# THINKING PROBABILISTICALLY VI

WARREN R HUGHES – 26th December 2021

*Thinking Probabilistically* [available on this website or at the digital object identifier (doi) https://doi.org/10.17265/1537-1506/2021.05.002] outlined the rationale for consideration of all possible scenarios when faced with making a decision or recommendation. In this sixth sequel, the uncertainty surrounding the origin of COVID-19 is further examined using the methodology outlined in *Thinking Probabilistically* or *Thinking Probabilistically Revisited* (this website).

In July 2020, the WHO organized a fact-finding mission to Wuhan in China to investigate the origin of the COVID-19 virus. Then in January 2021, a team of international and Chinese experts spent two weeks investigating the outbreak and concluded that four scenarios were possible. No probabilities were provided by the conference, but possible probabilities are explored here using the "more likely" methodology discussed further in *Structuring Probability Assessments*. The paper is available on this website or alternatively at https://doi.org/10.17265/1537-1506/2020.05.003. The four scenarios outlined at the conference are ranked in order of increasing likelihood in Table 1.

**TABLE 1: WHO POSSIBLE SCENARIOS AND LIKELIHOODS FOR THE ORIGIN OF COVID-19**

| Event | Scenario | Likelihood |
|---|---|---|
| LE | Transmission through a laboratory incident | Extremely unlikely |
| FF | Transmission through frozen food | Possible |
| BH | Direct transmission from bat to human | Possible to likely |
| AH | Transmission through an intermediate animal | Likely to very likely |

A good review of the lab-leak hypothesis is given in ***Nature*** at doi: https://doi.org/10.1038/d41586-021-01529-3. See also "The origins of viruses: discovery takes time, international resources, and cooperation," ***The Lancet***, September 30, 2021, at doi: https://doi.org/10.1016/S0140-6736(21)02180-2.

In Table 2, the routine probability calculations are shown for a variety of pairwise values on the four possibilities in Table 1. Note these are **not** probabilities derived by the WHO mission.

**TABLE 2: PROBABILITIES ON VIRUS ORIGIN FOR VARYING PAIRWISE JUDGMENTS**

| Scenario | Pairwise Values | Probabilities | | | | |
|---|---|---|---|---|---|---|
| | | LE | FF | BH | AH | SUM |
| Basic scenario | 1, 7, 2, 3 | 0.0156 | 0.1094 | 0.2187 | 0.6563 | 1.0000 |
| FF/LE 10 times more likely | 1, 10, 2, 3 | 0.0110 | 0.1099 | 0.2198 | 0.6593 | 1.0000 |
| Increased values for more likely origins | 1, 10, 3, 4 | 0.0062 | 0.0621 | 0.1863 | 0.7454 | 1.0000 |
| FF/LE drops, AH/BH increases | 1, 8, 3, 5 | 0.0065 | 0.0523 | 0.1569 | 0.7843 | 1.0000 |
| FF/LE drops further | 1, 7, 3, 5 | 0.0075 | 0.0522 | 0.1567 | 0.7836 | 1.0000 |
| FF/LE 10 times more likely | 1, 10, 3, 5 | 0.0052 | 0.0524 | 0.1571 | 0.7853 | 1.0000 |

The resulting probabilities in Table 2 were previously discussed in *Thinking Probabilistically IV* (see website). And in that paper, we formulated an alternative analysis to the above focused on the Non-China/China options for the origin of the virus. The alternative hypotheses are reproduced below.

1. NonChinese - The virus originated outside of China and is the least likely hypothesis.
2. WuLab -The virus initially escaped from a Chinese laboratory in Wuhan (Wuhan Lab Escape).
3. WuMkt - The virus emerged naturally (animal to human) from a Wuhan wet-market (most likely).

Initially it was considered that hypothesis 2 was the most likely origin of the virus but virologists consider the genome sequence of the virus is missing the fingerprints that would be present had it been engineered in a lab. Consequently, the "more likely" values in Table 3 have been re-thought with the routine probability calculations as detailed below.

**TABLE 3: PROBABIILITIES OF COVID-19 VIRUS ORIGIN**

| Virus Origin | Pairwise Value | Compound Likelihood | Probability | Percent Probabilities | |
|---|---|---|---|---|---|
| | | | | Current | Previous |
| NonChinese | 1.00 (base value) | 1.00 | 1/51 = 0.0196 | 2 | 3 |
| WuLab | 10.00 (10 x more likely) | 1.00 x 10.0 = 10.00 | 10/51 = 0.1961 | 20 | 24 |
| WuMkt | 4.00 (4 x more likely) | 10.00 x 4.0 = 40.00 | 40/51 = 0.7843 | 78 | 73 |
| | | | | | |
| | | **51.00** | **1.0000** | **100** | **100** |

.

The above probabilities in Table 3 now make the Wuhan market the most likely origin of the virus based on this analysis with a 5% increase in probability over the previous calculations. Table 3 shows odds of almost 4:1 on for a Wuhan market origin, up from 3:1 previously. According to scientists, however, we may never discover the exact origin of the virus. ***The Lancet*** article referenced above reviews the problems involved.

The virus origin analysis illustrates a problem involving order-of-magnitude in the pairwise values. Can we really be certain that one event is 10 times "more likely" than another and not 12 times "more likely"? Of course we cannot, even allowing for the fact that these values only get us to the "ballpark" within which the final distribution is determined. All we know is the difference is large for a pairwise value. Using sensitivity analysis as illustrated above gives a range of alternative probabilities and a possible solution. Table 4 summarizes the easily calculated probabilities for significant changes in the pairwise values.

**TABLE 4: PROBABILITIES ON VIRUS ORIGIN FOR VARYING PAIRWISE JUDGMENTS**

| Scenario | Pairwise Values | Probabilities | | | |
|---|---|---|---|---|---|
| | | NonChinese | WuLab | WuMkt | SUM |
| Basic scenario as in Table 3 | 1, 10, 4 | .0196 | .1961 | .7843 | 1.0000 |
| WuLab lower | 1, 8, 4 | .0244 | .1951 | .7805 | 1.0000 |
| WuMkt higher | 1, 8, 5 | .0204 | .1633 | .8163 | 1.0000 |
| WuLab higher | 1, 12, 4 | .0164 | .1967 | .7869 | 1.0000 |
| WuMkt lower | 1, 12, 3 | .0204 | .2449 | .7347 | 1.0000 |
| WuLab lower | 1, 10, 3 | .0244 | .2439 | .7317 | 1.0000 |
| WuMkt higher | 1, 10, 5 | .0164 | .1639 | .8197 | 1.0000 |

Table 4 shows that large changes in the pairwise values lead to a 2% - 3% percentage probability range for the NonChinese origin, 16% - 25% for WuLab and 73% - 82% for a WuMkt origin. A final percentage distribution for NonChinese, WuLab and WuMkt respectively using approximate midpoints of the ranges could be say 3%, 20% and 77%. Political sensitivities may see another allocation as in say 5%, 20%, 75% giving a slightly larger probability to an already small chance. Alternatively, perhaps 5%, 15%, 80% maximizes the natural transmission origin and minimizes the lab escape theory.

Word Count 952, 26th Dec 2021.