Supplementary material of:

Unravelling the role of mandatory use of masks for SARS-CoV-2 control in schools: A quasi-experimental study nested in a population-based cohort in Catalonia (Spain)

Appendix 1: Details of the regression discontinuity design (RDD)

Methods

In order to maintain the assumptions required for a RDD analysis and to increase the number of observations, we use age in months at the end of the study period instead of grade. This limit our analysis only for incidence since it is not possible to calculate SAR or R* by age in months. These two main transmission indicators are aggregate measures as they are calculated at the bubble group level. The cutoff of the RDD analysis was set at 72 months of age, coinciding with the limit between P5 (without FCM mandate) and year 1 of primary education (with mandatory use of FCM). In addition, to focus on two more comparable ages, the same analysis was performed only for children aged between 60 and 83 months.

Results

Figure S1 shows the result of the RDD analysis for SARS-CoV-2 incidence. We found a non-statistically significant absolute difference of -0.0089% (p-value: 0.930).

The same analysis was performed with the two nearest ages around the cutoff (5 and 6 years) and we also found a non-statistically significant difference of -0.0634% (p-value: 0.728) (**Figure S2**).

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Arch Dis Child

Figure S1. RDD analysis for SARS-CoV-2 incidence using age in months. Observed incidences by months of age are split at the cutoff at 72 months (discontinued black line). The shadow areas indicate the 95% CI of the corresponding regression lines (green: 36 to 71 months, blue: 72 to 143 months)
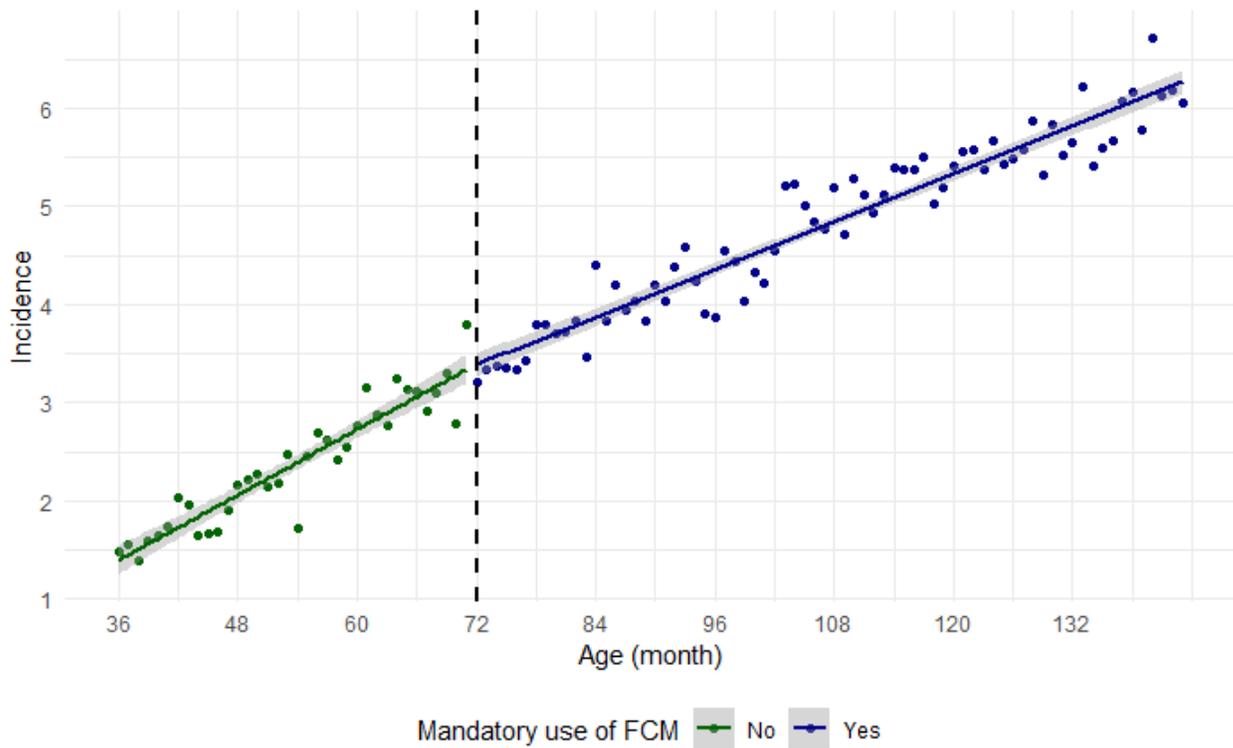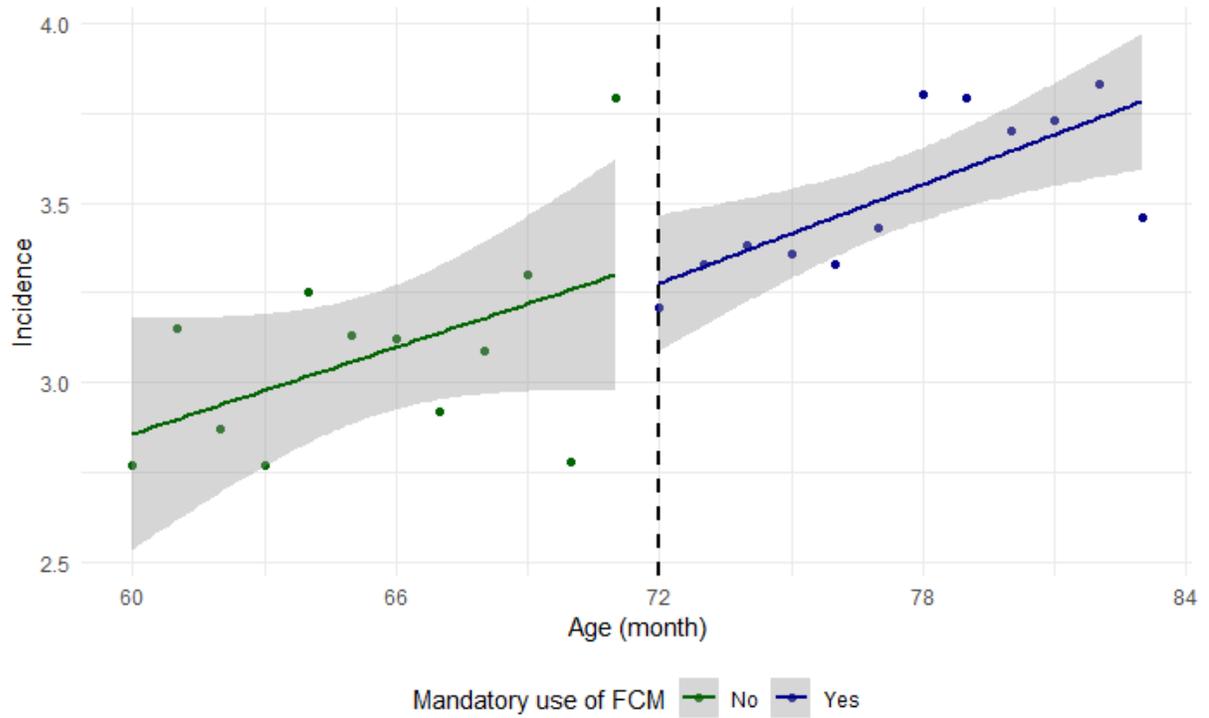
Figure S2. RDD analysis for SARS-CoV-2 incidence using age in months for children aged between 60 and 83 months (5 and 6 years old). Observed incidences by months of age are split at the cutoff at 72 months (discontinued black line). The shadow areas indicate the 95% CI of the corresponding regression lines (green: 60 to 71 months, blue: 72 to 83 months).

Appendix 2: Details of the regression model (simulation)

Methods

We performed a simulation analysis assuming that the age trend observed in previous studies is a parameter that should be maintained in our data across the different grades. First, we fitted a linear regression model to data of primary education (6 to 11 years of age) for incidence, SAR and R*. Then, we extrapolated the model to preschool groups and assessed which the expected values would be if the primary trend was followed. The fittings were performed using the function *fitlm* of Matlab 2021b. The 95% CI were assessed using the function *predict*. This function was also used to extrapolate the model to preschool grades.

Results

The linear regression model to data from primary education grades provides an R^2 of 0.99 (incidence-age), 0.95 (SAR-age) and 0.96 (R*-age). After extrapolating a backward regression to P5, we observe that both SAR and R* are, non-significantly, 18% higher than those expected from the regression model, while the incidence remains 2% below the expected value. The difference between the number of expected and observed secondary cases for P5 was 162 (95% CI: -28–352) in a population of 63,344 students (0·3%) during the whole study period. On the other hand, P3 and P4 show mean SAR values that are 19% (P3) and 18% (P4) lower than those expected from the regression model. The observed R* values would be 24% (P3) and 20% (P4) lower than those expected, and the incidences would be 21% (P3) and 14% (P4) below the expected values (**Figures S3, S4** and **S5**).

Figure S3. Linear regression model of incidence with age. The regression model is fitted to data of primary grades (6 to 11 years old). The shadow indicates the 95% CI of the fitting. Observed values are split between those that were used in the regression model (black points, primary education grades) and those that were not (blue points, preschool grades).
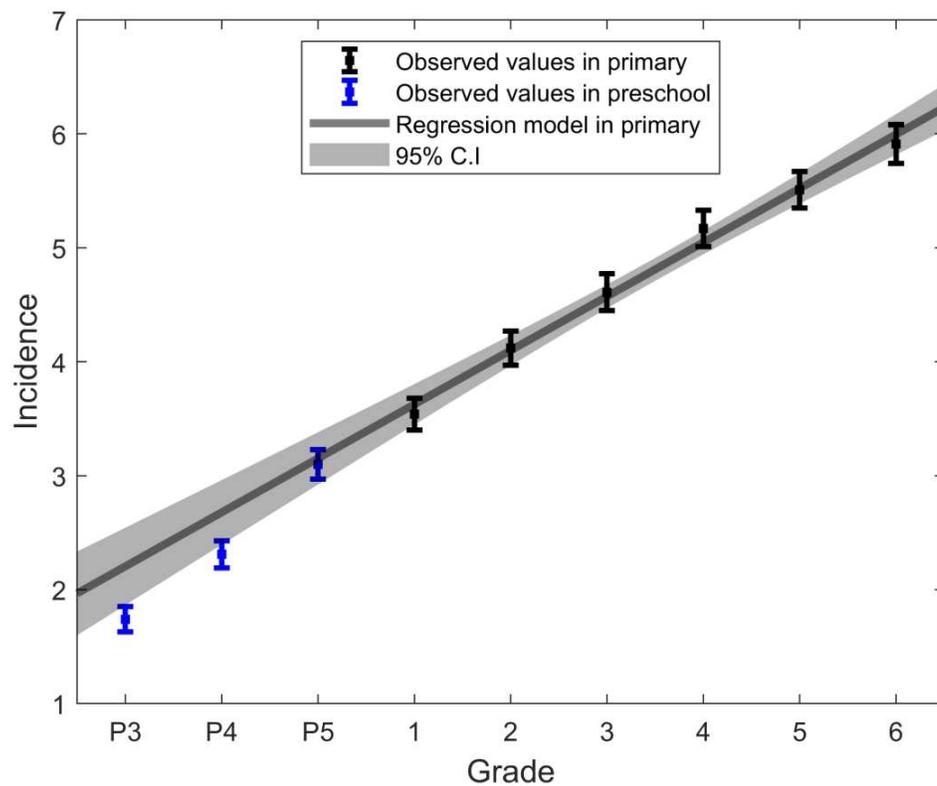
Figure S4 Linear regression model of secondary attack rate (SAR) with age. The regression model is fitted to data of primary education grades (6 to 11 years old). The shadow indicates the 95% CI of the fitting. Observed values are split between those that were used in the regression model (black points, primary education grades) and those that were not (blue points, preschool grades).
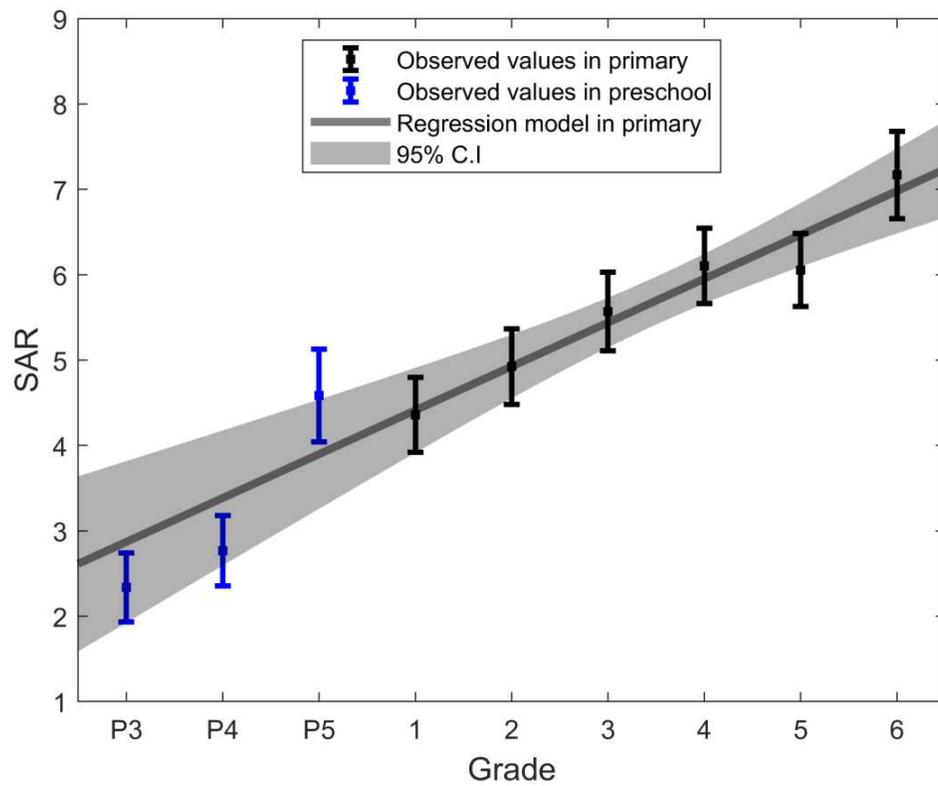
Figure S5 Linear regression model of effective reproduction number (R*) with age. The regression model is fitted to data of primary education grades (6 to 11 years old). The shadow indicates the 95% CI of the fitting. Observed values are split between those that were used in the regression model (black points, primary education grades) and those that were not (blue points, preschool grades).