="list-style-type: none"> Hubei [provincial] Animal CDC⁴ 	<ul style="list-style-type: none"> Wuhan Institute of Virology



Fig 3: Joint-team itinerary [source: WHO]

⁴ As per the Annex D6 of the [WHO joint-team study report](#) the Hubei Animal CDC has a BSL-3 laboratory which is not used due to the lack of staff (on top of its BSL-2 facility). Such lack of operational utility (including maintenance) due to budget constraints is unfortunately rather common and has been a constant subject of concern amongst Chinese experts. For instance Yuan Zhiming, director of the WIV P4, raised this precise issue as [recently as October 2019](#).

6.b Comparison to target list:

Let's now compare that list to the list of relevant labs and institutions as per 3. Above. In bold are the common institutions.

Institutions of interest for a tests review	Institutions for which tests are mentioned in the China-WHO joint-team report
<ul style="list-style-type: none">• Wuhan Institute of Virology (two sites: Xiaohongshan Park and Zhengdian)• Wuhan CDC• Hubei [provincial] Animal CDC• Huazhong Agricultural University (HZAU)• Wuhan University• Wuhan Institute of Biological Products	<ul style="list-style-type: none">• Wuhan Institute of Virology• Hubei [provincial] Animal CDC• Wuhan [city] CDC• Jiangnan [district] CDC

We can see that 3 of the 6 institutions of interest were asked about their tests.

6.c Misleading mention of Jiangnan CDC tests:

We further note that the mention of the testing at the Jiangnan [district] CDC in the WHO report can be rather misleading as it makes strictly no sense. It is mentioned by Peter Embarek as an example of testing that should dispel any idea that infected lab workers may have been responsible for the outbreak.

However the tests conducted at the Jiangnan CDC were viral [PCR] tests performed in June 2020 - hence active infection tests which are of no relevance at all for detecting infections as far back as 2019.

7. Who was actually tested?

7.a Limitations of the purported testing

The [statements](#) (see p. 1) made by the WIV or the joint-study about the testing procedure at various Wuhan labs are rather vague and somewhat contradictory. Here is how these three statements describe the populations tested:

- S1: staff and students in the [WIV] lab
- S2: staff from the WIV and three other institutions
- S3: workers and students from the laboratory of Zhi Zhengli at the WIV

In Statement S2, Peter Embarek explains that joint-study team asked the same question and got the same answer from the [four labs they visited while in Wuhan](#):

Putting all this back in our Lab-scenario table, we get:

Lab-Related Accident Scenario:	Field sampling accident LS1	Lab acquired infection (LAI) of Wuhan lab personnel LS2	Lab escape without LAI LS3
Entity of Interest	Any field sampling team	Any institution in Wuhan working on BatCoVs (may include WIBP) All BSL levels	Any institution in Wuhan working on BatCoVs (may include WIBP) All BSL levels
Entity tested	No mention	<ul style="list-style-type: none"> • WIV • Hubei [provincial] CDC • Wuhan [city] CDC • Jiangnan [district] CDC [<u>but active-infection tests irrelevant</u>] 	No mention
Who should be tested?	Anybody present during field sampling	<ul style="list-style-type: none"> • Direct Infection in lab suites: All people with access to the lab suites and lab wastes • Indirect Infection via contact: anybody sharing space/facilities with people working directly in lab suites or handling the lab wastes inside the institution. • Indirect infection via aerosols: Anybody within the lab institution 	<ul style="list-style-type: none"> • Indirect Infection via contact: anybody in proximity to any liquid wastes from the institution or handling the lab wastes outside of the institution. • Indirect infection via aerosols: Anybody in downwind proximity of the institution.
Population tested	No mention	No number given, just some staff/workers and students. Exact ab suites not clear (beyond SZL lab)	No mention

Table 2: expected vs. actual entities and populations tested

In particular we note that:

- There is no mention of any testing done that would cover a field sampling accident (LS1). And as per note (1) full PPEs were commonly not used, only ordinary protection (latex gloves, light mask, and plastic poncho inside a cave)
- There is no mention of any testing that would address the possibility of a Lab Escape with LAI (LS3).
- At least two more institutions with a history of work on BatCoVs should have been considered: the Wuhan University (A/BSL3) and the Wuhan Institute of Biological Products (WIBP).

7.b Target populations for the institutions mentioned

Let's try to evaluate how many people should have been tested in any case in the labs mentioned in the statements under the Lab acquired infection (LAI) scenario.

WIV Population

According to the [official website](#) of the Institute of Virology in August 2020, the Institute normally hosts 325 graduate students and 268 employees. That gives us a total of around 593 people. As is known, the Chinese government shutdown all “non-essential businesses” in Wuhan from February 13, 2020 to April 8, 2020 (see [Wuhan Lockdown Timeline](#)). It has not been disclosed how much of the WIV was closed during that time, but clearly those WIV staff who were not working during the lockdown but instead were staying home should have shown the COVID-19 prevalence of the general Wuhan population.

It is worth noting that the [WIV has two sites](#): the old site (Xiahongshan in Wuchang district) and the new site (Zhengdian in Jiangxia district) where the P4 is (with some P2s and P3s). However it is not clear whether Shi Zhengli [was referring](#) to both sites when discussing sampling staff and students in statement S1, or to the Zhengdian site only.

It is not clear also whether Shi Zhengli referred to all the labs that may have handled some BatCoV (be they P2, P3 or even P4) or to her lab only. Given that BatCoVs were normally handled at P2 and P3, and to avoid any suspicion on the P4, at the very minimum all the P2 and P3 laboratory handling BatCoVs plus the P4 should be tested - this would cover a *direct laboratory exposure*.

But a direct lab exposure is not the only possible exposure. As SARS-CoV-2 is highly transmissible via aerosols, and since it may also be transmitted via common surfaces and liquid wastes, a primary case could happen out of one of the laboratory suites handling BatCoVs. This is something which happened for instance during the Beijing Lab Leaks on 2004, when an ordinary electron microscope room (no biosafety level) got infected by a only partially inactivated SARS virus belonging to a P3 lab suite, with the additional twist that the virus was actually held in a fridge in the corridor (no biosafety level again) due to overcrowding of the lab suite.

To be able to understand how comprehensive the testing should be in the WIV, we would therefore have to know how the laboratory suites are configured within the Xiahongshan or Zhengdian site, whether these suites share common facilities with non-lab activities on each site, what the movement of people is within each site and between the two sites, etc.

At worse it may thus mean that every single staff or student within both sites should logically have been tested, whether working in a lab suite or not, and irrespective of whether they were at the Xiahongshan or Zhengdian site - since both sites were likely working on coronaviruses in the BSL-2 or BSL-3 labs. That would mean 590+ staff and students.

Strictly speaking we may have to also consider anybody with a prolonged physical presence in these institution precincts - given the potential risk of infection via aerosols. That would include any temporary construction workers on site, such as the ones using the temporary barracks that were only within 15 meters of the P2/P3 labs in the WIV compound (see [page 30 of this report](#)).

Four Wuhan labs (as per S4)

We do not have enough information to venture how large a candidate population for testing the four labs mentioned in statement S4 may represent. If we suppose very roughly maybe 50 people (researchers, staff and students) potentially exposed to a direct laboratory infection in each one of these 4 institutions, and another 50 in each institution potentially exposed to indirect infection via aerosols and shared facilities or wastes, we get already 400 persons in total that should be tested as priority over these 4 institutions.

Note that the exact Wuhan institutions and laboratories involved in coronavirus research - by opposition to the ones mentioned in statement S4 - were listed in [section 4](#).

8. Observed seropositivity rates in WIV population of interest in April 2020

In order to estimate how many employees and students at the WIV should have returned a positive antibodies test in March 2020 (as per the Science Magazine statements), we shall first make use of this [paper](#) (Oct 20) which establishes background IgG antibody seropositivity levels in April-May 2020 in Wuhan residents through a cross-sectional study of around 35,000 individuals.

Note that that paper only considered individuals without a COVID-19 history - leading actually to a slight underestimation of the seropositivity rate across the full population, which additionally includes people with a COVID-19 history - hence people with a likely much higher IgG positive rate.

In a very useful manner, the paper gives seropositivity prevalence levels for IgG testing by urban/non-urban districts, age and sex. It further gives a confidence interval (C.I.) for the prevalence level in the urban areas.

We thus obtain a 4.4% IgG+ prevalence level for urban Wuhan in April 2020, with a 95% C.I. of [4.0%, 4.8%].

'The seropositive prevalence in the urban districts was higher than that in the suburban and rural areas (4.4% [95% C.I.: 4.0%-4.8%] vs 2.9% [95% C.I.: 2.3%-3.6%]; $P < .001$), demonstrating an urban to suburban gradient'

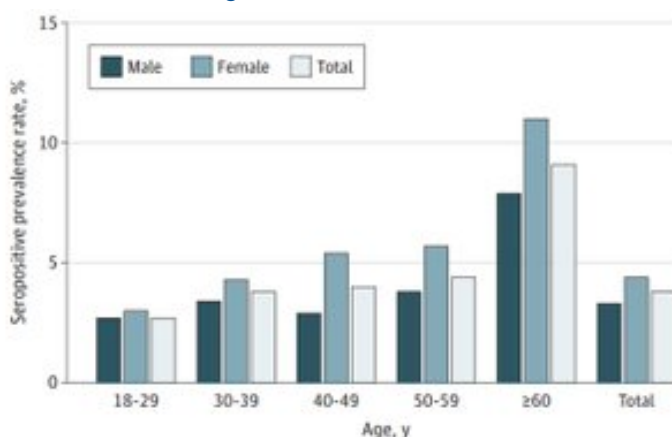


Fig 4.a: Severe Acute Respiratory Syndrome Coronavirus 2 Seropositive Prevalence by Sex and Age Group [source: [jamanetwork.com](#)]

Table. SARS-CoV-2 Seropositive Prevalence in Different Age Groups

Age group, y	Patients, No.	Male, %	SARS-CoV-2 seropositive prevalence, patients, No. (%) [95% CI]
18-29	8163	44.5	232 (2.8) [2.5-3.2]
30-39	13 471	44.3	524 (3.9) [3.6-4.2]
40-49	7713	52.3	313 (4.1) [3.6-4.5]
50-59	4932	64.7	221(4.5) [3.9-5.1]
≥60	761	49.3	70 (9.2) [7.1-11.3]

Fig 4.b: SARS-CoV-2 Seropositive Prevalence in Different Age Groups [source: [jamanetwork.com](#)]

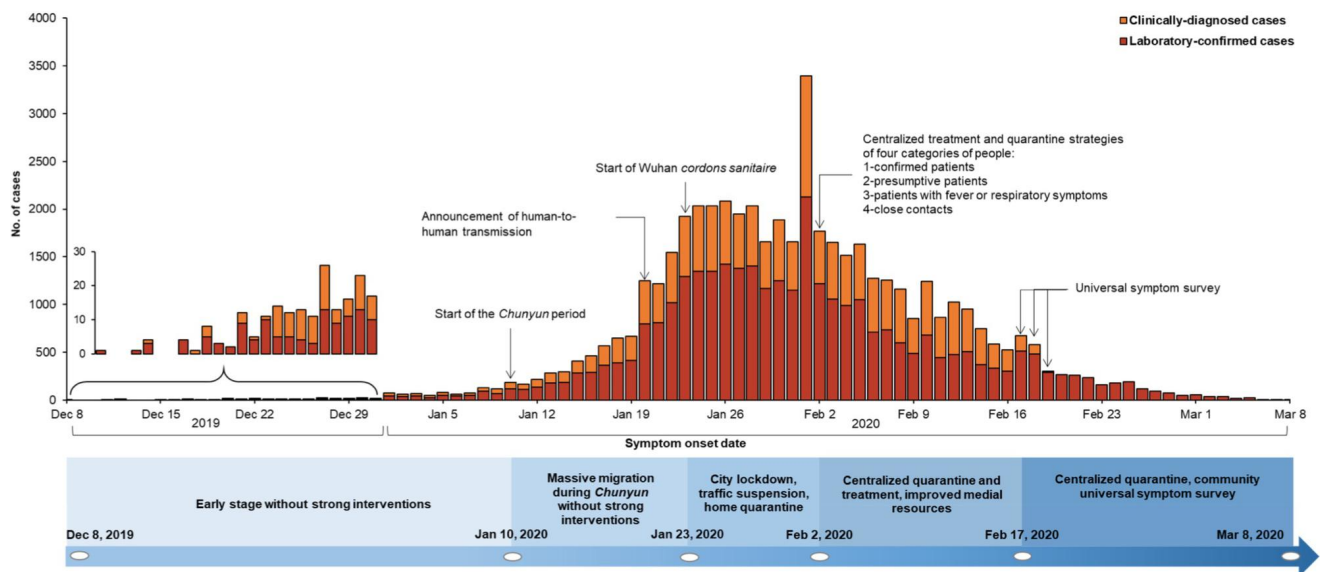
As a first pass, as most representative of the WIV population of interest we shall use the urban district IgG positive (IgG+) rate of 4.4% in April 2020.

9. Expected seropositivity rates in WIV population of interest in March 2020

We now need to assess what the same prevalent IgG+ rate for Wuhan urban would have been just one month earlier, in March 2020.

To do so we observe that there were very few COVID-19 cases with onset date past the 20th Feb 2020 and that IgG+ rates start flattening out at week 3 past onset. Hence the people who would have tested IgG+ in April cases would have in their large majority (90%+) also tested positive in March - with only some of the cases with onset in the last 3 weeks of February potentially not testing IgG+ in early March.

eFigure 1. Daily numbers of laboratory-confirmed and clinically-diagnosed COVID-19 cases by symptom onset date across the five periods



A total of 17 365 clinically-diagnosed cases and 32 583 laboratory-confirmed cases were included in the figure.

Fig 5: Daily numbers of laboratory-confirmed and clinically-diagnosed cases
[source: [joi200040supp1_prod.pdf](#)]

Conversely, given that various studies (see box 1) show that the vast majority of IgG positive individuals (90%+) would again test positive up to 9 months (if not more) after infection, then the vast majority of the people who would test positive in March would again test positive in April.

Hence the individuals who would test positive in April would also test positive in March (except for a small minority of late February onset cases), and vice-versa. We shall thus conservatively round down the 4.4% April IgG+ prevalence rate to a 4% March IgG+ prevalence rate.

We shall later consider a few alternatives in a sensitivity analysis.

10. How many of the WIV employees and students should have tested positive in March 2020?

Based on the above background seropositivity level we shall calculate the expected number of returned seropositivity tests at the WIV if they had been done in the period covered by the study: March 20.

To do so we will use a binomial distribution. That probability model assumes observations (draws) are independent of each other with the occurrence of a positive result having the same probability, p . The important point here is 'independent'. If there was an infection cluster in a lab, then that assumption would not be appropriate.

However, that is perfectly fine. We shall suppose no infection clusters in the lab and calculate how many employees/students should be expected to be seropositive for IgG just by virtue of the Wuhan background seropositivity level applicable to that population of interest, assuming a similar level of community-acquired infection for WIV staff. Any cluster at the WIV would in any case increase that rate, so at worst we will underestimate the true rate this way,

The probability of having x seropositive results in n individuals, supposing independence of the draws (with p the probability of a positive result and $q = 1-p$ the probability of a negative result), is given by the binomial distribution:

$$P_x = \binom{n}{x} p^x q^{n-x}$$

and the mean is simply np .

10.a Base case:

Using 593 employees and students (n) at the WIV (across both sites) and 4% Wuhan community prevalence rate in April-May 2020 (p) will find that:

- We should expect around **24 positive tests** at the WIV (0.04×593),
- The probability of returning no IgG positive test at all ($x=0$) at the WIV in March 2020 is ~ 0.000000000003 (11 zeros after the comma), basically null.

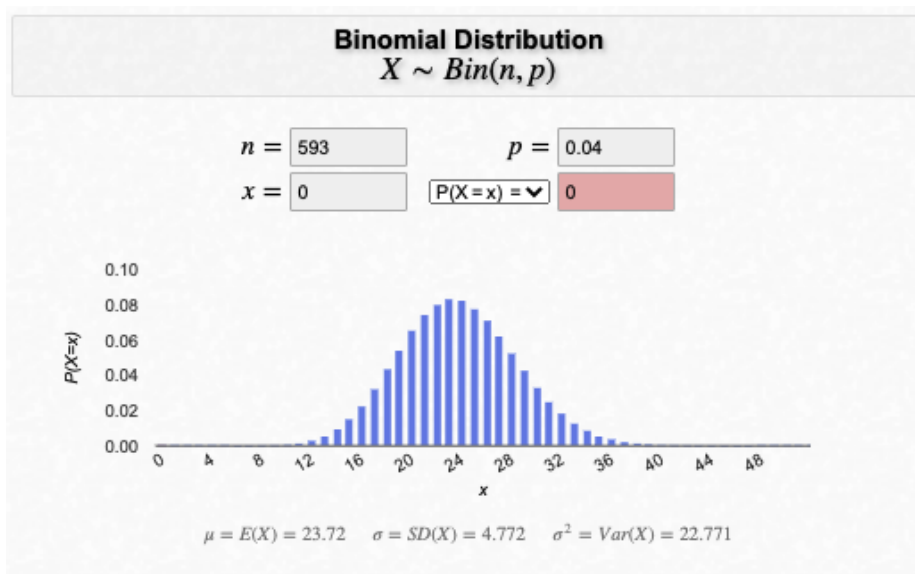


Fig 6.a: Binomial Distribution ($x=0$, $n=593$, $p=4\%$) [\[source\]](#)

Even assuming testing of only 51 employees and students, effectively a very small subset of the population that could include the index case for a lab-related accident, the probability of returning no IgG positive test at all is only 12.5%, and is halved with every additional 17 individuals.

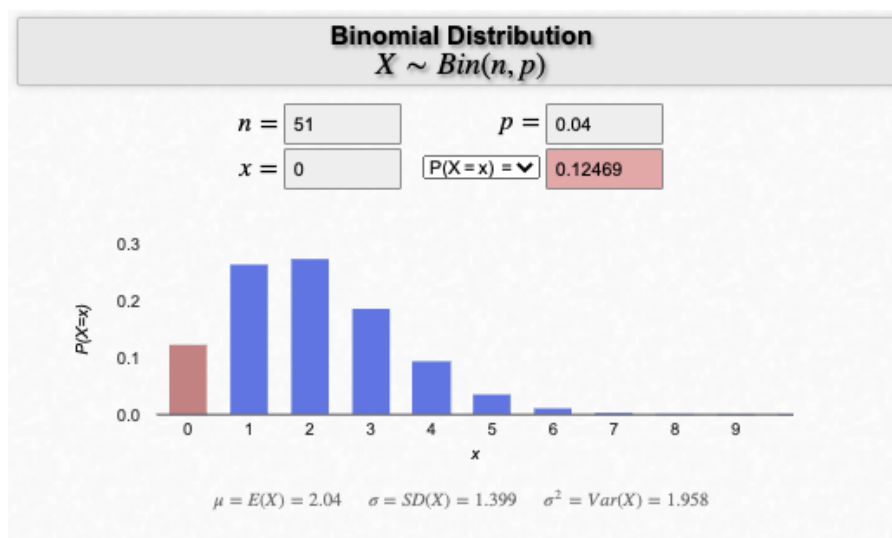


Fig 6.b: Binomial Distribution ($x=0$, $n=51$, $p=4\%$) [source]

10.b Sensitivity analysis:

To check how robust this result is, we shall consider a few variations based on the applicable background seropositivity rates. The 4.0% we noted above should be quite stable, being calculated on a fairly large population - but let's assume that it may be as low as 1.00% for illustration [an extremely low and very unlikely value, well beyond the 95% C.I. of roughly 3.6%-4.4%].

Even with 1% background seropositivity rate the conclusion remains essentially the same: the probability of returning 0 positive tests is negligible at 0.26%.

probability of 0 positive test		average number of positive tests	
bgr inf. \ WIV pop	593	bgr inf. \ WIV pop	593
1.00%	0.0026	1.00%	6
2.00%	6.27E-06	2.00%	12
3.00%	1.43E-08	3.00%	18
4.00%	3.07E-11	4.00%	24
4.40%	2.58E-12	4.40%	26

Conclusion:

Had the WIV population been tested in May 2020 it would have without any doubt returned positive cases. On average 24 with an assumed positivity rate of 4%.

Box 2: Mathematical corner

Here is how you can easily tabulate the probabilities in your head.

The key is to remember that with a 4.0% seropositivity rate there is very close to a 50% chance of no seropositive in a group of 17 people. This is a simple binomial distribution result. See for instance this online calculator.

If you add another 17 people, the probability of no seropositive in the group is again divided by 2. From there you can see that the probability of no seropositive at in a group of $(n * 17)$ people is $\sim 1/2^n$.

510 is still short of the 590 staff and students at the WIV, but conveniently $510 = 30 \times 17 = 3 \times 10 \times 17$. Also $1/2^{10} \sim 1/1,000$, so the probability of no IgG seropositive in a group of **510** people is around $1/1,000^3 = 1$ in a billion!

If one prefers to work in powers of 1/100 then the probability is about 0.01 for a group of 113 people. So for **565 people** (5×113) it becomes $1/10^{10}$, **a 10th of a billionth**.

11. How many of the WIV staff & students did get infected by March 2020?

Let's now move from seropositivity to infection by evaluating the expected infection rates in the WIV population of interest in March 2020:

If we ignore false positive and false negatives, the seropositivity prevalence levels above tell us that at least 4% of the WIV population of interest should have been infected with SARS-CoV-2 by May 2020.

The reality is that there should be some false positives and false negatives. False negatives cause an underestimation of the actual prevalence level, while false positives cause an overestimation. To be conservative we shall only consider false positives.

The relevant confusion matrix of the IgG test is thankfully given in the [supplement](#) of the study:

“The sensitivity, specificity, positive predictive value, and negative predictive value of IgG were 87.2 %, 99.3%, 98.8% and 91.7%, respectively.”

Let's include the false positives in the analysis, while ignoring the false negatives - a very conservative assumption. Applying the Positive Predicted Value (98.8%) to 4% of seropositives (and), we get a rate of True Positive of **3.95%**.

Stat	Value	
PPV	98.80%	$TP/(TP+FP)$
NPV	91.70%	$TN/(TN+FN)$
Sensitivity	87.20%	$TP/PC = TPR$
Specificity	99.30%	$TN/NC = TNR$

We could go further and include the false negatives to get the exact True Positive rate. This is done in the Supplement to this paper. However because the resulting rate is rather high (possibly due to issues with the confusion matrix of the IgG test), we prefer to conservatively use the 3.95% above.

From there we can deduce a conservative estimate of the number of WIV staff & students likely infected by March 2020.

Based on the conservative estimate of 3.95% for the True Positive Rate, the expected number of staff and students likely infected with COVID-19 by March 2020 is essentially the same as the expected number of positive tests (24).

12. Some possible interpretation

To be very clear, the fact that some personnel and students should have tested positive to BatCoV antibodies, at the WIV and other key labs around March 2020, does not necessarily mean that the pandemic began with a laboratory-acquired infection at some time in 2019. But such an ineluctable statistical conclusion of necessary infections, if it had been confirmed by the WIV and other labs, would have led to proper tracking and review of these cases - the absolute minimum requirement when investigating an outbreak.

What is surprising is that for some reason Chinese officials seem to have preferred avoiding any such standard line of epidemiological investigation by making statements that do not bear scrutiny today. Here we will venture a possible interpretation based on the available information at the time.

When Dr. Shi was asked in July 2020 ([statement S1](#)) about whether any staff at the WIV tested positive for COVID, the [data available](#) at the time was essentially that only around 50,000 COVID-19 cases had been recorded in Wuhan, with yet no broader seropositivity study being published.

Precisely 49,912 official COVID-19 cases over a Wuhan population of 11 mln (as of end March 2020, and little changed after that) could be interpreted as a very low prevalence rate of around 0.454%. Only proper IgG testing across large urban cohorts [would later convincingly show](#) that the real prevalence rate within Wuhan was actually much higher; 10 times higher at about 4.4%, a rate which is still low but much closer to the rates recorded in many cities worldwide (for instance [Madrid](#): 11.3%, [Boston](#): 9.9%, [Stockholm](#): 7.35%, all in May 2020).

Applying the binomial distribution to this 'prevalence' rate of 0.45%, we can see that the likelihood of no test returning positive over 51 staff members is around 79%, which is a much more credible rate than the 12.5% likelihood arrived at using the actual background rate of 4% .

It is therefore conceivable that Dr. Shi was only aware of this low rate of around 0.45% when she made that statement in July 2020. A combination of a (i) low number of staff/students tested (focussing - wrongly - strictly on her team) and (ii) such an underestimated prevalence rate of 0.45% would indeed give a credible probability (of 79%) of not returning any positive IgG test.

For the sake of clarity, this does not mean that Dr Shi was purposely lying. The exact information about the results of the various tests would have been carefully controlled, and it is very possible that Dr Shi was simply repeating what she was told and believed at the time, even if a simple analysis - as done here - should now convince her of the improbability of these results.

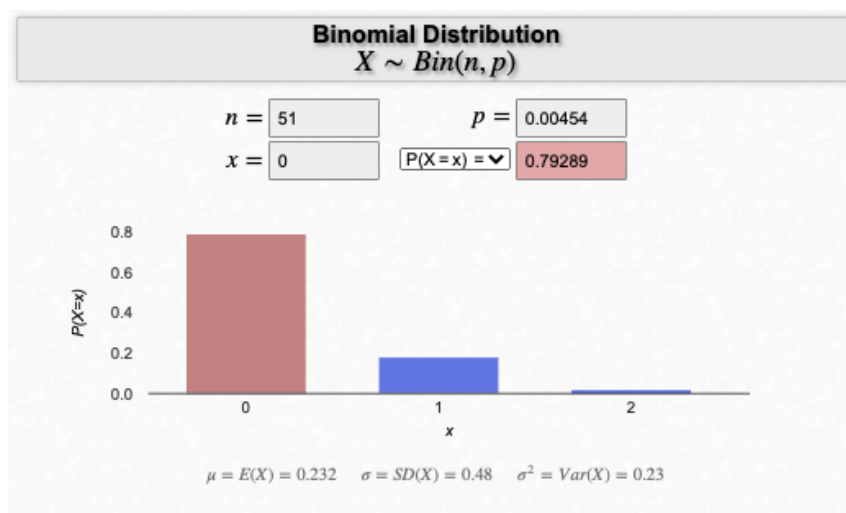


Fig 11.a: Binomial Distribution ($x=0$, $n=51$, $p=0.454\%$) [[source](#)]

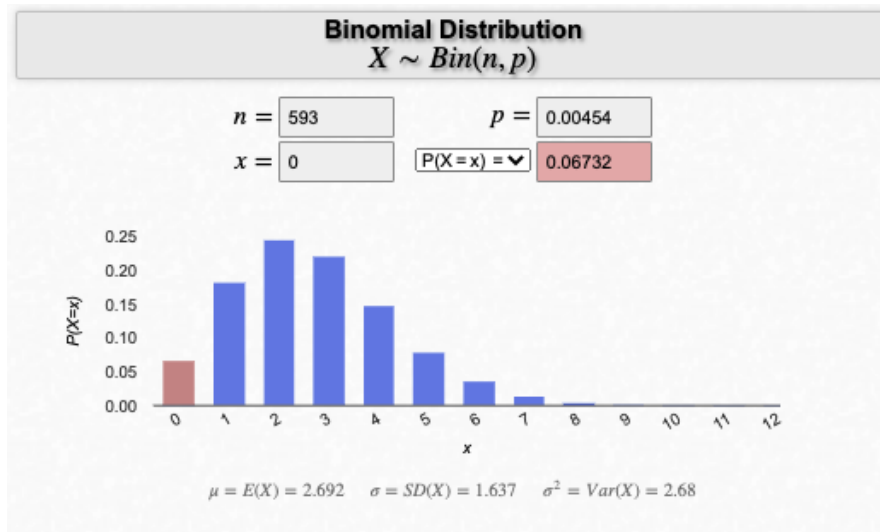


Fig 11.b: Binomial Distribution ($x=0$, $n=593$, $p=0.454\%$) [\[source\]](#)