School Air Filtration Study - Initial Findings

Sam Relins

2022-07-15

1 Introduction

The effect a school environment has on the health of its students is of paramount importance, particularly in the wake of the COVID-19 pandemic. Of particular interest is the relationship between classroom air quality and the transmission of airborne infections, and any interventions that might mediate this relationship. This report aims to study the effect of classroom air filters on illness-related absences. We are using data from the Class-Act study schools: 31 primary schools fitted with air quality monitoring devices, and divided into one of 3 intervention arms - 10 schools fitted with HEPA air filters, 10 schools fitted with UVC air purifiers and 11 control schools. Attendance records for each of the schools in the study were provided by Bradford Council's Information Management Team (IMT), and are analysed to identify any relationship between air quality interventions and illness-related absences.

2 The Data

2.1 Data Preparation

The dataset is composed of attendance records for each of the Class-Act schools spanning the period from the beginning of the 2021/2022 academic year to 4th July 2022. Bradford Council IMT provided raw attendance "codes" for each student (anonymised), detailing the one of 30+ attendance types for each of 14 weekly "sessions" (AM/PM, Monday-Sunday) as defined by DfE guidelines. These data were cleaned and aggregated into weekly attendance figures for each school, week and each of the recognised attendance types. The dataset was then labelled with the different arms of the Class-Act study: schools with HEPA/UVC air filters, and those without any filtration devices.

Weeks in which schools were closed were removed from the data, as they contain no information about attendance, illness or otherwise. Furthermore, a minority of codes were "junk" or did not correspond to a recognised attendance type. Pending potential future data quality improvement efforts, school weeks with a 1% or greater proportion of "junk" sessions were removed from the dataset - 2.4% of the aggregated data after removing school closures (Table 1)

	Affected Records	n
Aggregated Dataset	-	1,271
School Closures	160	1,111
>1% "Junk" Records	27	1,084

Table 1: Number of records per school/week in the attendance data, and numbers of records removed in the data preparation process

2.1.1 Attendance Measures

Key data to determining the effect of classroom air quality on attendance are:

1) The number of students present in school - we defined this as any of the following session types: “present AM” / “present PM” / “late arrival before register has closed” / “arrival after registration closed”
2) The number of students absent from school because of illness - we defined this as any of the following session types: “illness” / “illness confirmed covid” / “self-isolating with coronavirus symptoms”*

*Though we would have liked to study covid absences more specifically, ambiguities in the DfE guidelines on covid-related attendance records and the significant strain on resources resulting from the pandemic lead to numerous conflicting approaches to recording covid/non-covid absences.

We defined our outcome measure as the ratio of students absent because of illness as a proportion of the students present in school for each week (percentage).

2.1.2 UVC filters

Unlike the HEPA arm, schools in the UVC arm were not fitted with filters from the beginning of the 2021/2022 academic year, and the filters were not immediately active following installation. For the weeks in which the schools did not have an active UVC filter, they were considered part of the control arm. As a result, records of weeks in which UVC filtration was active are much lower in number than those for HEPA filtration/control arms can be seen in Figure 1 :

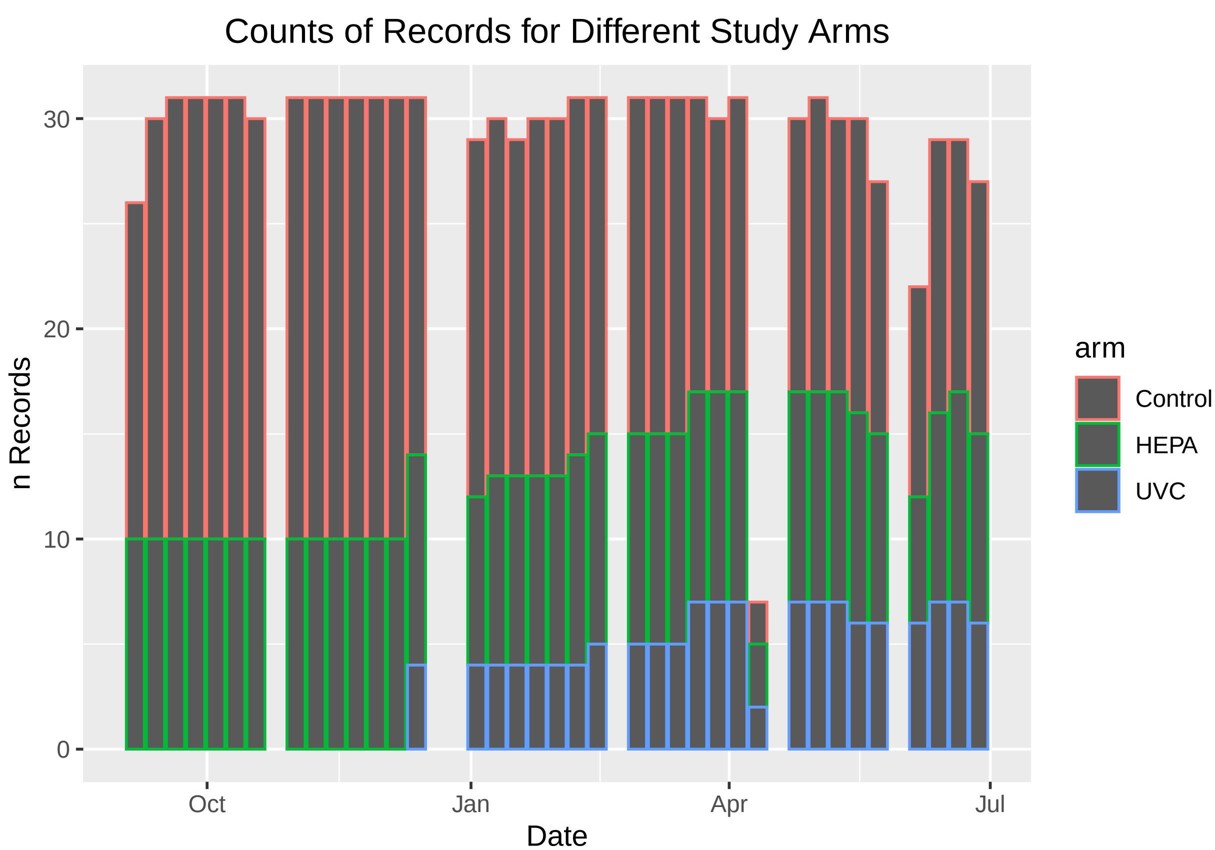

Figure 1: Number of records each week divided by the intervention type. The number of UVC records increases gradually from December 2021 as schools began activating their UVC filtration devices

3 Methods

We performed a basic descriptive analysis of the data, visualising the distribution of illness ratios across time and subdivided by the different air quality interventions - this included LOESS (Locally Estimated

Scatterplot Smoothing) plots to identify the trend in illness rates over time across the different interventions.

We then modeled illness rates using logistic regression, to estimate the effects of the different air quality interventions. We performed a basic regression of illness rates as a function of the different air quality interventions to estimate total effects. We also produced a more descriptive model that accounted for the increases in illness related absences as the new year approached: our initial analysis of the illness rates showed a clear peak in infections in the new year, likely affected both by the wave of Omicron infections over the Christmas period and also seasonal increases in cold/flu viruses (Figure 2). To reflect this peak, we fitted a regression model combining the different interventions with absolute proximity in days to 1st January 2022, in an effort to encode the effect of increased transmission of airborne illnesses as this wave rose and fell.

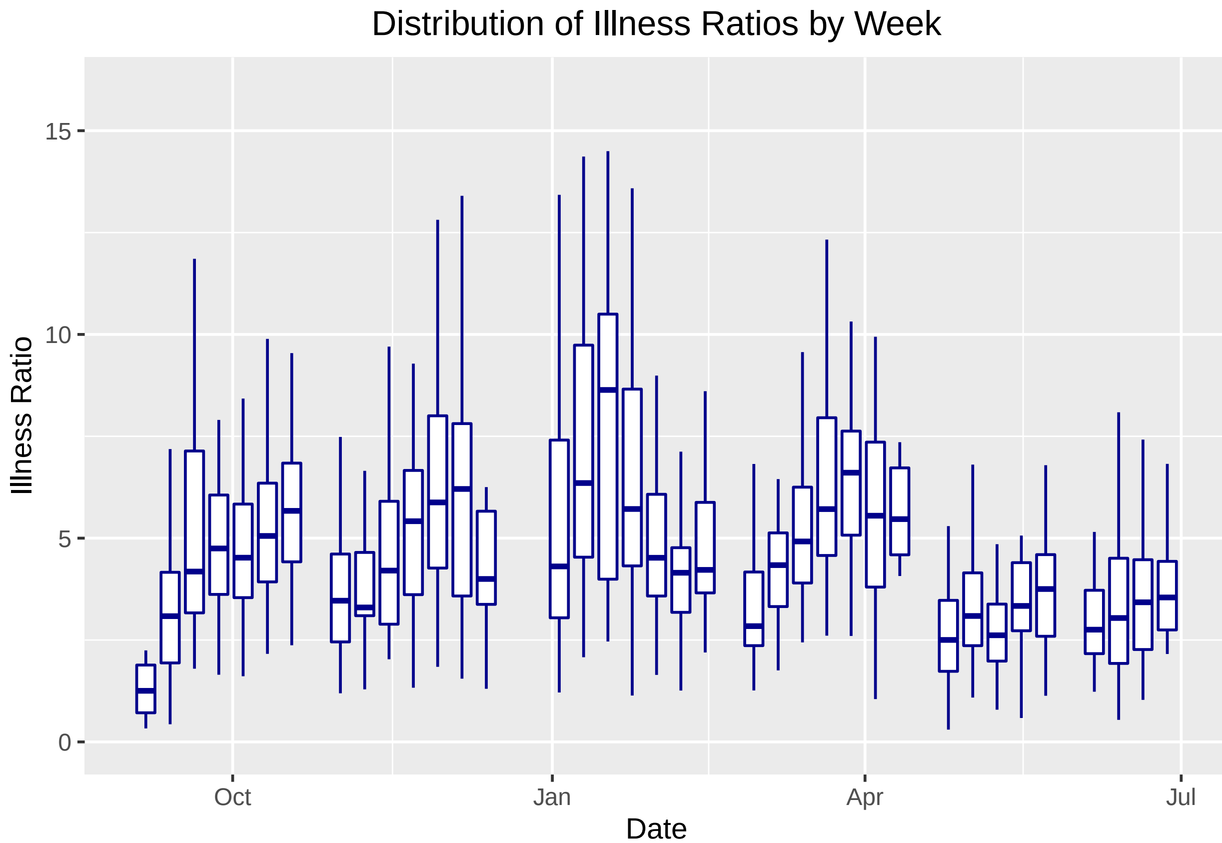

Figure 2: Distribution of illness rates by week for the full dataset. Note the clear increase of infection rates to a peak in early January.

4 Results & Discussion

Our analyses showed a significant reduction in illness rates among the schools using HEPA air filters. Figure 3 shows a clear peak in illness rates in December among the schools without any air filters and those with UVC filters; comparatively, illness rates in the schools using HEPA filters are relatively flat across the same period. This suggests the HEPA filters may be effective at mitigating person-to-person transmission of airbourne pathogens. As the spring and summer periods approach, the illness rates converge and little difference is seen between the different air quality interventions - this may result from greater ambient ventilation from windows/doors as outdoor temperatures rise throughout the year, and/or a reduction in transmission of airborne infections as overall illness in the population decrease.

Similarly the logistic regression results show a statistically significant relationship between the use of HEPA air filtration and a reduction in illness absence rates (Table 2). The multi-variable regression shows the best

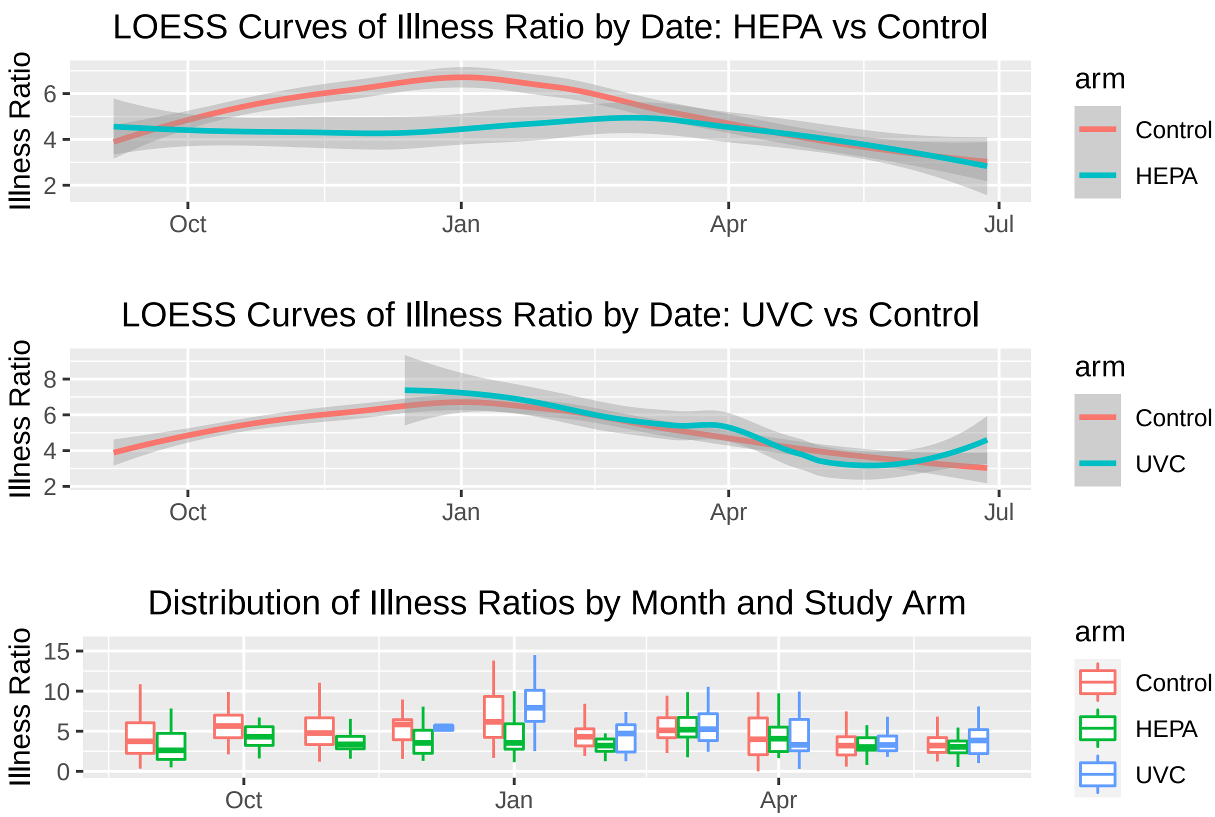

Figure 3: Box plots and LOESS smoothed curves for illness rates broken down by study arm. Note the peak of illness rates in December among the control/UVC arms whilst the HEPA arm remains flat.

fit to the data, and estimates that HEPA filters reduce the illness absence rate by 0.816% (95% CI -1.199, -0.433) - the estimated baseline illness rate is 6.61% (95% CI 6.241, 6.978), so this signifies a 12.3% mean reduction in illness related absences when comparing schools using HEPA filters to those not.

	Dependent variable:
	Illness Rate (%)
Baseline	6.610^{***}
	(0.188)
Distance from 01/01/2022 (Days)	-0.019^{***}
	(0.002)
HEPA	-0.816^{***}
	(0.195)
Observations	1,084
R^2	0.095
Adjusted R^2	0.093
Residual Std. Error	3.006 (df = 1081)
F Statistic	56.574^{***} (df = 2; 1081)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 2: Logistic regression results for illness ratios as a function of HEPA filtration (Yes/No) and absolute number of days from 1st January 2022. UVC filtration was removed as a factor as it was not statistically significant

The UVC intervention does not show a statistically significant relationship with illness rates and so were removed from the final model. This is unsurprising given the similarity in distribution of illness rates between the UVC and control schools seen in Figure 3 - the two are almost indistinguishable. It is difficult to draw conclusions from this result, as very few (if any) of the filters were active during the peak period of illness related absences from October 2021 to February 2022. We would require data covering the winter period 2022/2023, or the results of parralell work using air quality measurements to reach a more definitive conclusion.