Icahn School of Medicine at Mount Sinai

CHEAR Data Center
Department of Environmental Medicine and Public Health

One Gustave L. Levy Place
Box 1057/1058
New York, NY 10029-6574

Date 2/19/2019
From CHEAR Data Center
Subject CHEAR Project #2017-1762– Study data for inflammatory markers
To Dr. Mahabee-Gittens

The inflammatory markers in saliva data (N=272) analyzed by the Minnesota CHEAR Exposure Assessment Hub for CHEAR project #2017-1762 are now finalized and available on the CHEAR Data Submission and Review Portal (DSRP). Biomarkers in this dataset include C-reactive protein (CRP), Interleukin 8 (IL8), and Interleukin 10 (IL10), as per Lab Analysis Plan. Due to insufficient quantity of sample for analysis, 1 sample was not analyzed for CRP, 7 samples were not analyzed for IL8, and 40 samples were not analyzed for IL10.

Limit of Detection (LOD) values were reported for each biomarker and used to identify the valid values for each sample. Machine-read values were provided for all samples, including those less than LOD. The following approaches can be considered for utilizing biomarker values that are below the LOD:

- Investigators may choose to use a machine value.
- Investigators may substitute a surrogate value for all values <=LOD, which is often LOD/sqrt2, first recommended by Hornung and used by CDC.
- Investigators can consider using a multiple imputation method (see Lubin). In general, a surrogate value is not used in models with continuous variables unless >60% of the observations are >LOD, as described in Lubin.

The PI included 13 duplicate pairs, distributed across the batches. Of these, 12 pairs had sufficient volume for CRP, 11 pairs for IL8, and 10 pairs for IL10.

These laboratory results have been reviewed and approved by the CHEAR Lab Hub to assure that they conform to acceptable quality standards. Summary tables of the study sample data and relevant quality control data are appended at the end of this document.

Signed,

Matthew Mazzella
Statistical Analyst

Susan Teitelbaum, Ph.D.
Principal Investigator, CHEAR Center for Data Science, Mount Sinai

REPORT OF LAB RESULTS

1 - Summary Table of Sample Data:

Geometric mean and geometric mean standard deviation calculated for subjects with detectable levels (NC- not calculated when detection frequency was less than 70%).The LOD was utilized to determine detectable values.

Analyte	Number of Samples Analyzed	Number of Samples >LOD	Percent Detect	LOD	Geometric Mean	Geometric STD	Min	10th Percentile	25th Percentile	Median	75th Percentile	90th Percentile	Max
CRP (pg/mL)	272	272	100%	100	2498.40	2.94	289.6	788.4	1208	2098.65	3649.65	12326.4	53253.8
IL10 (pg/mL)	233	229	98%	0.28	1.19	2.10	<LOD	0.48	0.7	1.08	1.72	2.96	22.68
IL8 (pg/mL)	266	265	100%	0.16	423.32	3.07	<LOD	111.86	183.32	376.32	787.22	2004.42	17146.9

2- Summary Table of QC Data (sample duplicates):

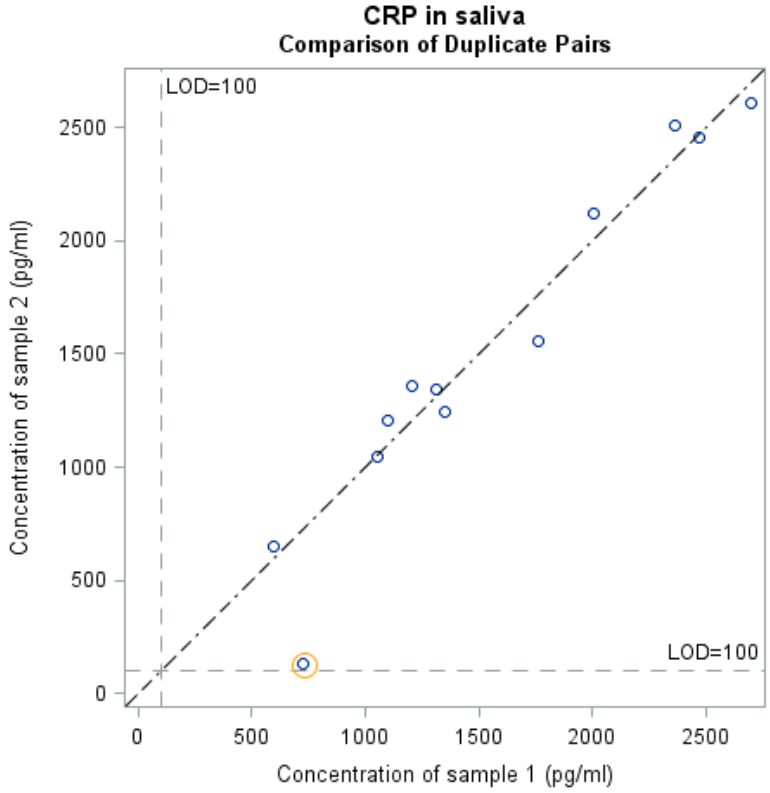
A) ***QC Duplicates Summary Table***- As a measure of precision based on the percent of the difference between duplicate sample concentrations relative to the mean of the duplicate sample concentrations per subject, the Relative Percent Difference (RPD)* was calculated for each blinded duplicate pair when both sample concentrations were above LOD for a specific analyte. With multiple duplicate pairs from different subjects, we produce summary statistics from the distribution of pair-specific RPDs; the range of concentrations will span the distribution of the study samples. The intra-class correlation (ICC)** is a composite measure of reliability and is used to assess both the degree of correlation and agreement between measurements.

Analyte	N total duplicate pairs	N valid duplicate pairs	% valid duplicate pairs	LOD	RPD Median	RPD P75	RPD P90	RPD Max	ICC
CRP (pg/mL)	12	12	100	100	7	10	12	140	0.718
IL10 (pg/mL)	10	10	100	0.28	14	25	42	55	0.978
IL8 (pg/mL)	11	11	100	0.16	6	6	6	50	0.990

**The relative percent difference (RPD) is* $RPD = \frac{|sample\ result - repeat\ result|}{(sample\ result + repeat\ result)/2}$ **100 and has been used by the EPA (New Jersey Department of Environmental Protection 2014) to evaluate duplicate samples. It is defined as a measure of precision that is based on a percent of the difference relative to the mean of the duplicate value per subject. For example, if one sample has a value of 2 and its duplicate pair has a value of 4, the mean is 3 and the RPD is 67%. A duplicate pair of 2000 and 4000 would also have the same RPD. An RPD of zero indicates that the concentrations of the two duplicate samples were equal. RPDs greater than zero indicate a difference between the two sample concentrations. Following the EPA RPD guideline values (New Jersey Department of Environmental Protection 2014), we have provided sample 1 by sample 2 plots (see below if applicable) when the median RPD exceeds 30% and/or the 90th percentile exceeds 50%.*

***The ICC presented is estimated from a one-way random effects model measuring absolute agreement with multiple raters/measurements (McGraw 1996; Shrout 1979). An ICC of 1 indicates that concentrations of the two samples in a duplicate pair, for each pair, were equal. Although there is no standard value for acceptable reliability using the ICC, Koo et al (2016) recommends that ICC values less than 0.5 indicate poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.90 indicate good reliability, and values greater than 0.90 indicate excellent reliability. We have provided sample 1 by sample 2 plots (see below if applicable) when the ICC is less than 0.75.*

B) ***Sample by sample plots of QC Duplicates***- Samples in the duplicate pair were randomly assigned as sample 1 or sample 2. A plot with the concentration of sample 1 on the x-axis and sample 2 on the y-axis was used to visually inspect agreement between measurements. For complete agreement, we would expect all measures to fall along the dotted reference 45 degree line. Plots for analytes with a measure below our criteria, as stated in part A (median RPD exceeds 30%, 90th percentile RPD exceeds 50%, or ICC is less than 0.75), are provided below. Duplicate pairs with an RPD exceeding 50% are circled in orange on the plot. Dashed lines mark the average LOD for this analyte.



References:

1. Hornung, R. W., & Reed, L. D. (1990). Estimation of average concentration in the presence of nondetectable values. Applied occupational and environmental hygiene, 5(1), 46-51.
2. Lubin, J. H., Colt, J. S., Camann, D., Davis, S., Cerhan, J. R., Severson, R. K. Hartge, P. (2004). Epidemiologic evaluation of measurement data in the presence of detection limits. Environmental health perspectives, 112(17), 1691.
3. Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of chiropractic medicine, 15(2), 155-163.
4. New Jersey Department of Environmental Protection (2014). Data Quality Assessment and Data Usability Evaluation Technical Guidance. Version 1.0, 1-132.