SUPPLEMENTARY MATERIAL

We applied Bayesian methods, as these approaches yield probability distributions of the parameters of interest (posterior probabilities) based on prior probability distributions and on the observed data. Bayesian probability is understood as a measure of belief or certainty about a parameter. This approach allows for the integration of prior information with current data, based on which posterior probability distributions are obtained.

In our analyses, the probability of losing one hour of school was assumed to follow a Beta distribution. The Beta distribution is bounded in the $(0\,;1)$ interval and is often used to model probabilities. The Beta distribution can be defined by two parameters, α and β, which themselves are defined by the mean and the precision. Reflecting minimal prior assumptions (i.e., we did not have strong prior beliefs on the probability of losing one hour of school), we assigned non-informative uniform priors (dunif[0,1] and dunif[0,1000]) to these parameters, allowing the data to predominantly shape the posterior distributions.

To obtain the probability distributions of the parameters of interest, we gathered data on academic absenteeism and presenteeism through the WPAI:AS+CIQ questionnaire in the MASK-air app. We assumed that the proportion of hours lost in a week (whether due to absenteeism, presenteeism or both) could be modeled by a binomial distribution defined by the probability of losing one hour in the total number of hours in class during a week.

To present our findings, we described the percentage of academic hours lost in terms of absenteeism, presenteeism and total hours lost using the medians and the 25th and 75th percentiles of the respective posterior predictive distributions (obtained by simulating from the model using the sampled parameter values from the posterior distribution). Importantly, in Bayesian statistics, uncertainty is

doi: 10.18176/jiaci.1010

expressed through these credible intervals which allow for a direct probabilistic interpretation, instead of classical confidence intervals.

We used simulations of two chains and assessed Markov chain Monte Carlo convergence by means of visual inspection of the trace and history plots of all parameters and the Gelman-Rubin diagnostics. An effective sample size of 4,000 and an $\hat{R} \leq 1.01$ were obtained for all analyses.

doi: 10.18176/jiaci.1010