



LIFE SCIENCE AND BIOMEDICINE  
REPLICATION

# Comparing the creatocrit of human milk before and after long-term freezing

Masako Fujita<sup>1,2,\*</sup>  and Eleanor Brindle<sup>3</sup>

<sup>1</sup>Department of Anthropology, Michigan State University, East Lansing, Michigan, USA, <sup>2</sup>Biomarker Laboratory for Anthropological Research, Michigan State University, East Lansing, Michigan, USA, and <sup>3</sup>Maternal, Newborn and Child Health & Nutrition, PATH, Seattle, WA, USA

\*Corresponding author. Email: [masakof@msu.edu](mailto:masakof@msu.edu)

(Received 19 April 2022; Accepted 24 May 2022)

## Abstract

**Objectives:** The creatocrit is a simple technique for estimating the lipid content of milk, widely adopted for clinical and research purposes. We evaluated the effect of long-term cryogenic storage on the creatocrit for human milk.

**Methods:** Frozen and thawed milk specimens ( $n = 18$ ) were subjected to the creatocrit technique. The specimens were reanalyzed after long-term cryogenic storage (10 years at  $<70^{\circ}\text{C}$ ). The correlation between pre- and post-storage values was tested, and their differences were analyzed using the Bland–Altman plot.

**Results:** The pre- and post-storage values were highly correlated ( $r = 0.960$ ,  $p < .0001$ ). The Bland–Altman plot revealed a positive association between their differences and means (Pitman’s test  $r = 0.743$ ,  $p < .001$ ), suggesting the presence of nonconstant bias across the creatocrit range. Long-term storage of human milk may introduce subtle bias to the creatocrit in replicating pre-storage values. Further research should evaluate whether this bias is statistically correctable.

**Key words:** creatocrit; cryogenic storage; fat content; human milk; lipid content

## Introduction

The creatocrit is a simple technique for estimating the lipid and energy content of human milk, based on the separation of cream and aqueous fractions of milk in a capillary tube using a microhematocrit centrifuge—the creatocrit value is determined as the percentage of the cream column length out of the total milk column (Lucas et al., 1978). The technique has been widely adopted for clinical practices and milk research, owing to its simplicity and minimal needs for equipment, supplies, or chemicals. It is an economical and safe alternative to other techniques such as the Gerber standard method (Collares et al., 1997).

Numerous studies have evaluated the validity and reliability of the creatocrit technique. These include the effect of storage temperatures (Silprasert et al., 1987), freezing and thawing (de Groot et al., 1998; Silprasert et al., 1987; Vazquez-Roman et al., 2014; Wang et al., 1999), varying thawing protocols (Thatrimontrichai et al., 2012), temperatures during centrifugation, and the inter- and intra-rater errors (Tie et al., 2021).

These studies clarify the validity of the creatocrit technique applied to frozen and thawed milk specimens for estimating milk lipid content. The creatocrit values tend to be lower in frozen and thawed specimens than the values from fresh specimens (de Groot et al., 1998; Silprasert et al., 1987;

Vazquez-Roman et al., 2014; Wang et al., 1999). Mathematical formulae can be applied to account for this short-term freeze–thaw effect for estimating the lipid and energy content through the creatatocrit of previously frozen milk (Wang et al., 1999). This is particularly important for research in remote sites or population-level research in which immediate laboratory analysis of the specimens is unfeasible.

We evaluated the effect of longer-term storage on the creatatocrit values through a comparison of the creatatocrit values from human milk specimens undergone 10 years in cryogenic storage to the original creatatocrit values from the same specimens prior to long-term storage (but have undergone one short-term freeze–thaw cycle). Human milk specimens in cryogenic archival storage are abundant in universities, milk banks, and other clinical and research facilities. These specimens hold potential for the retrospective determination of their nutritional content for research on human milk variation across time and place. It is currently unclear whether the magnitude of the long-term cryogenic storage effect would be equivalent to that of the shorter-term freeze–thaw effect. Characterizing the agreement or difference between the creatatocrit before and after long-term archival storage would help fill this knowledge gap and facilitate research opportunities using archival milk specimens.

## Methods

### *The creatatocrit*

Frozen and thawed milk specimens ( $n = 18$ ) originally collected among volunteers in rural Kenya were subjected to the creatatocrit in 2008 and again in 2018, separated by storage at cryogenic temperatures ( $<70^{\circ}\text{C}$ ). The specimen collection and the creatatocrit measurement in 2008 were approved and overseen by the institutional review boards of the University of Washington and the Kenya Medical Research Institute. The procedure for specimen collection can be found in Fujita et al. (2011). The creatatocrit re-testing of de-identified milk specimens in 2018 required no further approvals.

The published methodology for the creatatocrit technique (Lucas et al., 1978) explained more in detail by Wang et al. (1999) was used. In short, the homogenized milk was drawn into the capillary tube, and the dry end of the tube was sealed with sealing clay. The process was repeated two more times to produce three sealed tubes per specimen. The tubes were centrifuged to separate the cream/lipid from the aqueous fraction. The length of the cream column and the length of the total milk column of the centrifuged tubes were measured to the nearest 0.5 mm. The creatatocrit value was reported as the average percentage of the cream column length out of the total milk column length.

The 2008 (pre-storage) and 2018 (post-storage) creatatocrit analyses were performed, using the same protocol and equivalent equipment by the same trained observer (MF) albeit in two laboratories. The former analysis was performed at the Biodemography Laboratory of the University of Washington, and the latter at the Biomarker Laboratory for Anthropological Research of Michigan State University where specimens have been housed since 2011 under the Material Transfer Agreement between the two universities.

### *Data analysis*

Descriptive statistics were applied to the creatatocrit values and their natural logarithmic values. The correlation between these pre- and post-storage logarithmic values was tested. The scatterplot was constructed to visualize the relationship between pre- and post-storage logarithmic values.

A Bland–Altman plot was created to allow visual evaluation of the agreement between the pre- and post-storage creatatocrit values and identify any systematic bias. The distribution of the percentage difference between post-storage and pre-storage creatatocrit (calculated as post-storage minus pre-storage multiplied by 100, then divided by their mean; equivalently  $100 \times \ln(\text{post-storage}) - 100 \times \ln(\text{pre-storage})$ ; Cole & Altman, 2017) was plotted against their means. This plot displayed three reference lines (with 95% CIs [confidence intervals]): the observed mean percentage difference, the upper limit of agreement ( $\bar{x} + 1.96 \text{ SD}$ ), and the lower limit of agreement ( $\bar{x} - 1.96 \text{ SD}$ ; Bland & Altman, 1999;

Giavarina, 2015) along with the theoretical line for no difference (% difference = 0). The calculation of 95% CIs was based on an exact method using the noncentral *t*-distribution, proposed by Carkeet (2015), Chatfield (2022), and Shieh (2018). If post-storage values perfectly reproduce pre-storage values, then the percentage difference would be zero. The difference between the mean percentage difference and this theoretical zero was interpreted as the overall bias between pre- and post-storage values. The upper and lower limits of agreement were interpreted as the largest extent of biases.

The directional patterning of data points within the plane of the Bland–Altman plot was interpreted as the presence of bias that is nonconstant. The Pitman’s test of equality of variance evaluated the correlation between the paired logarithmic differences and the creatatocrit means; a significant correlation was interpreted as evidence for the presence of some linear patterning of the bias, owing to differing variances in the two sets of data (Bartlett & Frost, 2008).

A second Bland–Altman plot was constructed, replacing the three reference lines with the regression-based ones for the bias of percentage difference and limits of agreement (Bland & Altman, 1999; Chatfield, 2022; Giavarina, 2015; Ludbrook, 2010) for an additional characterization of the patterning of the data. Here, the bias equation was based on the regression of percentage difference in the natural logarithmic mean of post- and pre-storage values. The equations for the limits of agreement were respectively the bias equation  $\pm 1.96$  SD, where SD was obtained by a second regression of the residuals (the absolute values taken and multiplied by  $\sqrt{\pi/2}$ ) from the first regression (Bland & Altman, 1999; Chatfield, 2022).

We used Stata version 15 with the “BAPLOT” and “blandaltman” packages (Chatfield, 2022) for statistical computation.

## Results/conclusions

The pre- and post-storage creatatocrit percentages and their summary statistics are found in Table 1. These values were comparable to previously reported creatatocrit values, although the right-skewed distribution might have underrepresented the upper end of the physiological range. The mean (SD) total fat content calculated from the equation,  $0.54 \times \text{creatatocrit} + 0.39$  (Wang et al., 1999), was 3.0 (1.3) g/dl, ranging from 1.3 to 5.5 g/dl, slightly lower than the aggregated range 2.5–6.0 g/dl reported from several human populations (Miller et al., 2013).

Using the logarithmic values, the pre-storage and post-storage data were highly correlated ( $r = 0.960$ ,  $p < .0001$ ); the strong linear positive relationship is depicted in the scatterplot (Figure 1).

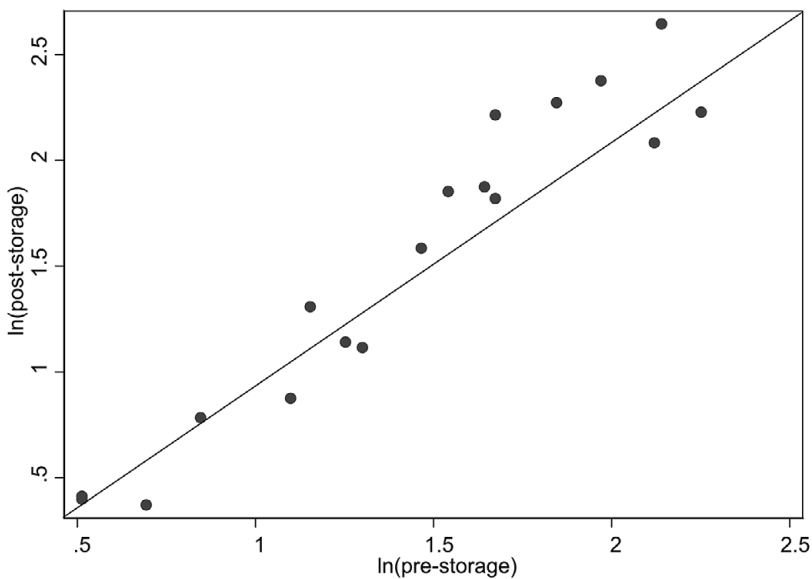
Figure 2 displays the Bland–Altman analysis for the percentage differences between post- and pre-storage creatatocrit percentages against their means, along with the reference lines. The mean percentage difference (dashed line) was 9.1% (95% CI  $-3.9$ , 22.0). The CI of the mean percentage difference includes the value zero (that is no difference), meaning that the overall bias was minor.

However, the upper limit of agreement (60%; 95% CI 43.7, 89.7) and the lower limit of agreement ( $-42\%$ ; 95% CI  $-71.5$ ,  $-25.5$ ) were wide apart from the mean, revealing that the extent of under- and overestimations were large. Furthermore, there was an upward slope in the percentage difference across the mean, suggesting that the bias was not constant. Namely, the bias changed from negative to positive as the mean values increased; the post-storage values underestimated (that is fell short of) the pre-storage values by 42% in the lower range of creatatocrit, whereas they overestimated (that is exceeded) by 60% in the higher range. This positive association was significant when evaluated by the Pitman’s test ( $r = 0.738$ ,  $p < .001$ ) between logarithmic creatatocrit differences versus creatatocrit means. This suggests that the variance of the pre-storage values differed from that of the post-storage values, contributing to nonconstant bias (Bartlett & Frost, 2008). In the second Bland–Altman plot, the regression-based bias and limits of agreement had prominent positive slopes, corroborating the finding of the positive association (Figure 3).

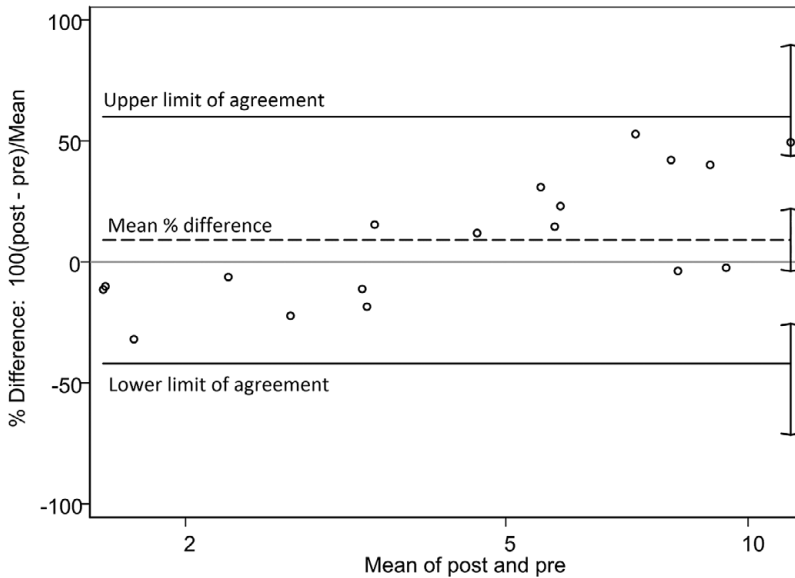
The widely separated limits of agreement and corresponding large extent of over- and underestimations can be attributed to the small sample size (Giavarina, 2015). Nonetheless, the strong positive

**Table 1.** Creatatocrit data and summary statistics

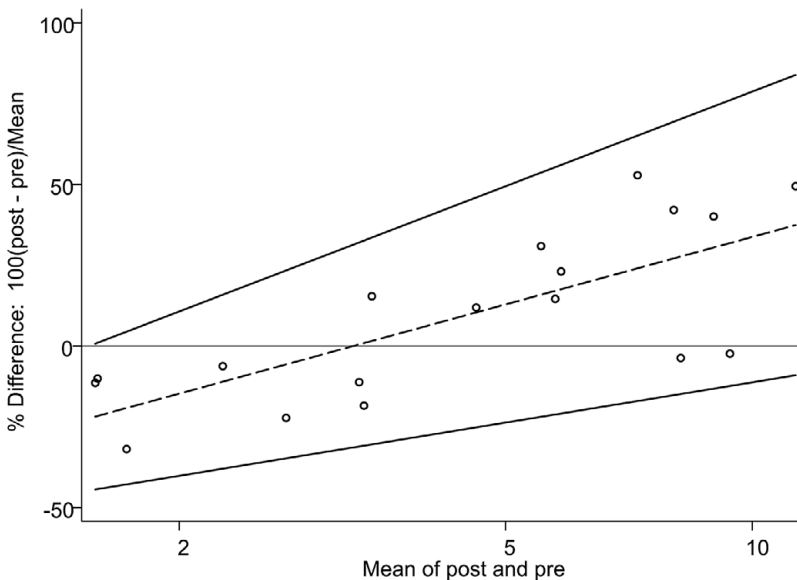
Specimen #	Pre-storage creatatocrit %	Post-storage creatatocrit %
1	3.00	2.40
2	5.33	6.17
3	8.50	14.09
4	8.33	8.03
5	4.67	6.38
6	2.33	2.19
7	9.50	9.28
8	3.67	3.05
9	1.67	1.51
10	2.00	1.45
11	1.67	1.49
12	4.33	4.88
13	3.17	3.70
14	5.17	6.52
15	7.17	10.77
16	3.50	3.13
17	5.33	9.16
18	6.33	9.71
Mean (SD)	4.76 (2.42)	5.77 (3.75)
Min, Max	1.67, 9.50	1.45, 14.09



**Figure 1.** Scatterplot of post-storage versus pre-storage creatatocrit using natural logarithm-transformed values. The diagonal reference line represents the perfect correlation.



**Figure 2.** Bland–Altman plot for percent difference. The percentage difference between post- and pre-storage creatocrit values versus the mean of post- and pre-storage creatocrit values. Mean % difference, 9.1% (95% CI –3.9, 22.0); upper limit 60% (95% CI 43.7, 89.7); lower limit –42% (95% CI –71.5, –25.5).



**Figure 3.** Bland–Altman plot with the regression-based percentage difference and limits of agreement. The same plot as Figure 2 displaying the regression-based bias of percentage difference ( $y = -35.6 + 30.1 \times \ln(M)$ ; dashed line) and upper and lower limits of agreement ( $y = -18.6 + 42.3 \times \ln(M)$  and  $y = -52.6 + 18.0 \times \ln(M)$ , respectively), where  $M$  stands for the mean of post- and pre-storage creatocrit values.

association between differences and means is apparent and suggests the presence of a systematic, concentration-dependent bias.

This study found that the creatocrit of human milk undergone 10 years of cryogenic storage can provide values in high correlation with creatocrit values obtained prior to storage; however, our results

suggest a systematic bias that may vary with milk lipid concentration. This bias may result in a subtle but systematic underestimation in the lower creatamocrit range and overestimation in the higher creatamocrit range. Further research with a larger sample size is needed to scrutinize the extent of this bias and determine if a mathematical formula could correct it.

**Acknowledgment.** Emma Bignall and Nerli Paredes Ruvalcaba assisted with creatamocrit analysis at the Biomarker Laboratory for Anthropological Research.

**Data availability statement.** The data used in this study are provided in Table 1.

**Funding statement.** Specimens and data used for this work were supported by the National Science Foundation (M.F.; grant numbers BCS 0622358 and BCS 1638167) and the Wenner-Gren Foundation (M.F.; grant numbers 7460 and 9278). The Eunice Kennedy Shriver National Institute of Child Health and Human Development research infrastructure grant (P2C HD042828) supported E.B. during a portion of the research.

**Conflicts of interest.** The authors declare none.

**Authorship contributions.** M.F. conceived and designed the study. E.B. and M.F. procured materials. M.F. performed laboratory and statistical analyses. M.F. drafted the article. E.B. and M.F. edited the article for intellectual content.

## References

- Bartlett, J. W., & Frost, C. (2008). Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. *Ultrasound in Obstetrics & Gynecology*, **31**, 466–475. <https://doi.org/10.1002/uog.5256>
- Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, **8**, 135–160. <https://doi.org/10.1177/096228029900800204>
- Carkeet, A. (2015). Exact parametric confidence intervals for Bland–Altman limits of agreement. *Optometry and Vision Science*, **92**, E71–E80. <https://doi.org/10.1097/OPX.0000000000000513>
- Chatfield, M. (2022). BLANDALTMAN’: Module to create Bland–Altman plots featuring differences or %differences or ratios, with options to add a variety of lines and intervals. (Stata package). <http://fmwww.bc.edu/RePEc/bocode/b>
- Cole, T. J., & Altman, D. G. (2017). Statistics Notes: Percentage differences, symmetry, and natural logarithms. *British Medical Journal*, **358**, j3683. <https://doi.org/10.1136/bmj.j3683>
- Collares, F. P., Goncalves, C. V., & Ferreira, J. S. (1997). Creatamocrit as a rapid method to estimate the contents of total milk lipids. *Food Chemistry*, **60**, 465–467. [https://doi.org/10.1016/S0308-8146\(97\)00014-9](https://doi.org/10.1016/S0308-8146(97)00014-9)
- de Groot, S. D. W., Kramer, L., Bulux, J., Solomons, N. W., Green, M. H., & West, C. E. (1998). Effect of freezing on the creatamocrit values of breast milk samples. *FASEB Journal*, **12**, A538–A538.
- Fujita, M., Shell-Duncan, B., Ndemwa, P., Brindle, E., Lo, Y.-J., Kombe, Y., & O’Connor, K. A. (2011). Vitamin A dynamics in breastmilk and liver stores: A life history perspective. *American Journal of Human Biology*, **23**, 664–673. <https://doi.org/10.1002/ajhb.21195>
- Giavarina, D. (2015). Understanding Bland Altman analysis. *Biochemia Medica*, **25**, 141–151. <https://doi.org/10.11613/bm.2015.015>
- Lucas, A., Gibbs, J. A. H., Lyster, R. L. J., & Baum, J. D. (1978). Creatamocrit—Simple clinical technique for estimating fat concentration and energy value of human milk. *British Medical Journal*, **1**, 1018–1020. <https://doi.org/10.1136/bmj.1.6119.1018>
- Ludbrook, J. (2010). Confidence in Altman–Bland plots: A critical review of the method of differences. *Clinical and Experimental Pharmacology and Physiology*, **37**, 143–149. <https://doi.org/10.1111/j.1440-1681.2009.05288.x>
- Miller, E. M., Aiello, M. O., Fujita, M., Hinde, K., Milligan, L., & Quinn, E. A. (2013). Field and laboratory methods in human milk research [Research Support, U.S. Gov’t, Non-P.H.S. Review]. *American Journal of Human Biology*, **25**, 1–11. <https://doi.org/10.1002/ajhb.22334>
- Shieh, G. (2018). The appropriateness of Bland–Altman’s approximate confidence intervals for limits of agreement. *BMC Medical Research Methodology*, **18**, 45. <https://doi.org/10.1186/S12874-018-0505-y>
- Silprasert, A., Dejsarai, W., Keawvichit, R., & Amatayakul, K. (1987). Effect of storage on the creatamocrit and total energy content of human-milk. *Human Nutrition—Clinical Nutrition*, **41C**, 31–36.
- Thatrimontrichai, A., Janjindamai, W., & Puwanant, M. (2012). Fat loss in thawed breast milk: Comparison between refrigerator and warm water. *Indian Pediatrics*, **49**, 877–880. <https://doi.org/10.1007/s13312-012-0219-x>
- Tie, W. J., Kent, J. C., Lai, C. T., Rea, A., Hepworth, A. R., Murray, K., & Geddes, D. T. (2021). Reproducibility of the creatamocrit technique for the measurement of fat content in human milk. *Food Chemistry*, **356**, 129708. <https://doi.org/10.1016/j.foodchem.2021.129708>

- Vazquez-Roman, S., Alonso-Diaz, C., Garcia-Lara, N. R., Escuder-Vieco, D., & Pallas-Alonso, C. R. (2014). Effect of freezing on the "crematocrit" measurement of the lipid content of human donor milk. *Anales De Pediatría*, **81**, 185–188. <https://doi.org/10.1016/j.anpedi.2013.09.001>
- Wang, C. D., Chu, P. S., Mellen, B. G., & Shenai, J. P. (1999). Creamatocrit and the nutrient composition of human milk. *Journal of Perinatology*, **19**, 343–346. <https://doi.org/10.1038/sj.jp.7200204>

---

**Cite this article:** Fujita M, Brindle E (2022). Comparing the creatocrit of human milk before and after long-term freezing. *Experimental Results*, 3, e11, 1–9. <https://doi.org/10.1017/exp.2022.9>

# Peer Reviews


**Reviewing editor:** Dr. Michael Nevels

University of St Andrews, Biomolecular Sciences Building, Fife, United Kingdom of Great Britain and Northern Ireland, KY16 9ST

Minor revisions requested.

doi:10.1017/exp.2022.9.pr1

## Review 1: Comparing the creatinocrit of human milk before and after long-term freezing

**Reviewer:** Prof. Oke Gerke PhD 

Odense University Hospital, Dept. of Nuclear Medicine, Odense, Denmark, 5000

Date of review: 29 April 2022

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

**Conflict of interest statement.** Reviewer declares none.

### Comment

Comments to the Author: In light of the data collection among volunteers in rural Kenya in 2008 and again in 2018, this agreement study is limited to 18 specimen (minimum 30-50 subjects would be clearly preferable in Bland-Altman analysis).

Abstract & Results: please indicate both p-values as ' $p < 0.0001$ ' instead of '0.000'. These are small, but positive probabilities, not 0.

Methods, Data Analysis:

(1) please indicate how 95% CIs for the Bland-Altman Limits of Agreement were derived. From Fig. 2 it is apparent that the confidence limits are not symmetric around the point estimates (those originally proposed by Bland & Altman that is the approximate Wald-type ones). Later, exact non-symmetric confidence limits for the Limits of Agreement have been proposed by Carkeet (doi: 10.1097/OPX.0000000000000513). Were these employed?

(2) I agree with the indicated directional pattern. As there are only  $n=18$  observations, I propose to stick to your analyses shown, but supplement the finding of the upwards trend in the data by another figure that shows Bland-Altman Limits of Agreement based on a regression approach (see Fig. 8 in doi:10.1177/096228029900800204, Fig. 2 in doi:10.1111/j.1440-1681.2009.05288.x, Fig. 3 in doi:10.11613/BM.2015.015).

(3) Other STATA package is concord.

Fig. 1: 1970's way of displaying (dis)agreement; Bland & Altman argued that disagreement may be hard to detect visually here (see their seminal paper in Lancet 1986;1(8476):307-10, Fig. 1 vs. Fig. 2). Could also apply here. However, as supplementary plot, you may as well stick to it, but please enhance the scatter plot by adding a 45 degrees line for interpretability, thanks.



---

## Score Card

### Presentation



Is the article written in clear and proper English? (30%)

5/5

Is the data presented in the most useful manner? (40%)

3/5

Does the paper cite relevant and related articles appropriately? (30%)

4/5

### Context



Does the title suitably represent the article? (25%)

5/5

Does the abstract correctly embody the content of the article? (25%)

5/5

Does the introduction give appropriate context? (25%)

5/5

Is the objective of the experiment clearly defined? (25%)

5/5

### Analysis



Does the discussion adequately interpret the results presented? (40%)

5/5

Is the conclusion consistent with the results and discussion? (40%)

5/5

Are the limitations of the experiment as well as the contributions of the experiment clearly outlined? (20%)

5/5