Work Should Not Hurt You: Reduction of Hazardous Exposures in Small Businesses Through a Community Health Worker Intervention

ClinicalTrials.gov ID NCT03455530

PI: Professor Paloma Beamer (University of Arizona)

Statisticians: Shannon Gutenkunst supervised by Professor Dean Billheimer (University of Arizona)

January 16, 2024

Statistical Analysis Plan

Descriptive statistics will be calculated for shop and participant characteristics. Differences in baseline characteristics between the immediate and delayed intervention groups will be tested using the Welch Two-Sample t-test for continuous variables and Fisher's Exact Test for categorical variables. For all analyses, all tests will be two-sided, with a p-value < 0.05 considered statistically significant.

To analyze the primary outcome TVOC data, we will use linear mixed models to account for the lack of independence from repeated measurements within a shop at an assessment and across assessments. We will use contrasts to test the change in TVOC concentration from pre- to post-intervention between the arms. To account for possible confounding, we will fit both unadjusted and adjusted models. The unadjusted model has log-transformed TVOC concentration as the outcome, with fixed effects for intervention group (immediate, delayed), assessment time (1, 2, 3), and their two-way interactions; random effects will be included for shop and assessment time within shop, with random intercepts to account for correlations of shifts within the same shop and the potentially stronger correlations of shifts within the same shop at the same assessment. Random effects are assumed to be independent. Models for auto and beauty shops will be run separately because of differences in potential covariates for adjusted models.

Covariates for the adjusted models are selected based on their logical relevance and potential impact on exposures. For auto shops, additional covariates will be outside ventilation (i.e., working outside) and average shift apparent temperature. For beauty shops, additional covariates will be baseline air exchange rate and whether the shop offered nail services. The main results will be obtained from the adjusted models, because they provide a more accurate representation by accounting for relevant covariates.

To explore the impact of different model choices, additional adjusted model results will provided in the Supplemental Material. Specifically, covariates that change the treatment effect by more than 10% when added one-at-a-time as a main effect to the unadjusted model for each sector will be included. The following shift-level covariates at each assessment will be considered: apparent air temperature and self-report activity and ventilation data. The following shop-level covariates at each assessment will also be considered: apparent air temperature; outside TVOC concentration; if the building was rented or owned; if an owner or manager wore a monitor during that shop assessment; self-report activity and ventilation data aggregated to the shop-level; if it was a home business; number of people working that day; number of rooms in shop; if the shop uses "green" products; shop volume (calculated from measured shop dimensions); ventilation data (same categories as those for self-report plus wall heater, personal heater, central heater,

and extractor); and for beauty shops only, air exchange rate and whether the beauty shop provided nail services in addition to hair.

To analyze the secondary outcome of hazard scores derived from specific VOC data, we will use similar mixed models as described above to test the change in hazard score from pre- to post-intervention between the arms (although there will not be a random effect for shift within a shop assessment, because shift-level data are not taken for this outcome). Data preparation, cleaning, and statistical analyses will be performed using R version 4.4.1, the tidyverse package for data wrangling, the lme4 package for linear mixed effects models, and other packages listed in the Supplemental Material.