New Data Demonstrates that Relationship of Salivary to Serum Progesterone in Bolivian Women is Comparable to that in other Populations

Jonathan Thornburg[1], Diva Bellido[2], Rose Stewart[3], Hilde Spielvogel[2], Lourdes Echalar[2], Jennifer Burch[4], and Virginia J. Vitzthum[4]

[1] Department of Astronomy and Center for Spacetime Symmetries, Indiana University, Bloomington, Indiana, USA
[2] Instituto Boliviano de Biología de Altura, La Paz, Bolivia
[3] Center for the Integrative Study of Animal Behavior, Indiana University, Bloomington, Indiana, USA
[4] Anthropology Department and Kinsey Institute, Indiana University, Bloomington, Indiana, USA

Abstract

Salivary progesterone (P) levels are widely used as a proxy for serum P in field studies. Chatterton *et al.* (2006) reported that the ratio of salivary to serum P in concurrently collected samples is substantially lower in Bolivian women ($n = 26$) than in Chicago women ($n = 20$) and concluded that interpopulational variation in salivary P does not indicate parallel variation in serum P. To further investigate the relationship between salivary P and serum P in Bolivians, here we present new data from $n = 33$ sets of concurrent saliva and serum samples collected from a new sample of $n = 31$ Bolivian women. Contrary to Chatterton *et al.*'s findings, our data show an excellent correlation ($r = 0.76$) between salivary P and serum P. Our data have typical uptake fractions (ratio of salivary P to serum P between 1.5% and 8.8%; geometric mean 2.5%), consistent with other published values for non-Bolivian populations. We argue that Chatterton *et al.*'s results are most parsimoniously explained by some combination of inappropriate statistical analyses, sample contamination, and/or inaccurate assays. Our data and analyses confirm that salivary P is a reliable proxy for serum P in Bolivian (and other) women.

Introduction

Salivary progesterone (P_{saliva}) is widely used as a proxy for serum P (P_{serum}, a biomarker of ovarian functioning). Unlike serum, saliva samples are easily collected in a field setting and (with a suitable preservative) can be stored for long periods at ambient temperatures.

To investigate the relationship between P_{serum} and P_{saliva}, and how this might differ in Bolivian women from those in an industrialized country, Chatterton *et al.* (2006) obtained concurrent mid-luteal saliva and serum samples from $n = 26$ Bolivian and $n = 20$ Chicago women, and assayed these samples for P. Chatterton *et al.* found that the mean Bolivian P_{saliva} was 48% of mean Chicago P_{saliva}, while mean Bolivian P_{serum} was 197% of mean Chicago P_{serum}. Chatterton *et al.* concluded that "the differences between concentrations of salivary P in these population groups are not indicative of the serum concentrations of these steroids [*sic*]", discussed various possible explanations, and posed several questions for further research "that must be answered before we can interpret differences in salivary steroid hormone levels between population groups".

If true, Chatterton *et al.*'s findings would call into question many studies using P_{saliva}. To further investigate the utility of P_{saliva} as a proxy for P_{serum}, we collected a new data set of concurrent serum/saliva samples from Bolivian women. Here we describe preliminary analyses of this new data.

For our analyses, we define the "uptake fraction" f of a serum/saliva sample pair as the ratio of salivary to serum P, i.e., $f = P_{\text{saliva}}/P_{\text{serum}}$.

Chatterton *et al.*'s Data and Analyses

Chatterton *et al.*'s analyses is based on ratios of group means of the Bolivian and Chicago samples. We (Thornburg *et al.* (2008)) have previously outlined several serious flaws in their analyses. Notably:

- Because Chatterton *et al.* inappropriately used arithmetic (rather than geometric) means to summarize each study sample's P_{serum} and P_{saliva}, the ratios of Chatterton *et al.*'s group means differ dramatically (by more than a factor of 2) from the means of the individual samples' uptake fractions.
- Group means—regardless of how they are calculated—give little information about the full *bivariate* statistical distribution of P_{serum} and P_{saliva}, or about how these are related to (correlated with) each other.
- Chatterton *et al.* failed to remove extreme outliners present in both their Chicago and Bolivia study samples. These outliers severely biased their analyses.

Figure 1 shows a scatterplot of Chatterton *et al.*'s data. In such scatterplots, an ideal data set (salivary levels highly correlated to serum levels) would have all the points falling on a narrow diagonal band of nearly-constant uptake fraction, with biologically reasonable values and ranges for each hormone and for the uptake fraction.

It's clear that Chatterton *et al.*'s data are far from this ideal, and have extreme outliers in both the Chicago and Bolivia study samples. Even neglecting the outliers, the range of variation within each study sample is still quite large, which renders the group means relatively uninformative about the actual correlations between P_{serum} and P_{saliva}.

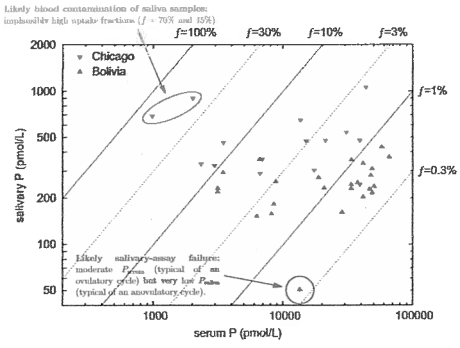

Figure 1: Scatterplot of Chatterton *et al.*'s P_{saliva} *vs.* P_{serum} data. The diagonals show lines of constant uptake fraction f. Note that the scales are logarithmic on both axes. The circled points are extreme outliers (likely due to contamination and assay failure). We excluded these outliers from our analyses of Chatterton *et al.*'s data.

Materials and Methods

We recruited study participants by announcements and word of mouth from members of IBBA and the Medical School of the Universidad Mayor de San Andrés in La Paz, Bolivia. All participants were between 20 and 40 years old, of normal weight for height, regularly cycling for at least the past 3 months, non-pregnant and non-lactating, and not using hormonal contraceptives or any other hormonal medication during the previous 4 months. All protocols had IRB approval.

Here we focus on the $n = 33$ pairs of concurrent serum/saliva samples collected from $n = 31$ women during their mid-luteal phases. Collection materials were made of either laboratory-quality glass or pure polypropylene. Blood samples were centrifuged, and serum and saliva aliquots were frozen at $-20°$ C shortly after collection and maintained frozen until they were assayed by standard enzyme-immunoassay methods using kits from Salimetrics, LLC. (State College, U.S.A.).

Data and Analyses

Figure 2 shows a scatterplot of our data. We excluded the circled outlier points from all analyses. Unlike Chatterton *et al.*'s data, our data does indeed fall in a relatively narrow diagonal band of roughly-constant uptake fraction, with biologically reasonable values and ranges for each hormone and for the uptake fraction.

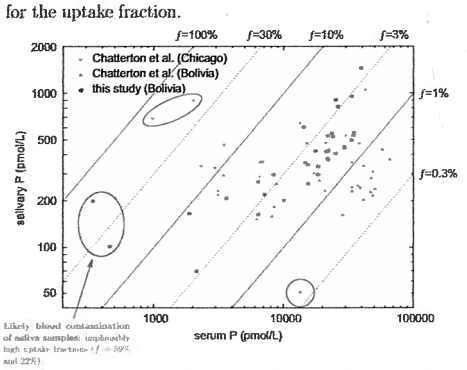

Figure 2: Scatterplot of our P_{saliva} versus P_{serum} data (green points) with Chatterton *et al.*'s data included (blue and red points) for comparison. The diagonals show lines of constant uptake fraction f. Note that the scales are logarithmic on both axes. We treated the circled points as outliners and excluded them from our analyses.

Figure 3 shows the distribution of uptake fractions in our data. The uptake fractions range from 1.5% to 8.8% (geometric mean 2.5%), and are consistent with other reported values in the literature for industrialized populations. For example, Boever *et al.* (1986) reported a luteal-phase uptake fraction of 1.8%, Choe *et al.* (1983) reported 3.7%, and Tallon *et al.* (1984) reported 1.0%.

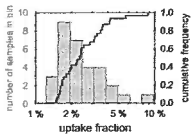

Figure 3: Histogram (left scale) and cumulative distribution (right scale) of uptake fraction in our data. Notice that the horizontal (uptake-fraction) scale is logarithmic.

Figure 4 shows the correlation of P_{saliva} *vs.* P_{serum} in the various study samples. Chatterton *et al.*'s data show low correlations (< 0.50) for *both* the Chicago and Bolivia samples. In contrast, our data show high correlations (0.76 for P_{saliva} *vs.* P_{serum} and 0.85 for $\log P_{\text{saliva}}$ *vs.* $\log P_{\text{serum}}$), in excellent agreement to those published for other industrialized-country populations. This agreement argues strongly against Chatterton *et al.*'s suggestion that steroid biochemistry is somehow unusual in Bolivians.

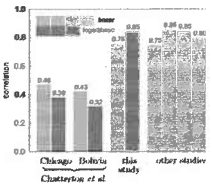

Figure 4: Pearson correlation c efficient r of P_{saliva} *vs.* P_{serum} a of $\log P_{\text{saliva}}$ *vs.* $\log P_{\text{serum}}$ for ca study sample (with outliers r moved as shown in figures 1 and 2 Lu *et al.* (1997) was the origin publication of the radioimmunoa say protocol used by Chatterton *al.*. Salimetrics refers to the val reported by the manufacturer f the assay kits used in this study.

Conclusions

- Chatterton *et al.* found that the ratio of mean salivary P t mean serum P differed substantially between their Chica and Bolivia samples, and suggested unusual biochemistry i the Bolivians as an explanation.
- We have identified a number of serious flaws in Chatterton statistical analyses:
  - Inappropriate use of arithmetic means instead of geome ric means when taking ratios (an error that changes som results by more than a factor of 2).
  - Failure to remove outliers from the data; these outlie strongly bias their analyses.
  - Failure to consider the full *bivariate* statistical distribu tion of (and correlation between) serum and salivary P.
- **We find a combination of sample contaminatio and/or inaccurate assays to be a more parsimo nious explanation for Chatterton *et al.*'s result** Indeed, if saliva/serum P ratios ("uptake fractions") a tually varied as much as Chatterton *et al.*'s data sugges then it's difficult to see how it is that many other publishe studies have obtained reasonable salivary P profiles.
- **We have obtained new data** to investigate the rel tionship between serum and salivary P in Bolivians.
- **Our results show a strong correlation ($r = 0.76$ between serum and salivary P**, in excellent agre ment with other published studies for industrialized-countr populations. We find uptake fractions ranging from 1.5 to 8.8% (geometric mean 2.5%), again in excellent agre ment with other published studies.
- We conclude that **salivary P is a reliable proxy fo serum P** in Bolivian (and other) populations.

Acknowledgements

We thank Esperanza Caceres and Gertrudis Nina for invalu able assistance with Bolivian data collection, and the stud participants for their time. Supported by Instituto Bolivian de Biología de Altura, La Paz, Bolivia; Anthropology D partment, Indiana University; Kinsey Institute for Researc in Sex, Gender, and Reproduction; Center for the Integrativ Study of Animal Behavior, Indiana University.

References

Boever, J. D., F. Kolen, and D. Vandekerckhove. 1986. Clinical Chemistry 32(5). 76

Chatterton, R. T., Jr., E. T. Mateo, D. Lu, and F. J.-H. Ling. 2006. Fertility and Steril 86(3). 723.

Choe, J. K., F. S. Khan-Dawood, and M. Y. Dawood. 1983. American Journal Obstetrics and Gynecology 147(5). 557.

Lu, Y.-C., R. T. Chatterton, Jr., K. M. Vogelsong, and L. K. May. 1997. Journal Immunoassay and Immunochemistry 18(2). 149.

Tallon, D. F., J. P. Gosling, P. M. Buckley, M. M. Dooley, W. F. Cleere, E. M. O'Dwye and P. F. Fottrell. 1984. Clinical Chemistry 30(9). 1507.

Thornburg, J., H. Spielvogel, and V. J. Vitzthum. 2008. American Journal of Physic Anthropology 135. 206