

From: 
Subject: Potential falsification of data in one of your publications
Date: December 10, 2021 at 3:57 PM
To: nature@nature.com

JK

Dear Editors,

I wanted to bring to your attention that data may have been falsified by the authors in one of your publications, [Systems vaccinology of the BNT162b2 mRNA vaccine in humans](#), specifically the [GSE171964](#) dataset (CITE-seq).

Removal of samples

Samples were taken from 6 participants on days 0, 1, 2, 7, 21, 22, 28 and 42 post-vaccination, however only on day 1 post-vaccination, data for 3 of the 6 participants is missing, ostensibly the data from two female participants and one male participant. I did not find any justification or explanation for this exclusion of samples in the article text, but the article is very long and I may have missed it.

Only samples from female participants were removed

Certain genes which are Y-linked and therefore only present in men are useful biomarkers for determining the sex of the study participants, for example ATP-dependent RNA helicase DDX3Y (DDX3Y) and histone demethylase UTY (UTY). Inspecting the raw counts for DDX3Y and UTY reveals that participants 2051 and 2052 appear to have “swapped” genders only on day 1 post-vaccination.

DDX3Y/ Sample ID	Sample ID Sex	Day 0	Day 1	Day 2	Day 7	Day 21	Day 22	Day 28	Day 42
2047	Male	2628	1716	1730	3136	2445	3370	2762	
2049	Male	3284	1226	4038	3257	3867	3831	3606	
2051	Female	0	1096	0	0	2	0	1	
2052	Male	3270	0	2893	3764	2803	2911	3874	
2053	Female	0	0	0	0	0	0	0	
2055	Female	0	0	0	0	0	0	1	
UTY/Sample ID									
2047	Male	1318	1002	920	1645	1369	1573	1258	
2049	Male	1656	542	1908	1551	1886	1651	1506	
2051	Female	0	423	0	0	0	0	0	
2052	Male	1903	0	1554	1989	1729	1641	2213	
2053	Female	0	0	0	0	0	0	0	
2055	Female	0	0	0	0	0	0	0	

Of course, this is biologically impossible but not necessarily evidence of malicious intent. The fact that this error occurred on the same day for which other samples are conspicuously missing to begin with, and additionally the fact that the error only occurred in the data for the sole female participant, make it exceedingly unlikely that this anomaly is the result of an honest mistake. In any case, the fact remains that all data available for day 1 post-vaccination is from men, and no data from women is available for day 1.

I'm not sure what the standard procedure for situations like these is, but I hope that this concern is at least addressed by the authors. You can download and verify the raw data from the GEO link above, and for your convenience I also attached the raw counts for DDX3Y and participants 2051/2052 as I calculated them.

Sincerely,

Johannes





DDX3Y_raw_cou
nt_2051.txt



DDX3Y_raw_cou
nt_2052.txt