Air purifiers in classrooms for infection control: a pilot study

Chris Rose, Norwegian Institute of Public Health (26 Aug 2024)

Generated using git revision: 1ba5e7f

# Introduction

This document presents methods and results for part 2 of the pilot study on air purifiers in classrooms for infection control (see https://zenodo.org/doi/10.5281/zenodo.12818264).

# Methods

All statistical analyses were performed as prespecified, except as noted, using Stata 18 (StataCorp LLC, College Station, Texas, USA). We had expected both sensors to measure the primary and secondary outcomes as counts, but the DigiRef sensor reported fractional counts. We therefore rounded these to integers. Negative binomial regression was selected over zero-inflated negative binomial and Poisson models for both the intention-to-treat (ITT) analyses of both outcomes on the basis of the Akaike information criterion and was therefore also used for the per-protocol (PP) analyses, which excluded measurements made when the air purifiers were not running as planned. All analyses accounted for the crossover design, sensor type, first-order autocorrelation (lagged outcomes), and clustering within classroom via cluster-robust standard errors. We used maximum pseudolikelihood in the ITT and PP analyses to estimate and hence account for undefined lagged outcomes (e.g., the first measurement of the day). Exposure was defined as the time between outcome measurements. Treatment effect was estimated as rate ratio (RR), where RR<1 disfavors the reference treatment of no air purification.

We used partialing-out lasso Poisson regression (Stata's xpopoisson command) to perform a prespecified exploratory analysis to identify time-varying covariates that may be associated with the primary outcome and assess the sensitivity of the treatment effect estimates to these covariates. The lasso penalty was chosen using 10-fold cross-validation to select among the following covariates: indoor relative humidity; the school's existing ventilation system setting (low or high); weekday; baseline morning PM2.5 (modelled on the log scale); number of students attending class; and mean outdoor temperature. We had planned to include indoor temperature and indicators of undefined lagged PM2.5 but had to omit these due to nonconvergence. We had also planned to include indoor CO2 and sound level but had to omit these because many values were missing.

Finally, we tested for superiority of any versus no air purification, and for noninferiority of portable versus ceiling-mounted air purifiers using a prespecified margin of RR=1.4. We report two-sided 95% confidence intervals and use the conventional p<0.05 significance criterion throughout.

# Results

## Intention-to-treat estimates of treatment effect for the primary and secondary analyses

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | None | Ceiling | Portable | Ceiling | | | Portable | | | Superiority of air purification | Noninferiority (portable vs ceiling) |
|  | Mean | Mean | Mean | RR | [95% | CI] | RR | [95% | CI] |  |  |
| PM2.5 | 1.03 | 0.84 | 0.55 | 0.85 | [0.65 | 1.11] | 0.71 | [0.57 | 0.89] | 0.011 | <0.001 |
| VOC | 192.39 | 197.78 | 202.85 | 1.05 | [0.99 | 1.12] | 1.03 | [0.90 | 1.18] | 0.136 | <0.001 |
| *Sample means are unadjusted and do no account for the crossover design. Rate ratios (RRs) are adjusted for the crossover design, sensor type, first-order autocorrelation, and clustering within classroom. RR < 1 disfavors the reference (no air purification).* | | | | | | | | | | | |

## Exploratory per-protocol estimates of treatment effect for the primary and secondary analyses

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | None | Ceiling | Portable | Ceiling | | | Portable | | | Superiority of air purification | Noninferiority (portable vs ceiling) |
|  | Mean | Mean | Mean | RR | [95% | CI] | RR | [95% | CI] |  |  |
| PM2.5 | 1.03 | 0.84 | 0.55 | 0.86 | [0.65 | 1.14] | 0.70 | [0.56 | 0.87] | 0.005 | <0.001 |
| VOC | 192.39 | 197.78 | 202.85 | 1.07 | [0.99 | 1.15] | 1.03 | [0.90 | 1.18] | 0.118 | <0.001 |
| *Sample means are unadjusted and do no account for the crossover design. Rate ratios (RRs) are adjusted for the crossover design, sensor type, first-order autocorrelation, and clustering within classroom. The estimation sample was restricted to exclude measurements made when air purifiers were not running as planned. RR < 1 disfavors the reference (no air purification).* | | | | | | | | | | | |

## Exploratory time-varying covariate adjusted analysis

The time-varying covariates selected by lasso were indoor relative humidity, weekday, baseline morning PM2.5, and the school's existing ventilation system setting.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | None | Ceiling | Portable | Ceiling | | | Portable | | | Superiority of air purification | Noninferiority (portable vs ceiling) |
|  | Mean | Mean | Mean | RR | [95% | CI] | RR | [95% | CI] |  |  |
| PM2.5 | 1.03 | 0.84 | 0.55 | 0.93 | [0.77 | 1.12] | 0.71 | [0.64 | 0.78] | <0.001 | <0.001 |
| *Sample means are unadjusted and do no account for the crossover design. Rate ratios (RRs) are adjusted for the crossover design, sensor type, first-order autocorrelation, the time-varying covariates selected by the cross-validated lasso, and clustering within classroom. RR < 1 disfavors the reference (no air purification).* | | | | | | | | | | | |

# Discussion

Treatment effect estimates are consistent across the three analyses. Portable air purifiers are estimated to be superior to no air purifiers with respect to PM2.5 but not VOC. For this reason, (any) air purification is estimated to be superior to no air purification with respect to PM2.5. Portable air purifiers are estimated to be noninferior to ceiling-mounted air purifiers (in fact, they appear to be superior to them).

The results of the prespecified exploratory analysis of the effect of time-varying covariates should be interpreted cautiously because we were unable to account for undefined lags and could not include all time-varying covariates. Further, it was not possible to use a negative binomial model for this analysis, as was selected over the Poisson model on the basis of AIC in the main analyses, because Stata does not currently provide xpopoisson-like commands for negative binomial models.