Supplemental Material

Oral Health/Hygiene

Levels of salivary markers of inflammation are altered by oral health/hygiene (Jaedicke et al., 2016; Slavish et al., 2015). For example, salivary IL-1β and salivary TNF-α—though not salivary IL-6 or salivary IL-8—are elevated in periodontal disease and decrease as a result of treatment for periodontal disease, leading to the conclusion that these salivary cytokines can be considered biomarkers for periodontal disease (Sexton et al., 2011; Rathnayake et al., 2012; for review, see Jaedicke et al., 2016). Similarly, dental caries (i.e., tooth decay) has been associated with elevated salivary IL-6 (Gornowicz et al., 2012; Menon et al., 2016) as well as elevated salivary TNF-α and salivary IL-8 (Gornowicz et al., 2012). These findings entail that care should be taken to ensure oral health/hygiene levels are controlled for and do not influence results in studies examining salivary cytokines. We quantified oral health/hygiene problems in the current study using the following interview:

Y N 1. Do your teeth or gums ache? (*If yes:* "Was this happening in the past week? What kind of pain?" *Code current pain suggestive of injury or infection, not cold/heat sensitivity*)

Y N 2. Do you have any sores or cuts in your mouth? (*If yes:* "Do you have them now?" *Rate only if current*)

Y N 3. Are your gums red and swollen?

Y N 4. Do your gums bleed easily? (*If yes:* "When does this happen?" *Meant to be ordinary activities like brushing teeth or flossing*)

Y N 5. Do you often see some blood after brushing your teeth?

Y N 6. Has a dentist ever told you that you have a gum disease, like gingivitis? (*If yes:* "Do you have it now?" *Rate only if current*)

Y N 7. Do you have cavities that have not been filled? (*If yes:* "Did a dentist tell you that you need a filling or said that it needs to be watched?" *Rate only if filling is needed now*)

Y N 8. Do you have any loose teeth in your mouth?

Y N 9. Have you lost a tooth in the past month?

Y N 10. Does your mouth often feel dry? (*If yes:* "Did you have this feeling almost every day this month?")

Y N 11. Have you been drooling more than usual? (*If yes:* "In the past week was it much more than usual?" *Rate only if that's the case*)

Y N 12. Are you taking any medication for your mouth? What is it called? Why are you taking it?

"Yes" responses from this interview were summed to indicate oral health problems. This sum was then used as a covariate in relevant analyses. As reported in the main text, controlling for oral health/hygiene produced virtually identical results.

Data Analyses

Short-term reliability at each assessment was calculated as the test-retest correlation between Sample 1 and Sample 2. We also calculated test-retest correlations between corresponding samples (i.e., Sample 1 with Sample 1, Sample 2 with Sample 2, or Composite with Composite) from the first and second assessment. Long-term stability was calculated by correcting the long-term test-retest correlations for measurement error using the correction for attenuation, i.e.,

$$r_{x'y'} = \frac{r_{xy}}{\sqrt{r_{xx} r_{yy}}}$$

Where r_{xy} is the correlation between 2-sample composites between baseline and follow-up, r_{xx} and r_{yy} equal the reliabilities of variables *x* and *y*, respectively, which were the Spearman-Browne reliabilities of the short-term test-retest correlations between Sample 1 and Sample 2 at each assessment, i.e., $\frac{2r}{1+r}$

Whenever a mean correlation is presented, correlations that went into its calculation were Fisher *z*-transformed, then averaged, and this average was then back-transformed from *z* to *r* for reporting. Missing values (e.g., biomarker level below the lower detection limit) were handled by listwise deletion.[1] Finally, although the provided detection rates include all samples, we removed outliers greater than ± 3 standard deviations from the mean for all analyses to prevent these more extreme scores from unduly influencing the results.

Intraclass Correlations

From the perspective of classical test theory, intraclass correlation coefficients (ICCs) are optimal measures of reliability when the absolute number is actionable (e.g., systolic blood pressure > 130 mm Hg indicates need for treatment or driving with blood alcohol level > .08 is illegal), so it is important to ensure that instruments agree on specific values with the gold standard and each other. However, if relative standing in the sample (i.e., rank-order stability) is of interest, then test-retest Pearson's r is the most relevant statistic (Schmidt et al., 2003). In other words, ICCs reflect both rank-order stability (Pearson's r) and mean-level changes in the sample between the two time points. In our 18-month retest, mean-level changes include effect of maturation (i.e., all participants were older at retest), batch effects for reagents, changes to study staff who collected samples, and similar confounds. These effects are not relevant to most research applications of salivary inflammation markers, as actionable cutoffs tied to absolute values of concentration exist only for CRP, and even those cutoffs were developed only for

[1] We chose to delete missing values listwise and to remove outlying cases rather than winsorizing because most studies measuring salivary cytokines that we know of follow these approaches. However, when missing values were imputed with the lower limit of detection and extreme values were winsorized rather than removed, the same findings emerged in all analyses. Although there are too many tests to report all of these results using both approaches, we note that the composite of the two samples still dramatically outperformed the individual samples even under this alternate approach (i.e., imputing missing values with the lower limit of detection and winsorizing extreme values); the composite showed moderate 18 month test-retest correlations, mean $r = .24$, $p = .009$, whereas the individual samples did not, $ps \geq .07$.

values in blood and in relation to physical health. For these reasons, we consider test-retest *r*s the most relevant statistic and focused on it in the paper. Nonetheless, we acknowledge there may be readers curious about the ICCs, so we conducted these analyses. The results are as follows:

	r	*d*	ICC
Short-Term Baseline			
TNF-α	0.65	0.13	0.64
IL-1β	0.70	0.36	0.62
IL-6	0.72	0.01	0.72
IL-8	0.51	0.51	0.40
IL-18	0.55	0.08	0.54
CRP	0.81	0.02	0.80
MCP-1	0.64	0.07	0.64
Short-Term Follow-up			
TNF-α	0.54	-0.04	0.55
IL-1β	0.61	0.17	0.60
IL-6	0.66	-0.08	0.63
IL-8	0.39	0.26	0.37
IL-18	0.59	0.16	0.59
CRP	0.72	0.03	0.70
MCP-1	0.49	0.18	0.46
Long-term			
TNF-α	0.22	1.82	-0.35
IL-1β	0.30	0.77	0.10
IL-6	0.10	0.68	-0.03
IL-8	0.27	0.57	0.16
IL-18	0.37	0.62	0.21
CRP	0.31	0.49	0.19
MCP	0.31	-0.02	0.28

The above Table shows test-retest *r*, Cohen's *d* for mean-level change, and ICC (two-way mixed effects, absolute agreement). ICC equaled *r* when there was no mean-level difference, but was greatly reduced and even made negative when mean-level difference were high, such as at the long-term retest; the pattern inherent to definition to ICC as described above. For reasons we

described in the paragraph above, we believe test-retest correlations are more informative than ICCs to those interested in utilizing salivary markers of inflammation in studies.

Validity of Salivary Cytokines

Although our results speak to the reliability of salivary markers of inflammation, important considerations regarding the validity of salivary markers of inflammation as indicators of immune system function remain. Salivary markers of inflammation are indices of immune system function (e.g., interleukins are proteins through which immune system cells in the mouth communicate), though they may reflect local immune system function rather than systemic, central, or circulating immune system function (e.g., Riis et al., 2015). Consistent with this concern, correlations between salivary and circulating cytokines are modest (Riis et al., 2014). Nonetheless, similar to circulating inflammatory biomarkers, sleep deprivation (El-Sheikh et al., 2007), acute stress (Slavish et al., 2015; Szabo et al., 2019), early adversity (Szabo et al., 2019; Tyrka et al., 2015), and obesity (Goodson et al., 2014) all increase or are associated with increased levels of salivary inflammatory biomarkers. Moreover, salivary inflammatory biomarkers are higher in individuals suffering from various diseases—such as cancer (Koizumi et al., 2018), rheumatoid arthritis (Mirrielees et al., 2010) and asthma (Little et al., 2014)—and these levels decrease with systemic anti-TNFα treatment (Mirrielees et al., 2010). Additionally, the importance of local immune system activity to overall health should not be understated, as localized inflammatory activity can nonetheless contribute to glucocorticoid resistance (Rook et al., 2000), which has implications for systemic immune system activity and therefore health. Similarly, localized immune system activity is an important indicator of local diseases, such as periodontal disease (Zhang et al., 2016). Therefore, salivary markers of inflammation provide a health-relevant index of at least local immune system function and possibly a window to the broader immune system milieu.

References

El-Sheikh, M., Buckhalt, J.A., Granger, D.A., Erath, S.A., Acebo, C., 2007. The association between children's sleep disruption and salivary interleukin-6. J. Sleep Res. 16, 188–197. https://doi.org/10.1111/j.1365-2869.2007.00593.x

Goodson, J.M., Kantarci, A., Hartman, M.L., Denis, G. V., Stephens, D., Hasturk, H., Yaskell, T., Vargas, J., Wang, X., Cugini, M., Barake, R., Alsmadi, O., Al-Mutawa, S., Ariga, J., Soparkar, P., Behbehani, J., Behbehani, K., Welty, F., 2014. Metabolic disease risk in children by salivary biomarker analysis. PLOS ONE 9. https://doi.org/10.1371/journal.pone.0098799

Gornowicz, A., Bielawska, A., Bielawski, K., Grabowska, S.Z., Wójcicka, A., Zalewska, M., & Maciorkowska, E., 2012. Pro-inflammatory cytokines in saliva of adolescents with dental caries disease. Ann. Agric. Environ. Med. 19, 711–716.

Jaedicke, K.M., Preshaw, P.M. and Taylor, J.J., 2016. Salivary cytokines as biomarkers of periodontal diseases. Periodontol. 2000 70, 164–183. https://doi.org/10.1111/prd.12117

Koizumi, T., Shetty, V., Yamaguchi, M., 2018. Salivary cytokine panel indicative of non-small cell lung cancer. J. Int. Med. Res. 46, 3570–3582. https://doi.org/10.1177/0300060518775563

Little, F.F., Delgado, D.M., Wexler, P.J., Oppenheim, F.G., Mitchell, P., Feldman, J.A., Walt, D.R., Peng, R.D., Matsui, E.C., 2014. Salivary inflammatory mediator profiling and correlation to clinical disease markers in asthma. PLOS ONE 9. https://doi.org/10.1371/journal.pone.0084449

Menon, M.M., Balagopal, R.V., Sajitha, K., Parvathy, K., Sangeetha, G.B., Arun, X.M., Sureshkumar, J., 2016. Evaluation of salivary interleukin-6 in children with early childhood caries after treatment. Contemp. Clin. Dent., 7, 198–202. https://10.4103/0976-

237X.183059

Mirrielees, J., Crofford, L.J., Lin, Y., Kryscio, R.J., Dawson, D.R., Ebersole, J.L., Miller, C.S., 2010. Rheumatoid arthritis and salivary biomarkers of periodontal disease. J. Clin. Periodontol. 37, 1068–1074. https://doi.org/10.1111/j.1600-051X.2010.01625.x

Rathnayake, N., Åkerman, S., Klinge, B., Lundegren, N., Jansson, H., Tryselius, Y., Sorsa, T. and Gustafsson, A., 2013. Salivary biomarkers of oral health–a cross-sectional study. J. Clin. Periodontol. 40, 140–147. https://doi.org/10.1111/jcpe.12038

Riis, J.L., Granger, D.A., Dipietro, J.A., Bandeen-Roche, K., Johnson, S.B., 2015. Salivary cytokines as a minimally-invasive measure of immune functioning in young children: Correlates of individual differences and sensitivity to laboratory stress. Dev. Psychobiol. 57, 153–167. https://doi.org/10.1002/dev.21271

Riis, J.L., Out, D., Dorn, L.D., Beal, S.J., Denson, L.A., Pabst, S., Jaedicke, K., Granger, D.A., 2014. Salivary cytokines in healthy adolescent girls: Intercorrelations, stability, and associations with serum cytokines, age, and pubertal stage. Dev. Psychobiol. 56, 797–811. https://doi.org/10.1002/dev.21149

Rook, G., Baker, R., Walker, B., Honour, J., Jessop, D., Hernandez-Pando, R., Arriaga, K., Shaw, R., Zumla, A., Lightman, S., 2000. Local regulation of glucocorticoid activity in sites of inflammation: Insights from the study of tuberculosis. Ann. N. Y. Acad. Sci. 917, 913–922.

Schmidt, F.L., Le, H., Ilies, R., 2003. Beyond alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual-differences constructs. Psychol. Methods, *8*, 206–224.

Sexton, W.M., Lin, Y., Kryscio, R.J., Dawson III, D.R., Ebersole, J.L. and Miller, C.S., 2011. Salivary biomarkers of periodontal disease in response to treatment. J. Clin. Periodontol. 38, 434–441. https://doi.org/10.1111/j.1600-051X.2011.01706.x

Slavish, D.C., Graham-Engeland, J.E., Smyth, J.M., Engeland, C.G., 2015. Salivary markers of inflammation in response to acute stress. Brain. Behav. Immun. 44, 253–269. https://doi.org/10.1016/j.bbi.2014.08.008

Szabo, Y.Z., Fernandez-Botran, R., Newton, T.L., 2019. Cumulative trauma, emotion reactivity and salivary cytokine levels following acute stress in healthy women. Anxiety, Stress Coping 32, 82–94. https://doi.org/10.1080/10615806.2018.1524377

Tyrka, A.R., Parade, S.H., Valentine, T.R., Eslinger, N.M., Seifer, R., 2015. Adversity in preschool-aged children: Effects on salivary interleukin-1β. Dev. Psychopathol. 27, 567–576. https://doi.org/10.1017/S0954579415000164

Zhang, C.Z., Cheng, X.Q., Li, J.Y., Zhang, P., Yi, P., Xu, X., Zhou, X.D., 2016. Saliva in the diagnosis of diseases. Int. J. Oral Sci. 8, 133–137. https://doi.org/10.1038/ijos.2016.38