

Twitter Thread by [Ian Rennie](#)



[Ian Rennie](#)

[@theangelremiel](#)



[@KatyMontgomerie](#) [@EIPolloDiabloX](#) [@AbigailShrier](#) [@chadfelixg](#) Okay, so you haven't read the study in the British Journal of Sports Medicine, you've read a Guardian story about it.

I have now read the study, and I teach critical appraisal as part of my job, so let's dive into it a bit.

[@br0oklynzo0](#) [@KatyMontgomerie](#) [@EIPolloDiabloX](#) [@AbigailShrier](#) [@chadfelixg](#) First things first, the sample size in this study is tiny. 46 people. Plus, the point of comparison was "all women and men under the age of 30 in the Air Force between 2004 and 2014" rather than a longitudinal group.

[@br0oklynzo0](#) [@KatyMontgomerie](#) [@EIPolloDiabloX](#) [@AbigailShrier](#) [@chadfelixg](#) So no matching for age, background, or physical condition. The result of these is that you're going to get a lot of statistical artefacts.

[@br0oklynzo0](#) [@KatyMontgomerie](#) [@EIPolloDiabloX](#) [@AbigailShrier](#) [@chadfelixg](#) Second of all, there were three points of comparison here and you're ignoring two of them. The researchers measured for push-ups and sit-ups as well, both of which showed physical advantage removed entirely by hormones. Which means you're cherry-picking.

[@br0oklynzo0](#) [@KatyMontgomerie](#) [@EIPolloDiabloX](#) [@AbigailShrier](#) [@chadfelixg](#) So what do you get when you have a tiny sample size and cherry-picked results? That's right, you get noise in the data. Take a look at the results graph from this paper you haven't read.

[@br0oklynzo0](#) [@KatyMontgomerie](#) [@EIPolloDiabloX](#) [@AbigailShrier](#) [@chadfelixg](#) You'll notice in the times for the 1.5 minute mile that the mean times go pretransition 708, 0-1 years 758, 1-2 years 791, 2-2.5 years 765. So the data seems to think that trans people get... faster after two years? What?

[@br0oklynzo0](#) [@KatyMontgomerie](#) [@EIPolloDiabloX](#) [@AbigailShrier](#) [@chadfelixg](#) And why does the 2-2.5 years category exist? Why not 2-3? Or for that matter why include over 2 years at all, especially when it contains this odd anomalous reading?

@br0oklynzo0 @KatyMontgomerie @ElPolloDiabloX @AbigailShrier @chadfelixg What's also notable is that there is clearly an outlier in the 2.5 year readings. In the range for the 1.5 minute mile, the fastest time goes from 753s in the 1-2 years category to 685s in the 2-2.5 year category. That's an anomaly, and one the paper never mentions.

@br0oklynzo0 @KatyMontgomerie @ElPolloDiabloX @AbigailShrier @chadfelixg Had I been conducting this study, that anomaly would have been enough for me to question my data. if I'm being charitable I'd say that perhaps they didn't notice. If I'm not being charitable I'd say perhaps the anomaly was the point.

@br0oklynzo0 @KatyMontgomerie @ElPolloDiabloX @AbigailShrier @chadfelixg In summary, nothing here is strong enough to base very much on.

@br0oklynzo0 @KatyMontgomerie @ElPolloDiabloX @AbigailShrier @chadfelixg incidentally, MLA citation:

Roberts, Timothy A., Joshua Smalley, and Dale Ahrendt. "Effect of gender affirming hormones on athletic performance in transwomen and transmen: implications for sporting organisations and legislators." *British Journal of Sports Medicine* (2020).