Origin of Covid-19: Lab Leak

A case by Rootclaim, presented by Saar Wilf (Founder)
Rootclaim’s Mission

Rootclaim attempts to solve the problem of intelligent people reaching contradictory conclusions while having access to the same evidence.

Former CDC chief Redfield says he thinks COVID-19 originated in a Chinese lab

CDC chief says lab origin of Covid is ‘one possibility,’ but animal host is most common for coronaviruses

The Huanan market was the epicenter of SARS-CoV-2 emergence with vendors selling live animals. Together, these analyses provide dispositive evidence for the emergence of SARS-CoV-2 via the live wildlife trade and identify the Huanan market as the unambiguous epicenter of the COVID-19 pandemic.

Endonuclease fingerprint indicates a synthetic origin of SARS-CoV-2 arisen by random evolution. Both the restriction site fingerprint and the pattern of mutations generating them are extremely unlikely in wild coronaviruses and nearly universal in synthetic viruses. Our findings strongly suggest a synthetic origin of SARS-CoV-2.

One IC element assesses with moderate confidence that the first human infection with SARS-CoV-2 most likely was the result of a laboratory-associated incident, probably involving experimentation, animal handling, or sampling by the Wuhan Institute of...

Four IC elements and the National Intelligence Council assess with low confidence that the initial SARS-CoV-2 infection was most likely caused by natural exposure to an animal infected with it or a close progenitor virus—a virus that probably would be...
What Rootclaim Doesn’t Do

Fact-Checking
Investigative Journalism
Leaks/Uncovering new evidence
Conjecture when evidence is lacking

Rootclaim excels when experts struggle to reach a consensus despite abundant evidence
There are a few common flaws people experience when dealing with high uncertainty and multiple pieces of evidence.

- Reliance on gut feelings and conventional wisdom to assess likelihoods
- Weak intuition for compound probabilities
- Ignoring filters on evidence
- Ignoring priors - Prosecutor's fallacy
- Overlooking dependencies
The goal of the Rootclaim methodology is to accurately quantify the strength of the evidence so it can be fed into a proven probabilistic model.

1. Gather all likely hypotheses (prior probabilities) of each hypothesis
2. Assess the initial likelihoods of each hypothesis
3. Gather all the available evidence, without cherry picking and accounting for filters
4. Sort information into evidence groups to account for dependencies
5. Quantify how each group of evidence affects the likelihood of each hypothesis
6. Assign a best explanation for each hypothesis “steel man”
7. Assess the reliability of our own estimates to prevent overconfident results
8. Once all elements have been quantified, proven mathematical formulas calculate the final likelihoods of each hypothesis

Human inference methods
Probabilistic inference methods
Example Rootclaim Cases

Murder of Tair Rada in 2006 (Israel)

Tair Rada, a 13-year-old Israeli schoolgirl, was found dead, covered in her own blood and her throat sliced, inside a locked bathroom stall in her school, Nofey Golan, in the town of Katzrin. Initial suspects were Roman Zdorov, a floorer temporarily working at the school and a new immigrant to Israel from Ukraine, as well as Tair’s schoolmates. During the police investigation, Zdorov confessed to the crime but recanted soon after. Zdorov was convicted of the murder on 14th September 2010.

Conclusion: Ola Kravchenko murdered Tair Rada (98.4%)

Ghouta Chemical Attack in 2013 (Syria)

On August 21, 2013, rockets with chemical payloads landed around the neighborhood of Zamalka, located in Ghouta (an agricultural area outside of Damascus), resulting in hundreds of civilian casualties. There was initial skepticism that the Syrian government would have ordered such an attack, risking retaliation by the US. But information published by the US, UN, and Human Rights Watch during the following weeks convinced Western public opinion that the Syrian government was indeed responsible.

Conclusion: Opposition forces in Syria (Liwa al-Islam) carried out the chemical attack (96.4%)
<table>
<thead>
<tr>
<th>Why did authorities get these conclusions wrong?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chemical Attacks in Syria</strong></td>
</tr>
<tr>
<td>&quot;Trapped Prior&quot; (Confirmation Bias)</td>
</tr>
<tr>
<td>Claimed “Smoking Gun” - In reality fabricated or “sharpshooter fallacy”</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Why did authorities get these conclusions wrong?</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td><strong>Chemical Attacks in Syria</strong></td>
</tr>
<tr>
<td>Contrary Evidence Dismissed</td>
</tr>
<tr>
<td>Contrary Strong Evidence Dismissed</td>
</tr>
<tr>
<td>Video of the rocket launch claimed to be “fabricated”</td>
</tr>
<tr>
<td>Fabricated videos are very rare</td>
</tr>
<tr>
<td><strong>Murder of Tair Rada</strong></td>
</tr>
<tr>
<td>Rare Coincidences Dismissed</td>
</tr>
<tr>
<td>Mitochondrial DNA match of 1:700 - Dismissed as “10,000s of people”</td>
</tr>
<tr>
<td>Since it matched a previously identified suspect - very strong evidence of guilt</td>
</tr>
<tr>
<td>Different suspect with desire to kill random people living close to the crime scene</td>
</tr>
<tr>
<td><strong>Origin of Covid-19</strong></td>
</tr>
<tr>
<td>Impossibility is irrelevant, the likelihood is very low</td>
</tr>
<tr>
<td>Furin cleavage site insertion - “not impossible in nature”</td>
</tr>
<tr>
<td>Outbreak occurring next to a major coronavirus lab</td>
</tr>
</tbody>
</table>
Later Evidence Supported Rootclaim’s Conclusions

In June 2021, the videos were authenticated. They were matched to a field, within opposition-controlled territory, which was at the intersection of seven trajectories calculated from images of the impact sites.

The mtDNA was fully sequenced, strengthening the match from 1:700 to near certain. After spending over a decade in prison, Zdorov’s conviction was overturned.
Summary: Why Lab Leak IS Most Likely

**Priors**

Zoonosis is likely a more common cause, but lab leaks are more frequent than previously believed, and recent advances in Gain-of-Function (GoF) increase pandemic risk.

Lab leak: 23.5%

**Location**

Wuhan is a more probable origin under the lab leak hypothesis due to its prominence in GoF coronavirus research, whereas zoonotic outbreaks could have started in any other city.

Updated probability

Lab leak: 87.7%
Summary: Why Lab Leak IS Most Likely

**WIV Activity**

WIV has been likely conducting precisely the kind of research that would produce SARS-CoV-2

**Market**

Claims that the Huanan market is the epicenter are based on selection bias, and are contradicted by a substantial body of evidence

**Updated probability**

Lab leak: 96.9%

**Updated probability**

Lab leak: 94.8%
Summary: Why Lab Leak IS Most Likely

Genetics:
The genetics of SARS-CoV-2 and especially the specific way in which the Furin Cleavage Site appears are highly unlikely to occur in nature.

Lab leak: 99.9%

Other:
Some behaviors of China and WIV match a lab leak and some don’t. Same for missing evidence.

Lab leak: 99.5%

Note: “Genetics” will be discussed in Session 2, and “Behavioral” in Session 3
Results of other probabilistic analyses

Three analyses estimate the probability of a lab leak to be over 99%

- M. Weissmann: “our point estimate of the probability of a laboratory-modified source, leaked in some lab mishap is ~99.96%”
- S. Quay: “The outcome of this report is the conclusion that the probability of laboratory origin for CoV-2 is 99.8%”
- L. Nemzer: “the estimated probability of a natural origin for SARS2 is ~0.001, which is 0.1%”

Demaneuf & de Maistre: “under a reference set of input probabilities, the relative probabilities are at least 55% for a lab-related event”

Note: Not included here are purely qualitative Bayesian analyses (Washburne: “The totality of the circumstances around the emergence of SARS-CoV-2 gives probable cause for a laboratory origin”) or analyses based solely on certain type of evidence (geographical for Seymour).
3 Sessions

1. Priors: previous pandemics
   Location: Wuhan and the market

2. Genetics

3. Additional Evidence and integration of all evidence
Session 1: Contents

**Priors**
- Origins of Viral Pathogens
- Previous Pandemics
- Gain Of Function & Risk Assessments

**Location**
- Wuhan
- Huanan Market
Priors are an initial estimate of a hypothesis's probability, before examining the specific evidence of the case. Explicitly requiring it prevents prosecutor’s fallacy.

Your friend went for a routine exam and got a HIV+ result. You find out this type of test is 99% accurate - that is, the probability that the test would indicate that a person has HIV when he doesn’t is only 1%

What is the probability that your friend really has HIV?

**Common answer:** 99%

**In reality:**
- Only 0.1% of the population has undiagnosed HIV
- It is actually 10x more likely the test produced a false positive.

“Extraordinary claims require extraordinary evidence” - Carl Sagan
Historic data on pandemic origins provides a basis for estimating the initial likelihoods of each hypothesis.
Challenges in Estimating Priors for Pandemics

**Rarity of Pandemics and Lab Leaks**

Infrequent historical occurrences make estimating priors challenging.

Lab leaks, in particular, have limited documented cases.

**Limited Data on Emerging Technologies**

Rapid advancements in genetic engineering complicate current risk estimation relative to past experience.
Prior: Confirmed & Suspected Lab Leaks

Confirmed Lab Leaks
- 1979 - Anthrax
- 2004 - SARS (leaked 4 times)
- 2014 - Anthrax

Suspected Lab Leaks
- 1977 - H1N1
  - Initially, the outbreak was blamed on the consumption of contaminated meat - Eventually being confirmed to have originated from a lab
  - Suspected to be a Lab Leak due to similarity with H1N1 strain from 1946-1957
- 2021 - Zaire Ebolavirus
  - Suspected to be a Lab Leak due to similarities with strain from 2014
SARS Outbreaks: Spillovers & Lab Leaks

Lab Leaks
- Singapore - 2003
- Taiwan - 2003
- China - 2004

Spillover Events
- Foshan, Guangdong Province, China (Nov 2002) - Started with a farmer
- Guangdong Province, China (Jan 2003) - Hotel Guest
- Guangdong Province, China (Jan 2004) - Restaurant serving civets

Sources:
- https://encyclopedia.pub/entry/29846
- https://iris.who.int/handle/10665/204834

Limitation: Only known events
Prior Used: Major Flu Pandemics

- Spanish Flu (1918-1920)
  - Asian Flu (1957-1958)
  - Russian Flu (1977-1979)
  - Hong-Kong Flu (1968-1969)
- Swine Flu Pandemic (2009-2010)

Zoonosis

Lab Leak

80% zoonotic
20% lab leaks.
Major epidemics (200K dead+):

- 1918–1930 - Encephalitis Lethargica Epidemic
- 1918–1922 - Russia typhus epidemic
- 1918 influenza pandemic ('Spanish flu')
- 1957–1958 influenza pandemic ('Asian flu')
- 1968–1970 - Hong Kong flu
- 1977 - Russian Influenza
- 2009 - Swine Flu Influenza
Adjusting Priors for growth in Gain-Of-Function

No precise measure of GoF work was found. The following two indicators are used as proxies:

1. Affordable genome editing tech advances
2. Increase in BSL-4 labs

Dr. Ralph Baric, a professor at the University of North Carolina Chapel Hill and an expert in coronaviruses, provided an overview of the selection and design of pathogen properties, along with information gaps and barriers. He noted the rapid pace of advance and decreasing cost of nucleic acid synthesis; the first coronavirus to be synthesized cost roughly $42,000, a price that would now be $6,000. The largest genome currently synthesized is a 520kb mycobacterium, indicating that it is now possible to synthesize the genomes of most RNA and DNA viruses. In addition, high fidelity sequences are available for many viruses, rendering it possible to synthesize viral genomes and recover viable virus for many strains.

(1) Affordable genome editing tech advances
Adjusting Priors for growth in Gain-Of-Function

(2) Increase in BSL-4 labs

BSL-4: Maximum containment for high-risk agents, often lethal without treatment

Growth industry
Europe has the most biosafety level-4 (BSL-4) labs, and three-quarters are in urban areas. (Ten existing labs without known start dates are not shown.)
These Concerns Have Resulted In Action Against GoF

National Institute of Health (NIH) placed a 3 year ban (2014 - 2017) on GOF research related to SARS, MERS and Influenza
NASA estimated that the space shuttle programme would reach catastrophic failure once every 100,000 launches. In practice, there have been two instances in 136 attempts.

Likelihood of a catastrophic failure during a mission: 1:100,000
Mission Count: 136
Failures: 2

Nuclear Power
While the risks of nuclear power are well known, there were still several major failures.
## Prior Probabilities, After Adjustment

<table>
<thead>
<tr>
<th>Prior</th>
<th>Low</th>
<th>Impartial</th>
<th>High</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARS1</td>
<td>0.3</td>
<td>0.5</td>
<td>0.7</td>
<td>3 vs 3 cases</td>
</tr>
<tr>
<td>Flu</td>
<td>0.15</td>
<td>0.2</td>
<td>0.25</td>
<td>1 vs 4</td>
</tr>
<tr>
<td>Major pandemics</td>
<td>0.100</td>
<td>0.125</td>
<td>0.150</td>
<td>1 vs 8</td>
</tr>
<tr>
<td>Geomean</td>
<td>16.51%</td>
<td>23.21%</td>
<td>29.72%</td>
<td></td>
</tr>
<tr>
<td>Adjust for GoF and growth</td>
<td>0.5</td>
<td>1</td>
<td>1.5</td>
<td></td>
</tr>
</tbody>
</table>

|                          | 0.08254818122 | 0.2320794417 | 0.4457941465 |
| Adjusted                | 9.00%         | 23.21%       | 38.81%        |

Weighted: 23.5%
Location

Disease spillover hotspots (Bats to Humans)
Wuhan, Hubei Province, China

How likely is Wuhan to be a zoonotic epicenter given its distance from bat populations in Southern China?

How unique is the Gain of Function research being done in the WIV in Wuhan?

To what extent does this work increase the likelihood of the Covid-19 Pandemic occurring?
East Asia is home to a diverse distribution of bat species, which were the source for SARS in 2002, reaching humans through civets sold at a meat market in the Guangdong Province in southern China. The origin of another dangerous coronavirus (MERS) was in the middle east.

In order to account for the greater likelihood of a zoonotic outbreak in East Asia (the original SARS outbreak location, and where there are more wildlife-human interactions), we will measure the population of Wuhan relative only to urban areas in relevant East Asian countries.

In 2020, Wuhan's population stood at 12.3 million / Total relevant East Asian urban population 1.66 billion. 2x for large cities.

Calc: 12.3m / 1.66bn * 2 = Wuhan’s Share of east asia urban population = 1.48%.

The likelihood of Wuhan as the outbreak location is 1.48%, as Wuhan is no more likely for a bat coronavirus zoonosis than other large East Asian cities.
Likelihood of Wuhan as the outbreak location

Under the lab leak hypothesis

WIV is located in Wuhan, and is the second largest lab for coronaviruses. They regularly collect coronaviruses from SE Asia and conduct GoF research.

North Carolina lab is probably more likely, as the volume of GoF research they do on coronaviruses is greater.

WIV estimated to take 20% of all coronavirus GoF
There have been numerous reports of safety issues and biocontamination problems at the Wuhan Institute of Virology (WIV) over the years. However, it’s hard to assess whether WIV is significantly worse in that aspect than other labs.
**Biosafety Levels (BSL) & The Study of Dangerous Pathogens**

**BSL-1:** Basic safety precautions for low-risk agents
- Controlled access
- Good hygiene
- Sharps disposal
- Protective eyewear
- Laboratory bench
- Autoclave

**BSL-2:** Enhanced safety measures for moderate-risk agents
- Controlled access
- Good hygiene
- Sharps disposal
- Sharps disposal
- Personal protective equipment
- Laboratory bench
- Autoclave

**BSL-3:** Containment for indigenous or exotic agents causing serious diseases
- Controlled access
- Good hygiene
- Sharps disposal
- Minimizing spillover
- Physical containment devices
- Laboratory bench
- Autoclave

**BSL-4:** Maximum containment for high-risk agents, often lethal without treatment
- Controlled access
- Good hygiene
- Sharps disposal
- Minimizing spillover
- Physical containment devices
- Laboratory bench
- Autoclave
- Efficient decontamination system

Potentially harmful pathogens like SARS and MERS usually require at least a BSL-3. WIV had up to BSL-4.
### Designation of Bat Coronaviruses at WIV

On the official website of the Wuhan Institute of Virology in 2017, Bat coronaviruses are handled in BSL-2.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Designation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canine infectious hepatitis virus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Bovine viral diarrhea virus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Bovine infectious rhinotracheitis virus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Bovine papular stomatitis virus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Bovine rotavirus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Vesicular stomatitis virus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Capripox virus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Equine infectious anemia virus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Sika deer epidemic rabies virus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Rabbit hemorrhagic disease virus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Simian rotavirus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Mousepox virus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Guinea pig herpes-like virus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Rat coronavirus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Feline coronavirus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Grass carp hemorrhagic disease virus</td>
<td>BSL-2</td>
</tr>
<tr>
<td>Bat coronavirus</td>
<td>BSL-2</td>
</tr>
</tbody>
</table>

---

[Link to the official website](https://web.archive.org/web/20171023053516/http://www.virus.org.cn/resource/)
BSL-2 Research at Wuhan Institute of Virology

Much more troubling than the possible safety breaches, is that dangerous GoF research on bat coronaviruses was likely being conducted under BSL-2 conditions.

In this study, we constructed a full-length cDNA clone of SL-CoV WIV1 (rWIV1), an ORFX deletion mutant (rWIV1-ΔX), and a green fluorescent protein (GFP)-expressing mutant (rWIV1-GFP-ΔX).

“All experiments using live virus was conducted under biosafety level 2 (BSL2) conditions.”

## Summary of Probabilities

<table>
<thead>
<tr>
<th>Prior</th>
<th>Low</th>
<th>Impartial</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARS1</td>
<td>0.3</td>
<td>0.5</td>
<td>0.7</td>
</tr>
<tr>
<td>Flu</td>
<td>0.15</td>
<td>0.2</td>
<td>0.25</td>
</tr>
<tr>
<td>Major pandemics</td>
<td>0.100</td>
<td>0.125</td>
<td>0.150</td>
</tr>
<tr>
<td>Geomean</td>
<td>16.51%</td>
<td>23.21%</td>
<td>29.72%</td>
</tr>
<tr>
<td>Adjust for GoF and growth</td>
<td>0.5</td>
<td>1</td>
<td>1.5</td>
</tr>
<tr>
<td>Adjusted</td>
<td>0.08254818122</td>
<td>0.2320794417</td>
<td>0.4457941465</td>
</tr>
<tr>
<td>Location Wuhan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zoonosis</td>
<td>1.48%</td>
<td>1.48%</td>
<td>1.48%</td>
</tr>
<tr>
<td>Lab Leak</td>
<td>10%</td>
<td>20%</td>
<td>30%</td>
</tr>
<tr>
<td>Ratio</td>
<td>6.74</td>
<td>13.48</td>
<td>20.22</td>
</tr>
<tr>
<td>Adjust for BSL-2</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Total Wuhan</td>
<td>13.48331126</td>
<td>53.93324503</td>
<td>121.3498013</td>
</tr>
<tr>
<td>Updated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57.14%</td>
<td>94.22%</td>
<td>98.72%</td>
<td></td>
</tr>
</tbody>
</table>
Inadequate handling of dangerous pathogens at the Wuhan Institute of Virology. + Specific GoF activities in 2019 = Pandemic potential.
In 2019, virology saw a notable emphasis on understanding the furin cleavage site (FCS). This focus coincided with the period just before the potential virus leak, which also featured the presence of a furin cleavage site.

“Studies to alter pathogen properties of viruses can use several approaches, including selection pressure to drive evolution toward a phenotype as well as deliberate design. Potential opportunities might include building chimeric viruses with altered structures for the receptor for viral entry, or those that incorporate changes to other virulence determinants or that modulate host-pathogen interactions.”
DEFUSE Proposal

Components of the DEFUSE proposal are also found in SARS-COV-2.

Screen for and optimise human ACE2 binding.

Introduce human specific cleavage sites if missing from the coronavirus.

Experiment with n-glycans

A spike that is unusually well adapted to human ACE2 from the earliest cases. (Unlike SARS1)

Has a FCS, first one ever in sarbecovirus

Missing n-glycan that increases infectivity in human lung cells (but not for enteric)

The specific genetics of SARS-CoV-2, and their influence on the probabalistic analysis will be discussed in 2nd session.
### Summary of Probabilities

<table>
<thead>
<tr>
<th>Prior</th>
<th>Low</th>
<th>Impartial</th>
<th>High</th>
<th>Geomean</th>
<th>Adjusted</th>
<th>Location Wuhan</th>
<th>Matching WIV activity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SARS1</strong></td>
<td>0.3</td>
<td>0.5</td>
<td>0.7</td>
<td><strong>16.51%</strong></td>
<td>9.00%</td>
<td><strong>1.48%</strong></td>
<td>FCS: 1.5</td>
</tr>
<tr>
<td><strong>Flu</strong></td>
<td>0.15</td>
<td>0.2</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Major pandemics</strong></td>
<td>0.10</td>
<td>0.125</td>
<td>0.150</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Geomean</strong></td>
<td></td>
<td>23.21%</td>
<td>29.72%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Adjust for GoF and growth</strong></td>
<td></td>
<td>0.5</td>
<td>1.5</td>
<td>0.0825468122</td>
<td>0.2320794417</td>
<td>0.4457941465</td>
<td></td>
</tr>
<tr>
<td><strong>Adjusted</strong></td>
<td></td>
<td>23.21%</td>
<td>38.81%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Location Wuhan</strong></td>
<td></td>
<td><strong>1.48%</strong></td>
<td>1.48%</td>
<td>1.48%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Zoonosis</strong></td>
<td></td>
<td>10%</td>
<td>20%</td>
<td>30%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ratio</strong></td>
<td>6.74</td>
<td>13.48</td>
<td>20.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Adjust for BSL-2</strong></td>
<td>2</td>
<td>4</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total Wuhan</strong></td>
<td>13.48331126</td>
<td>53.93324503</td>
<td>121.3498013</td>
<td>1.11</td>
<td>12.52</td>
<td>54.10</td>
<td></td>
</tr>
<tr>
<td><strong>Updated</strong></td>
<td>57.14%</td>
<td>94.22%</td>
<td>98.72%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Matching WIV activity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FCS: 1.5</td>
</tr>
<tr>
<td>FCS</td>
<td>1.5</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human ACE2</td>
<td>3</td>
<td>5</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-glycans</td>
<td>1</td>
<td>1.5</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total WIV Activity</strong></td>
<td></td>
<td>4.5</td>
<td>15</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Updated</strong></td>
<td>85.71%</td>
<td>99.59%</td>
<td>99.98%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Best explanations under Zoonosis:**
- **FCS** - This is what enabled the jump
- **ACE2** and **N-glycans** - Coincidence and cryptic transmission

**Weighted:** 96.9%
Alternative Calculation / Sanity Check

Estimating directly the rate of emergence, rather than the ratio between the hypotheses. Definition: Probability per year of a bat coronavirus pandemic starting in Wuhan.

Zoonosis
- We have SARS and MERS which were not infectious enough to cause a pandemic and HKU-1 which wasn't lethal enough.
- So a generous estimate is once every twenty years, or 5% a year.
- 1.5% for it to happen in Wuhan
  Total: 0.075% per year

In a lab leak the following need to occur:
- Despite DEFUSE rejection, WIV started a similar project: 40%
- This involves screening for human ACE2 match. Given BANAL-52 is identical to SARS2 in that aspect, and WIV had 180 unpublished viruses, it's likely they would find a good match: 50%
- They would then optimize it for human infection (e.g. adding an FCS).
- Given that this work was done in BSL-2 a lab worker infection is quite likely: 15% per year
- Given infectivity, this would start a pandemic in Wuhan: 40%
  Total: 1.2% per year

Final Result 94% Lab Leak
Location: Huanan Seafood Market

How likely was the Huanan market to be the first location noticed?

Probability under: Zoonosis Hypothesis

Probability under: Lab Leak Hypothesis
The Zoonosis Case

The case for zoonosis in the market rests on the following claims:

- The market is one of very few places in Wuhan with live wild life.
- The market is only the 1600th most visited place in Wuhan, making this a very strong coincidence.
- One market environmental sample contained both raccoon dog DNA and SARS2.
- The positive environmental samples from the market appear to be near wild life stalls.
The Issue with Retrospective Studies on Large Datasets

- "If you torture the data long enough, it will eventually confess to anything." (Ronald Coase)
- In extensive data analysis, patterns can emerge due to the sheer volume of data.
- To claim significance, it's crucial to account for the scale of the search and the size of the search space.

Texas Sharpshooter Fallacy, related to cherry picking, hindsight bias, p-hacking, multiple comparisons problem
The handwriting evidence in the Dreyfus affair, like the shoeprint in the Tair Rada case was based on finding a few matches from a large search space, and claiming together they form a strong match. Both accused were later exonerated.

Exposed by Henri Poincaré
The proximal origin of SARS-CoV-2

Kristian G. Andersen, Andrew Rambaut, W. Ian Lipkin, Edward C. Holmes & Robert F. Garry

The proximal origin of SARS-CoV-2

NATURE MEDICINE | 26, 450–452 (2020) | Cite this article

The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic

MICHAEL WOROBEY, JOSHUA I. LEVY, LORENA MALPICA SERRANO, ALEXANDER CRITS-CHRISTOPH, JONATHAN E. PEKAR, STEPHEN A. GOLDSTEIN, ANGELA L. RASMUSSEN, MORITZ U. G. KRAEMER, CHRIS NEWMAN, MARION P. G. KOOPMANS, MAR C. A. SUCHARD, JOEL O. WERTHEIM, PHILIPPE LEMEY, DAVID L. ROBERTSON, ROBERT F. GARRY, EDWARD C. HOLMES, ANDREW RAMBAUT, AND KRISTIAN G. ANDERSEN

The emails among the authors of Proximal Origins exposed their biases in favor of zoonosis and willingness to manipulate public discourse. 4 of these authors are in the market study.

Motivations to Find Patterns

This problem becomes much worse when researchers are particularly motivated to find patterns in support of a certain conclusion. In private discussions, the researchers saw a lab leak as plausible while dismissing it in public. These emails reveal a deliberate intent to manipulate public discourse of COVID origins, driven not by scientific considerations but by political factors. They are therefore especially susceptible to the sharpshooter fallacy, finding false patterns in data.
Motivations to Find Patterns

These messages unmistakably demonstrate the researchers' awareness of their colleagues' biases for specific outcomes, as highlighted in the quote about Ron. However, it's important to note that Ron is not included in the HSM paper.

February 2, 2020, Rambaut discussed omitting the likely possibility of a lab leak to avoid a political “shit show”. Andersen said, “I totally agree.” Additionally, Andrew fears the Chinese reaction to being accusing of even an “accidental release”.

Kristian Andersen 11:47
Reading through Ron’s comments again I agree en pretty much everything he’s saying - I come to the same conclusions. Where we differ is that he’s looking for very specific evidence proving that this is unnatural (which is understandable), but except for the most simple scenario where somebody plugged a gene into a preexisting backbone, that would simply be impossible to prove.

Natural selection and accidental release are both plausible scenarios explaining the data - and a priori should be equally weighed as possible explanations. The presence of furin a posteriori moves me slightly more towards accidental release, but it’s well above my paygrade to call the shots on a final conclusion.

Andrew Rambaut 11:53
Given the shit show that would happen if anyone serious accused the Chinese of even accidental release, my feeling is we should say that given there is no evidence of a specifically engineered virus, we cannot possibly distinguish between natural evolution and escape so we are content with ascribing it to natural processes.

Kristian Andersen 11:56
Yup, I totally agree that that’s a very reasonable conclusion. Although I hate when politics is injected into science - but it’s impossible not to, especially given the circumstances. We should be sensitive to that. (plus none of this matters at the moment)
Motivations to Find Patterns

The emails revealed efforts to deflect questions about the possibility of COVID-19 originating from a lab, using strategies that appear more political than scientific.
Retrospective Study Risk

In a robust probabilistic analysis, we should be extremely careful when incorporating studies relying on retrospective pattern matching in large datasets. The Huanan Market study in its current form should be ignored.

We of course apply this standard to all hypotheses. In Bruttel et al, researchers found patterns supporting genetic manipulation at a P-value of 1 in 10,000,000. While we generally agree with the study, we couldn’t incorporate it due to these limitations.

It is very difficult to identify the source of the biases in such studies. Luckily, the market study has already been thoroughly analyzed by others.
Bias in Early Cases

Early in the pandemic, diagnoses were partially given based on whether a patient had a connection to the Huanan Seafood Market.

Until January 18th, 2020 a connection to the market was a requirement for a Covid-19 diagnosis because it was thought to be the source. This led to a selection bias wherein only those connected to the seafood market were diagnosed with covid-19, which further fueled speculation that Huanan was the epicenter and the likely spillover location.

Early in the pandemic, diagnoses were partially given based on whether a patient had a connection to the Huanan Seafood Market.

Due to SARS experience, healthcare workers were more inclined to report new diseases with connection to a market.
Evidence of Bias

Furthermore, there seem to be around 90 December cases missing. There is no way to know why this happened and what biases it introduces. Any claims based on this data are of low confidence.
According to an analysis of 168 COVID-19 cases with or without a history of exposure to the Huanan Market, the first case had no history of exposure to the Huanan Market, and the incidence of cases with and without history of exposure to the Huanan Market basically increased simultaneously. However, the cases with exposure history decreased after reaching the peak on 25 December, and the cases without exposure history reached the peak on 30 December, as shown in Fig. 10.

Rates with and without connection to HSM “increased simultaneously”. If HSM is GO a likelier distribution would be mostly HSM-related patients at first and only then the rest of Wuhan catching up - but HSM and Wuhan are neck and neck.

A large portion of the early cases were not linked to the Huanan market. In general, early cases are unreliable, as there was no published effort to contact trace or find the earliest cases. This in itself is evidence for a lab leak.

https://www.science.org/doi/10.1126/science.abp8715#
Both of the first two cases were visitors (11 and 12 December); the first vendor case appeared on 13 December.

Given that the first cases were only a day apart, they could feasibly have been infected at the same time due to the length of the incubation period. However, if the market were the source, we would expect a significant majority of the first cases to be vendors.
Bias: Case of Mr. Chen

A good indication of selection bias is the case of Mr Chen:

- Fever 16 December
- Did not go to any wet market, lived and worked exclusively locally in rural Jiangxia (30km south)
- Was diagnosed only because a relative worked in the ER of Wuhan Central Hospital, across the river.

This implies widespread infection early December.
The claim that the market was not a likely location (absent wild life) due to its low-traffic was based primarily on social check-in data. Social check-ins are popular among younger crowds, at trendy locations. The median age of those infected at the market was 56, and a seafood market is not a place people would check-in to. In reality, Huanan market is fairly busy with 10,000 visitors a day.

Many of the early cases were associated with the Huanan market, but a similar number of cases were associated with other markets and some were not associated with any markets. Transmission within the wider community in December could account for cases not associated with the Huanan market which, together with the presence of early cases not associated with that market, could suggest that the Huanan market was not the original source of the outbreak. [...] No firm conclusion therefore about the role of the Huanan market in the origin of the outbreak, or how the infection was introduced into the market, can currently be drawn.”

Page 7 of the WHO report
Source: https://www.who.int/publications/i/item/who-convened-global-study-of-origins-of-sars-cov-2-china-part
Environmental samples from the market show a number of areas with a stronger rate of positive samples in the south-west corner of the market place that also had live animal stalls, specifically racoon dogs.

Unlike environmental samples, cases are evenly distributed across the market.

Initially, positive samples were shown to be close to wildlife stalls. However, the sample was biased, as those areas were oversampled so, naturally, they produced more positive samples.

Once corrected, the positive samples better match the toilets and a Mahjong room, ideal for airborne virus transmission like SARS-CoV-2.

https://www.science.org/doi/10.1126/science.abp8715#
An environmental sample from a raccoon dog cage (not the animal itself) was positive for SARS2, indicating that raccoon dogs could potentially be the intermediate host for the virus.

There was only one SARS2 RNA read, and raccoon dog mtDNA was in fact negatively correlated with SARS2 RNA.

https://www.npr.org/sections/goatsandsoda/2022/03/03/1083751272/striking-new-evidence-points-to-seafood-market-in-wuhan-as-pandemic-origin-point

https://www.biorxiv.org/content/10.1101/2023.04.25.538336v1
Wildlife Inversely Correlated with SARS2

Positive environmental samples around raccoon dog stalls were in fact lower than other animals. This negates the hypothesis that raccoon dogs were the intermediary species.

Note that in any case, infected raccoon dogs in the market isn't strong evidence for zoonosis. If they are as susceptible to SARS-CoV-2 as claimed, they could have contracted it from humans. Strong evidence would be them being infected by an earlier strain phylogenetically.

**Association between SARS-CoV-2 and metagenomic content of samples from the Huanan Seafood Market**

Jesse D Bloom

*Virus Evolution*, Volume 9, Issue 2, 2023, vead050,
https://doi.org/10.1093/ve/vead050

Published: 24 August 2023
Positive environmental samples are not positively correlated with wildlife stalls in particular.
Market Strains not Ancestral

Two Lineages of SARS-CoV-2 observed:
Lineage A & Lineage B

Lineage A is ancestral to Lineage B. This is a contested claim, and we will explain why there is strong evidence for this in session 2.

All the market cases were lineage B, not the earlier lineage A.

Early appearance of two distinct genomic lineages of SARS-CoV-2 in different Wuhan wildlife markets suggests SARS-CoV-2 has a natural origin

SARS-CoV-2 coronavirus
Unbiased Cases not Centered on HSM

When looking at the spread of reported cases on the Weibo social media platform, we can see they are not centered around the Huanan market, but instead correlate with population density.

https://www.science.org/doi/10.1126/science.abp8715#

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8545724/
Unbiased Cases not Centered on HSM

Interesting to note the study did not show the population center in the Weibo data...

DOI: 10.1126/science.abp871
In fact, when looking only at early Weibo reports, the epicenter is actually near the WIV. We think this is a coincidence as we don’t really expect lab leak cases to center around a lab, but rather a location frequented by a lab worker. However, it demonstrates the risk of retrospective pattern matching.

https://www.mdpi.com/2220-9964/9/6/402/htm
Deaths data (which lag behind cases 2-3 weeks) don’t show any bias towards the market district (Jianghan). All cause mortality is actually more dominant across the river, where WIV is located.

Source: Fig. 20 and 21 in WHO-convened global study of origins of SARS-CoV-2
Unbiased Cases not Centered on HSM

The district of the Huanan market actually showed lower seropositivity rates in April 2020, relative to neighboring districts.

Wuhan went into lockdown on 23rd January 2020, thus seropositivity is a good indication of early cases.

When looking at unbiased data - seropositivity, social networks, mortality - the bias to the market disappears.

Table S3. Case counts, Weibo help seeker counts, total population, seropositivity, and case fatality rate per district in the city of Wuhan.

<table>
<thead>
<tr>
<th>Neighborhood</th>
<th>Group</th>
<th>WHO cases (Dec)</th>
<th>Total cases (March)**</th>
<th>Total cases (March)***</th>
<th>Weibo help seekers***</th>
<th>Population *</th>
<th>Cases per 100,000 (March)***</th>
<th>Seropositivity (April)*</th>
<th>Seropositivity (April)**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hanyang</td>
<td>North-urban</td>
<td>9</td>
<td>4691</td>
<td>4820</td>
<td>84</td>
<td>664,202.00</td>
<td>706.26</td>
<td>7.88</td>
<td>4.6</td>
</tr>
<tr>
<td>Jiangan</td>
<td>North-urban</td>
<td>28</td>
<td>6570</td>
<td>7030</td>
<td>111</td>
<td>962,695.00</td>
<td>682.46</td>
<td>11.12</td>
<td>6.5</td>
</tr>
<tr>
<td>Jianghan</td>
<td>North-urban</td>
<td>54</td>
<td>5199</td>
<td>5448</td>
<td>72</td>
<td>729,704.00</td>
<td>712.48</td>
<td>7.83</td>
<td>5.2</td>
</tr>
<tr>
<td>Qiaokou</td>
<td>North-urban</td>
<td>12</td>
<td>6863</td>
<td>7143</td>
<td>110</td>
<td>868,702.00</td>
<td>790.03</td>
<td>13.08</td>
<td>11.2</td>
</tr>
<tr>
<td>Hongshan</td>
<td>South-urban</td>
<td>10</td>
<td>6990</td>
<td>7403</td>
<td>1</td>
<td>1,677,298.0</td>
<td>0</td>
<td>416.74</td>
<td>5.62</td>
</tr>
<tr>
<td>Qingshan</td>
<td>South-urban</td>
<td>4</td>
<td>2788</td>
<td>2979</td>
<td>63</td>
<td>528,894.00</td>
<td>527.14</td>
<td>5.54</td>
<td>3.8</td>
</tr>
<tr>
<td>Wuchang</td>
<td>South-urban</td>
<td>6</td>
<td>7484</td>
<td>7726</td>
<td>133</td>
<td>1,282,800.0</td>
<td>0</td>
<td>583.41</td>
<td>5.15</td>
</tr>
<tr>
<td>Dongxihu</td>
<td>North-suburb</td>
<td>12</td>
<td>2465</td>
<td>2544</td>
<td>40</td>
<td>584,803.00</td>
<td>421.51</td>
<td>5.07</td>
<td>2.3</td>
</tr>
<tr>
<td>Huangpu</td>
<td>North-suburb</td>
<td>13</td>
<td>2126</td>
<td>2204</td>
<td>10</td>
<td>1,011,897.0</td>
<td>0</td>
<td>210.10</td>
<td>1.62</td>
</tr>
</tbody>
</table>
“Despite the testing of more than 80,000 samples from a range of wild and farm animal species in China collected between 2015 and March, 2020, no cases of SARS-CoV-2 infection have been identified.”
Comparing to SARS1

We would expect a coronavirus that also originated from a market to share similar outbreak characteristics to SARS1.

SARS1:
9 out of 23 of the animal vendors tested positive for SARS1.

SARS2:
None of the animal vendors tested positive for SARS2.

SARS1:
Multiple locations had SARS1 outbreaks, including a restaurant.

SARS2:
All other outbreaks were connected to the original outbreak in Wuhan.

Source: https://www.washingtonpost.com/opinions/2023/03/02/china-covid-origins-answer/
Comparing to SARS1

SARS Infections Over Time
Comparing to Xinfadi

The Xinfadi market outbreak

- On June 11, 2020, an outbreak of COVID-19 was traced to Xinfadi Market in Beijing, after 56 days without any new cases in the city.
- From the genetic testing of the virus from the infected and the market, we know the source was external to the market and not a spillover.
- Both Xinfadi and HSM outbreaks lasted a month before the markets were closed, giving us an identical time frame to look at.
- This provides us with a real-life simulation of a COVID-19 market outbreak from an exterior source, which according to the zoonosis hypothesis, should not look like a market outbreak based on a spillover.
## Comparing to Xinfadi

<table>
<thead>
<tr>
<th></th>
<th>HSM</th>
<th>Xinfadi</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of infected with direct exposure to location</strong></td>
<td>55 (33%) (WHO, annexes, p178)</td>
<td>261 (78%) (Article 4.1)</td>
</tr>
<tr>
<td><strong>Vendors \ Workers % of direct exposure patients</strong></td>
<td>30 (55%)</td>
<td>177 (52.8%)</td>
</tr>
<tr>
<td><strong>Positive environmental samples</strong></td>
<td>Equipment for animals and floor drain</td>
<td>Chopping board and floor drain</td>
</tr>
</tbody>
</table>

When the market is known to be the source of an outbreak, we see 2.4 times the number of direct exposure patients, even though HSM had no other sources of infections outside the market (under zoonosis), whereas Xinfadi is during the pandemic.

This increases the likelihood HSM was not the single source in Wuhan or, at the very least, eliminates this factor which was used as support for Zoonosis.

The proportion of vendors is the same, indicating that contrary to our opponents’ assertions, the number of vendors is unrelated to spillover and is the expected outcome in any market outbreak.

Zoonosis proponents claim positive drain samples in HSM came from animals but Xinfadi shows positive drain samples are expected even without a spillover in the market, which was also corroborated by a [simulation experiment in Xinfadi](#) to test how COVID spreads in markets.
Comparing to Xinfadi

Papers describing how Xinfadi became such a large outbreak point to the same conditions that existed in HSM and most markets:

"The complete genome sequence analysis of the SARS-CoV-2 from Xinfadi further confirmed that the source of the epidemic was not a new "overflow" from the host or intermediate host. The Xinfadi market has become the SARS-CoV-2 transmission "hub", the most important reason being that the market environment is relatively wet and cold. One of the main characteristics of the virus is that it is sensitive to heat and not to cold. In such an environment, the virus can survive for a long time. Moreover, it is also closed and poorly ventilated, which also contributes to the spread of the virus. For example, if a person infected with SARS-CoV-2 sneezes, it is difficult to spread out, and the droplets may settle on the ground and pollute other places after flushing and sweeping. These conditions help to explain the occurrence of a large number of cases in a short period of time. However, this does not mean that the market itself is the source of the virus."
Why was the HSM the First Superspreader Event to be Noticed?

Best Explanation for the Huanan market being the first location noticed

**Zoonosis**

- **“Animal”**
  - Under zoonosis we want to weigh the conditional probability of each and choose the higher one - the "Best Explanation"

- **“Other”**

**Lab Leak**

- **“Other”**
  - Under lab leak it must be “other”, as there is no animal

We want to assess the conditional probability of the evidence under both explanations, "Animal" and "Other", and choose the best explanation for each hypothesis.
For animal, we need to estimate the probability that HSM would be where the virus spillover occurs, but also account for all the evidence.

HSM is one of several places with wildlife. There are three other known markets. There could be others (e.g. illegal trade), but there are also other places that would be connected to wildlife in retrospect such as restaurants (happened in SARS1) and TCM hospitals.

We estimate HSM accounts for 15%, conservatively 30%.

Best explanation: Only a single infected animal, and its vendor not infected. 15%, conservatively 25%

No animal vendors were infected, compared to 9 out of 23 in SARS.

Best explanation: The single animal was missed or RNA was lost (e.g. cage cleaned) 40%, conservatively 60%

All testing in the market looks unrelated to wildlife: Positive raccoon dog sample, but only one read and negative correlation overall. Vendors evenly distributed. All animals in the market and related farms were negative.

Best explanation: Only a single infected animal, and its vendor not infected.
All infections are lineage B, as are all but one environmental sample. While elsewhere A is a third of cases.

Best explanation: The animal was A and infected two (or more) people. One infected people mostly outside the market, the other mutated to B and infected the market 20%, conservatively 30%.

Biased data (WHO cases) centered on the market, but unbiased (seropositivity, mortality, Weibo) isn’t centered on the market, but

Best explanation: The noise introduced by traffic of a modern city makes centering less likely.

50%, conservatively ignored

Early cases associated with market, but not strongly, Chen far away

Best explanation: Lineage A (despite being less infectious) spread outside the market for a while before lineage B dominated the market.

30%, conservatively 50%
Mahjong: hall in the market, at highest positive rate area, and many early cases connected (Not evaluating Mahjong specifically, but any high risk activity)

Best explanation: One of the early patients played mahjong and the conditions caused most of the spread.

30%, conservatively 40%

We estimate this is far likelier than the current explanation offered by zoonosis supporters of multiple animals infecting only the HSM, with two lineages, one of them infecting only outside the market, and the animals leaving no traces anywhere.

Final: 0.0081%, conservatively 0.27%
For other, we need to estimate the probability of the coincidence of HSM being the first place noticed, which is not strong as claimed, because:

How likely is the infected lab worker (or their close contact) to visit HSM? HSM has more traffic than claimed by check-ins: 10000/d

0.1% of population (assuming people visit one place a day where they're likely to infect others), conservatively 0.05%

Location that can become a superspreader location. Ideally has:

1. High traffic - to bring the virus in when it is still rare (already accounted for above)
2. Permanent residents - to amplify the virus locally. Without this (e.g. a train station) the location just serves to amplify the virus all over the city
3. Enclosed, poor ventilation. Few places have all three.

- Of the top 20 Weibo check-in locations (where HSM is 1,676th), 6 are shopping areas, mostly outdoor, 5 are universities, 5 are transportation hubs, and 1 is a commercial building. None of these are as susceptible as HSM. The indoor malls are close, but are likely far better ventilated, and involve less close interaction among the permanent residents.
- Interestingly, the next superspreader event after lockdowns was also a market (Xinfadi wholesale agricultural market), demonstrating it is a likely location regardless of wildlife.

x10 increase, conservatively x4
Check-ins Analysis: The market has high superspreader location potential

Note the relationship is not linear: A small advantage in the criteria is a large increase in the probability of becoming the first superspreader location.
Health system was tuned to notice new diseases in markets following SARS1. In addition, the older population is much more likely to present to hospitals, cases may have been censored by China, and the market is next to top-tier hospitals (this may be a major factor, with small hospitals not keeping samples, not having advanced diagnostics etc. This is discussed in detail in https://washingtonpost.com/documents/a5102da1-9b47-4e11-b615-9f59b7d3a3c3.pdf, which was too long for us to analyze properly). This is to say there were likely a few more superspreader locations at the time but they took longer to notice, and were thus considered part of the city-wide outbreak despite starting at the same time as HSM.

So the probability of an animal spillover at HSM (with that evidence) is far smaller than the HSM being noticed first for other reasons. Therefore “other” is the best explanation.

Since “other” is chosen for both hypotheses the likelihood ratio is 1, which means no weight, and can be ignored.
The Mahjong Halls Hypothesis

While all the evidence points against the market being the epicenter, it is true that it was the earliest superspreader event to be noticed. The HSM Mahjong hall explains this while matching the evidence far better than the wildlife hypothesis.

- Mahjong rooms are perfect:
  - Crowded
  - Unventilated
  - Older population
  - Repeat visitors
    - Allow for an exponential growth within that population, even when the overall infection rate is low.

- The disease initially spread through Mahjong rooms in Wuhan.
- The room in the HSM was especially susceptible due to its high density location near a major train station.
- From there it spread to the vendors and the entire HSM, due to similar reasons.
- The HSM was then noticed early due to bias from SARS1, and proximity to top tier hospitals.
"I heard that four people were playing at one table, and all four got sick" [source]

"had heard about [...] some mah-jongg infections" [source]  

"most of the 27 infected people first reported by the government[56] had a common hobby - they like to play mahjong." [source]

"a friend of mine is in the hospital with it. [...] He thinks he might have got it while playing mahjong, reckons the air in the room was stuffy, a lot of people in there." [source]

"heard on the morning of December 31 that three store owners who often played mahjong at stalls near 10th Street were hospitalized two days ago. He sent WeChat greetings to one of the bosses, who replied "infected with the virus"." [source]

"Our boss was sick on December 19 (2019), sick before playing mahjong with a player who sells frozen chicken and frozen duck, who was also admitted to hospital for pneumonia." [source]

"The owner of the shop was a man in his 40s and 50s who usually likes to stay up late and play mahjong. Recently, he “didn't care” after catching a cold. He didn't go to the doctor at first. After a few days, he found that his condition was getting worse before going to the hospital. He was eventually diagnosed with pneumonia." [source]

"At the end of December, we had already judged that this infectious virus was related to the South China Seafood Market because most of their patients knew each other. [...] They are either a family or people who play mahjong together." [source]  

Testimonies from 8 people, collected by Chinese & Western media, include people working in HSM, a Wuhan journalist and the head of the respiratory department at Jinyintan hospital. Together, they say more than 35 early HSM cases were among Mahjong players.
The Mahjong Halls Hypothesis

Location

The Majhong Hall was located in a small corner of the market, in a closed space - next to the public bathrooms of the Huanan Market.

Huanan Market Study Results

Environmental samples exhibit a strong bias toward a specific corner of the market, coinciding with the location of the Majhong Hall.

Official data records only 2 Majhong players, but witness accounts suggest around 35.
The Mahjong Halls Hypothesis

Better fits the evidence

High positivity rate near the Mahjong room

Even distribution of vendor cases across the market, and visitors among early cases

All evidence against wildlife: Lineage B, early cases not connected to market, no infected animals, no wildlife vendors infected, reverse correlation with wildlife, unbiased location mismatch (seropositivity, mortality, Weibo)
If the Market is the Epicenter

Another potential for market as source: CDC Move

The CDC moved its laboratories on the 2nd of December 2019 to a location close to the market, leading to a higher probability of a breach of containment occurring during the move and thus strengthening the lab-leak hypothesis.

CDC move could be relevant if they interacted with the WIV, for example by:
- exchange of equipment
- meetings between personnel

“In 2003, during the SARS outbreak a 27-year old doctoral student was infected from SARS while working on the virus. An investigation showed that because the Department of Pathology BSL-2 laboratories were being renovated, mixed BSL-2/BSL-3 activities were in progress in the BSL-3 facility - jeopardising good safety practices. The move between laboratories leaves considerable room for mistakes to occur and procedures to go astray as the 2003 incident showcases.
Expectations vs reality

If the Huanan seafood market were the epicenter of the pandemic and the virus arrived to humans via zoonosis, we would expect most data showing Covid traveling from an intermediate animal in HSM outwards to the rest of the market and then into Wuhan, yet this is what we find:

<table>
<thead>
<tr>
<th>Expected (under zoonosis)</th>
<th>Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earliest patients would be HSM-related and Wuhan “catching up” with some delay</td>
<td>Wuhan &amp; HSM cases “increased simultaneously” (WHO)</td>
</tr>
<tr>
<td>Infected animals in the market</td>
<td>None</td>
</tr>
<tr>
<td>Infected vendors selling possible intermediate species</td>
<td>None</td>
</tr>
<tr>
<td>First cases from market would be people spending the longest time there, especially next to the infected animals (e.g., vendors)</td>
<td>First 2 HSM-related cases are visitors</td>
</tr>
<tr>
<td>First cases would have a shared location inside HSM or repeat locations individually</td>
<td>“it was found that none of the 5 [earliest] cases had similar exposure history [Inside HSM]”[WHO]</td>
</tr>
<tr>
<td>Amongst infected vendors, infection rate would be highest among (alive or frozen) animal vendors</td>
<td>Highest infection rate was among the vegetable vendors</td>
</tr>
<tr>
<td>Market infections would be from the strain closest to nature</td>
<td>Lineage B dominates, while the earlier lineage A is found elsewhere in the city</td>
</tr>
</tbody>
</table>
### Summary

<table>
<thead>
<tr>
<th>Prior</th>
<th>Low</th>
<th>Impartial</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARS1</td>
<td>0.3</td>
<td>0.5</td>
<td>0.7</td>
</tr>
<tr>
<td>Flu</td>
<td>0.15</td>
<td>0.2</td>
<td>0.25</td>
</tr>
<tr>
<td>Major pandemics</td>
<td>0.100</td>
<td>0.125</td>
<td>0.150</td>
</tr>
<tr>
<td>Geomean</td>
<td>16.51%</td>
<td>23.21%</td>
<td>29.72%</td>
</tr>
<tr>
<td>Adjust for GoF and growth</td>
<td>0.5</td>
<td>1</td>
<td>1.5</td>
</tr>
<tr>
<td>Adjusted</td>
<td>9.00%</td>
<td>23.21%</td>
<td>38.81%</td>
</tr>
<tr>
<td>Location Wuhan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zoonosis</td>
<td>1.48%</td>
<td>1.48%</td>
<td>1.48%</td>
</tr>
<tr>
<td>Lab Leak</td>
<td>10%</td>
<td>20%</td>
<td>30%</td>
</tr>
<tr>
<td>Ratio</td>
<td>6.74</td>
<td>13.48</td>
<td>20.22</td>
</tr>
<tr>
<td>Adjust for BSL-2</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Total Wuhan</td>
<td>13.48331126</td>
<td>53.93324503</td>
<td>121.3498013</td>
</tr>
<tr>
<td>Updated</td>
<td>57.14%</td>
<td>94.22%</td>
<td>98.72%</td>
</tr>
<tr>
<td>Matching WIV activity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCS</td>
<td>1.5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Human ACE2</td>
<td>3</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>N-glycans</td>
<td>1</td>
<td>1.5</td>
<td>2</td>
</tr>
<tr>
<td>Total WIV Activity</td>
<td>4.5</td>
<td>15</td>
<td>60</td>
</tr>
<tr>
<td>Updated</td>
<td>85.71%</td>
<td>99.59%</td>
<td>99.98%</td>
</tr>
</tbody>
</table>

**Weighted: 95%**