Statistical Logic of NIPT (Distinguishing Noise from Trisomy) +++ 2026

Tuan Tran, MD. PhD., Unit of Genetic & Genomic Medicine, Family Hospital, Da Nang.

A reader has asked a QUESTION:

In a NIPT, when we already have a complete “count” of DNA reflecting 46 chromosomes, including XY or XX. Then we have another 21, how to know that 21 is from a new shedding or really a third 21, making Trisomy 21?
If this sentence could be rephrased, it would be:
In a pool of cell-free DNA where we already expect a complete genetic blueprint (46, XX or 46, XY), how can we be certain that an increased concentration of chromosome 21 fragments represents a **fetal trisomy** rather than just **statistical noise** or a higher rate of placental shedding?

ANSWER:

This is a **FANTASTIC** question that gets to the heart of "bioinformatics"—the math used to interpret the DNA fragments. To a student, it might seem like the machine is just counting to 46, but in reality, the machine is performing a **relative proportion analysis**.
Here is how we differentiate a "normal" count from a "third" chromosome 21.

1. The Genomic "Baseline"

The lab doesn't just count the baby's DNA in isolation. First, the sequencer maps millions of fragments to a **reference human genome**. We know exactly what percentage of the "total DNA soup" should come from each chromosome if the count is normal.

- **Chromosome 1** is huge, so it might normally make up 8% of all fragments.
- **Chromosome 21** is very small, so it might normally only make up **1.3%** of all fragments.

2. The "Expected" vs. "Observed" Calculation

The computer is looking for a **statistically significant deviation** from that 1.3% baseline.

- **Scenario A (Normal):** The machine sees exactly 1.3% of fragments coming from Chapter 21. This matches the "Normal 46" profile.
- **Scenario B (Trisomy 21):** If the baby has three copies of Chromosome 21, the amount of "21-confetti" in the blood will increase. Instead of 1.3%, the machine might detect **1.4% or 1.5%**.

While 0.2% sounds tiny, when you are counting **millions** of fragments, that 0.2% jump is a massive red flag.

3. The Z-Score: The "How Sure Are We?" Number

To be sure this isn't just "new shedding" (random noise), labs use a statistical value called a **Z-score**.

- **Z-score < 3:** The "extra" fragments are likely just random variation or "noise." This is a **Negative** result.
- **Z-score > 3:** The amount of Chapter 21 confetti is so much higher than the average that it is statistically impossible for it to be a mistake. This is a **Positive** (High Risk) result.

4. Normalizing against the "Internal Control"

The machine "checks its own work" by comparing Chapter 21 to other chromosomes that are almost never affected, like Chromosome 1.

If *all* chromosomes show "new shedding" (more fragments), the total volume goes up, but the **ratios** stay the same—that's a normal baby. If **only** Chapter 21 fragments increase while Chapter 1 stays the same, that is a Trisomy.

Summary:

"NIPT doesn't count 1, 2, 3... 46. It measures the **density** of fragments. If the 'cloud' of DNA for Chromosome 21 is thicker than the 'clouds' for other chromosomes, we have found a Trisomy."

Clinical Tip

In patients with **metabolic syndrome or high BMI**, the mother's own "normal" DNA shedding is higher. This can "dilute" the baby's extra Chapter 21 fragments, making that 0.2% jump harder to see. This is why a **high fetal fraction** is so important for a confident Z-score.

The "Elegant" Answer: The Principle of Relative Representation

If you were to explain this "elegantly" to a class, you could use the **"Orchestra"** analogy:

Imagine the baby's DNA is an orchestra where every chromosome is an instrument.

- In a **Normal** pregnancy, the "Violin section" (Chromosome 21) plays at a specific, predictable volume relative to the "Cello section" (Chromosome 1).
- The lab isn't looking for a single new violinist to walk onto the stage (that's too hard to see). Instead, they are measuring the **total volume** of the violin section.
- If the violins are **50% louder** than they should be compared to the rest of the orchestra, the conductor knows there must be an extra player (a third chromosome), even if they can't see the person's face.

Summary of the Mechanism

1. **Massive Parallel Sequencing:** We scan millions of fragments to get a high-resolution "volume" reading.
2. **Normalization:** We compare the volume of Chapter 21 to "stable" chapters (like Chromosome 1).
3. **The Statistical Threshold:** We only call it a Trisomy if the "volume" exceeds a strict mathematical limit (the **Z-score**), ensuring we aren't just hearing an echo or "new shredding."

In other words, the statement "if fragment 21 gets denser, and other segment categories like # 1 are not proportionally higher, then the fragment 21 is real and extra".

Spot on. That is the exact logic of the **Normalized Chromosome Ratio**.

In clinical bioinformatics, we call this Massively Parallel Sequencing (MPS). The above statement highlights the core "internal control" of the test: **proportionality.**

The "Steady State" Rule

If a mother is slightly dehydrated or has a high rate of placental cell turnover (more "shedding"), the **total amount** of DNA in the "soup" goes up. However, in a healthy pregnancy, it goes up **proportionally** across all categories:

- **Segment #1:** Increases by 10%
- **Segment #2:** Increases by 10%
- **Segment #21:** Increases by 10%

The **ratio** stays the same. The computer sees this and says: *"There is more confetti today, but the 'recipe' for the baby hasn't changed. Result: Normal."*

The "Red Flag" (Trisomy)

A problem is detected only when there is a **disproportionate** spike. If the density of Segment #21 increases by 50% while Segment #1 remains at its baseline, the "recipe" has changed.

This method allows the test to work regardless of how much total DNA is in the blood. Whether the mother has a lot of "baby confetti" (high fetal fraction) or a little, the machine is always looking at the **relative density** of one chromosome compared to the others.

Key Takeaway: It isn't the *amount* of #21 that matters—it's the **concentration** of #21 relative to the "anchors" (like Segment #1).

A Final Note:

Sometimes, a mother might have a tiny, non-cancerous growth or a specific health condition that sheds extra DNA from a *different* chromosome. If Segment #18 gets denser but the baby is fine, the computer might get confused. This is why NIPT is a **screening** tool—it is an incredibly accurate "counting" machine, but it can occasionally be "tricked" by the mother's own unique biology.

REFERENCES:

1. Mulavagili Vijayasimha, Borderline Z-scores in NIPT: from ambiguity to actionable reporting. Archives of Gynecology and Obstetrics (2025) 312:2351–2353. https://link.springer.com/article/10.1007/s00404-025-08232-y

2. Ziqi Zhao, Yangsai Zhou. Optimization and anomaly judgment of NIPT detection based on multivariate statistical model and machine learning. December 2025. Journal of Computing and Electronic Information Management 19(3):38-41. DOI:10.54097/aqmjsx77

3. Emily Gaudet, MS, CGC; et al. https://www.ajog.org/article/S0002-9378(24)01194-3/pdf#:~:text=zPTR%20vs%20PPV:%20a%20clinical,following%20a%20positive%20NIPS%20result.