Title of Study: Historical trauma and resilience as a biological state and its association with the effects of the traditional Indigenous food chokeberry

PI: Dr. Nicole Redvers, ND, MPH, University of North Dakota School of Medicine & Health Sciences

Date of Document: December 21, 2021

Statistical Design and Power of Study

All statistical analyses will use two-tail tests with the significance level set at 0.05, using version 9.4 of SAS software (SAS v9.4, Carry, NC). Statistical reports will emphasize our desire to characterize the evidence, in addition to reporting significance levels. Prior to final analyses, we will screen the data, consult with the overarching COBRE PI, project mentors and our COBRE admin core biostatistician as needed. Statistical assumptions will be tested, & appropriate data transformations or model adjustments will be used as needed. Our COBRE admin core biostatistician Dr. Cristina Oancea has extensive experience with these methods and is equipped for supporting this study's needs.

Experimental Design:

Fifty participants will be recruited in a rolling sample collection from the Greater Grand Forks area of North Dakota and Minnesota. A respondent driven rolling sampling method will be used for adults between the ages of 18-65 years of age. Ages have been specified as such due to the substantially changing metabolic profiles in children and older adults. We have proposed a first step cross-sectional baseline data collection process followed by a longitudinal intervention design in this exploratory study examining the effects of consuming traditional Indigenous chokeberry on metabolic, epigenetic and mental health markers where subjects will serve as their own control.

Randomization Note: It is considered almost unethical at this time to conduct double blind, placebo control studies with Indigenous subjects. There has been long-standing historical mistrust that Indigenous communities have with research. This is due to often being taken advantage of in the past by researchers, and even having been the subject of experimental studies that would have been considered illegal by today's IRB standards. The idea of 'tricking' somebody with a placebo agent is not a way to build trust with this population, so adjustments to the standard study design are imperative to build trust, research interest and capacity in Indigenous settings. By supporting Indigenous investigators to lead research with their own populations gives further opportunity for community trust building with institutional settings. This is consistent with the NIH's own 2019-2023 Strategic Plan for Tribal Health Research which helps define the research variables and goals for research in Indigenous communities. The project has been designed with the best possible rigor that is realistic in an American Indian participant setting with involvement of the American Indian community including an Elder to ensure involvement and success.

Sample Size, Power and Attrition:

The proposed project aims to test the differences in gene expression, epigenetic (e.g. methylation) and metabolic endpoints (e.g. IL-6, 8-OHdG, lipids) between baseline and post 6 weeks of daily chokeberry consumption. Our sample will include fifty American Indian participants. With a sample of 50 and a conservative effect size as small as 0.4, we have a power of 79.2% with a two-sided paired test and level of significance at 0.05. We consulted the literature on the effects of our proposed chokeberry intervention on the metabolic outcome measures proposed in our study. The median effect size from these studies in the literature (significant findings) is 1.12, which is much higher than 0.4. Therefore, we expect to have a power higher than 79.2%. As to epigenetic endpoints, to the best of our knowledge, there is no study that has been published to

test chokeberry's effects on human IL-6 promoter CpG island methylation changes. Only a few cell culture studies have shown DNA methylation changes using anthocyanin rich fruits with reported effect size>0.4. Thus, we believe that we should have sufficient power to perform the proposed exploratory study with a sample of size 50. As American Indian populations (which are the basis for this study) have unique research needs and history, we consulted the literature to determine the average attrition rate of studies with the American Indian population. We were able to find literature specific to the reservation areas of North Dakota and used this information in our calculations.

Phase 1 Specific Aim:

For this portion of the statistical analysis, an established method for analyzing epigenetic differences at time points will be utilized. It must be noted that due to the small sample size in this exploratory study design the following statistical procedures will be employed. Normality of continuous variables will be assessed. For non-normal data, appropriate transformations (e.g. log transformations) and/or non-parametric tests will be applied for further statistical analysis. Back-transformations will be used for the interpretation of the final results. Correlation analyses will be performed by the Pearson's correlation test (or Spearman correlation for non-normal data). A general linear model with repeated measurements will be used to assess differences between baseline and post 6 weeks of daily chokeberry consumption with confounding variables adjusted. If multiple epigenetic markers are tested, family-wise Type I error rate will be controlled with standard methods (e.g. Bonferroni, False Discovery Rate method). The Student's t-test for independent samples will be performed to assess gender differences. A p-value of <0.05 will be considered as statistically significant for all statistical assessments in this research project. Future research steps could include a more complex mixed model design where interaction and mediation effects can be tested with a larger sample size pending the outcomes noted here.

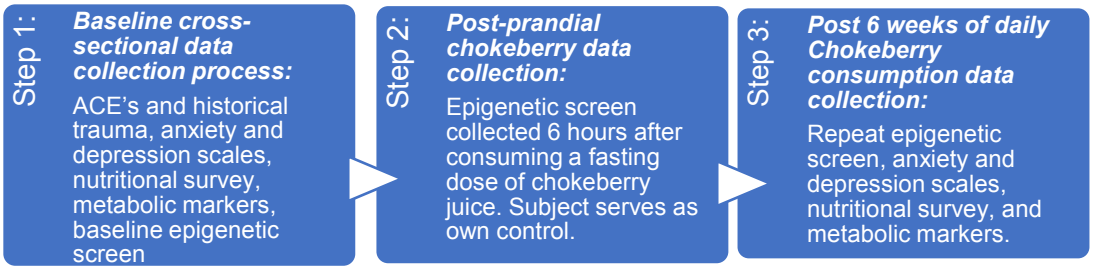

Figure 2. Three step research strategy

Phase 2 Specific Aim:

<u>Baseline and post 6 weeks of chokeberry:</u>
All statistical analyses associated with this cross-sectional portion of the study will be performed using a statistical analysis system software (SAS v9.4, Carry, NC). To investigate the association between the cross-sectional measures (i.e., adverse childhood experiences, historical trauma, resiliency, presence of anxiety and depression, metabolic markers (i.e., 8-OHdG, IL-6, lipids, glucose, CRP, blood pressure)) and baseline epigenetic measures, simple and multivariable linear regression models will be run. Adjustment will be made for the confounders of interest (i.e., smoking status, BMI, physical activity levels, presence of comorbidities, anti-depressant and/or anti-anxiety medication use, supplement use, nutritional intake, age and gender).
For the metabolic end points, the values will be expressed as means ± SD. Statistical tests will be performed using the same commercial software package (SAS v9.4, Carry, NC) as described above. The Kolmogorov–Smirnov test will be applied to determine whether the continuous variables are normally distributed. For within group comparison vs baseline, variables showing normal distributions will be evaluated using the paired *t* test, while variables showing non-normal distributions will be evaluated using the Wilcoxon signed rank test. Our null hypothesis is that chokeberries will not have an effect on gene expression, metabolic markers or mental health status in Great Plains Indian subjects. As noted previously, a p-value of <0.05 will be considered as

statistically significant for all statistical assessments in this research project. Again, future research steps could include a more complex mixed model design where interaction and mediation effects can be tested with a larger sample size pending the outcomes noted here. See *Figure 2*, above for a summary of the study steps.