One-Page Summary
Subject: Statistical concerns – *Kiani et al, Nature (2025)*, $CD8^+$ T cell stemness precedes post-intervention control of HIV viremia
DOI: 10.1038/s41586-025-09932-w
Overview: I am requesting your independent evaluation of statistical issues in the above Nature paper. The manuscript's central conclusions rely on analyses that violate core principles of independence, unit-of-analysis validity, and unbiased sampling.

Below is a concise outline of the concerns:

1. Pseudo-replication as the primary statistical error

Across multiple figure panels (e.g., Fig. 2c, 2h-i, 4c-f), the authors treat **epitope-specific measurements from the same individual** as **independent biological replicates**.

However:

- Multiple epitope responses originate from **the same person on the same day**
- These responses are **not independent observations**
- Standard tests used (Wilcoxon, unpaired and paired t-tests, Spearman correlations) assume independence
- The paper explicitly states on lines 720-721:*"All replicate measurements reflect distinct biological samples or epitope-specific responses"*. This sentence confirms the unit-of-analysis error.

When data are correctly aggregated at the **participant level**, all reported significant differences in Fig. 2c and 2h-i disappear (p-values shift from <0.01 to ~0.06–0.3). This single issue invalidates the major claims.

2. Differential weighting of individuals

Because the number of epitope responses varies significantly per participant, individuals with a high number of detected responses have a disproportionate influence on the results.

Example:

- Some participants contribute **3-5 responses**
- Others contribute **1-2**

This unequal weighting:

- biases between-group comparisons
- violates assumptions of equal contribution
- inflates the apparent sample size

Under correct weighting, effect sizes contract sharply.

3. Selection bias in Fig. 2h-j ("stemness" phenotype analyses)

The comparison of "stem-like" phenotype frequencies is based on a **non-representative selection of responses**:

- In PICs, 8/9 tetramer+ responses included are *proliferative* (%CD8 CFSE low > 2.5%)
- In PINCs, 6/7 included responses are *non-proliferative* (%CD8 CFSE low < 1%)
- Yet Fig. 2f shows that non-proliferative responses exist in PICs, and proliferative responses exist in PINCs

This indicates:

- non-random selection of data points
- biased enrichment of proliferative responses in PICs
- biased enrichment of non-proliferative responses in PINCs

Given the established correlation between stemness and proliferation, the between-group differences (Fig. 2h) are artifactual.

4. Lack of sensitivity analyses

The paper asserts that results remain significant after excluding a single outlier (PID-314), but:

- no corresponding p-values are reported
- no sensitivity graphs are shown
- no mixed-effects or participant-level re-analysis is provided

Preliminary re-analysis reveals that removing this participant eliminates the apparent significance when the data are aggregated.

5. Combined impact

Together, pseudo-replication, differential weighting, and selection bias:

- systematically inflate significance
- artificially strengthen between-group differences
- undermine all central claims

These appear to be structural rather than minor analytical errors.

Request

If you find the statistical concerns substantiated, your independent assessment – whether via PubPeer or other forum – would provide clarity for the scientific community.