

THINKING PROBABILISTICALLY III

WARREN R HUGHES – 20 August 2021

Thinking Probabilistically (available on this website) outlined the rationale for consideration of all possible scenarios when faced with making a decision or recommendation. In this sequel, the uncertainty surrounding the origin of COVID-19 is further examined using the methodology outlined in *Thinking Probabilistically*.

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. The four scenarios are ranked in order of increasing likelihood in Table 1. No probabilities were provided by the conference, but possible probabilities are explored here using the “more likely” methodology.

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>. The following re-examines the virus origin problem using the “more likely” methodology. The routine calculations (see Table 3) are summarized below in a Table 2 sensitivity analysis. The methodology is discussed further in *Structuring Probability Assessments*, which is available on this website or alternatively at <https://doi.org/10.17265/1537-1506/2020.05.003>.

TABLE 2: PROBABILITIES ON VIRUS ORIGIN FOR VARYING PAIRWISE JUDGMENTS

Scenario	Pairwise Values	Probabilities				
		LE	FF	BH	AH	Sum
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

Table 2 shows that the higher the latter pairwise values the lower the probabilities of the preceding events. High pairwise values for the less likely events early in the ordering generate higher probabilities for the more likely events. Even though the FF/LE “more likely” value changes in the range 7 – 10, the probability of LE remains close to 1%.

Previously 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. NonCh - The virus originated outside of China and is the least likely hypothesis (Non-Chinese).
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).

Previously 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: PROBABILITIES OF COVID-19 VIRUS ORIGIN

Virus Origin	Pairwise Value	Compound Likelihood	Probability	Percent Probability
NonChinese	1.00 (base value)	1.00	$1/25 = 0.04$	4
WuLab	8.00 (8 x more likely)	$1.00 \times 8.0 = 8.00$	$8/25 = 0.32$	32
WuMkt	2.00 (twice as likely)	$8.00 \times 2.0 = 16.00$	$16/25 = 0.64$	64
		25.00	1.00	100

Note that if the 8 times “more likely” were replaced by 10 times, the probabilities would be respectively 0.032, 0.323 and 0.645. That is, the probability of a Non-Chinese origin drops by 1% with the increase to 10 times “more likely” while the probabilities of the Chinese origins rise slightly. As noted, the higher the initial pairwise values, the higher the probabilities for events later in the ordering. The first pairwise number can be chosen to result in (say) a 4% Non-Chinese probability irrespective of its effect on later probability values. That is, the fixed point here is the decision-maker’s (DM’s) 4% chance for a Non-Chinese origin with the first pairwise value (8 to 10) varied until this probability is achieved. The subsequent 96% probability will be shared between the other hypotheses as determined by their respective “more likely” pairwise values later in the ordering.

The above probabilities in Table 3 now make the Wuhan market the most likely origin of the virus based on this analysis. According to scientists, however, we may never discover its exact origin. Current analysis seems to suggest the most likely route was from animal to human possibly at the Wuhan wet-market. This makes the original WHO conclusion of “very likely” as in Table 1 prescient. Of course, the lab-leak hypothesis is appealing as it makes for better media speculation and the fact that there have been previous leaks from other labs. In this case, the proximity of the WIV lab to the Wuhan wet-market naturally induces the lab-leak, LE or WuLab explanations for the COVID-19 phenomenon.

Word count 826, 20th August 2021