nature portfolio

Peer Review File

Retrospective evaluation of school-related measures on pre-vaccination transmission dynamics of SARS-CoV-2

Corresponding Author: Mr Rey Audie Escosio

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Dear Editor,

Thank you very much for the honour and pleasure for allowing me to review the article “Retrospective evaluation of school-related measures on pre-vaccination transmission dynamics of SARS-CoV-2” by Canforra and Escosio et al. Please find below my recommendation and comments.

As a limitation, I would like to state that while I have completed substantial work on the epidemiology of SARS-CoV-2 in schools as well as assessments of interventions targeting this contact domain, I am an expert in epidemiology rather than an expert in modeling. Thus, my review will focus primarily on the assumptions, use of data, and epidemiological logic of the paper whereas I would recommend an expert of modelling methodology as a second reviewer to reliably assess this methodological aspect of the paper.

Regarding the novelty of the finding, the effect and use of school closures is well-studied and the findings of the review do not provide a completely novel insight for the field. It is true that this analysis is the first of its kind for Portugal and the approach of retrofitting is not commonly used. However, given the scarcity of empirical data for retrofitting the model to the Portuguese context, other countries with more available and specific data on contacts (and their change over time) and more extensive seroprevalence data could have served as a more appropriate basis to answer the central study questions. Additionally, as the authors state themselves within the discussion, a mere analysis of the epidemiological effects of school closures within the current discourse is of less interest if it does not also include a measure of the collateral effects. In particular a discussion of scenario 2 would have benefitted from a consideration of the unintended consequences of such prolonged school closures complementing the epidemiological analysis.

See for example:

European Center for Disease Prevention and Control. COVID-19 in children and the role of school settings in transmission - second update. 8 July 2021. Stockholm: ECDC; 2021. doi:10.2900/48630
Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, Gavenciak T, et al. Inferring the effectiveness of government interventions against COVID-19. Science. 2021;371(6531):eabd9338. doi:10.1126/science.abd9338
Walsh S, Chowdhury A, Braithwaite V, Russell S, Birch JM, Ward JL, et al. Do school closures and school reopenings affect community transmission of COVID-19? A systematic review of observational studies. BMJ Open. 2021;11(8):e053371. doi:10.1136/bmjopen-2021-053371
Hume S, Brown SR, Mahtani KR. School closures during COVID-19: an overview of
systematic reviews. BMJ Evidence-Based Medicine. 2022;28(3):164-174. doi:10.1136/bmjebm-2021-111849
Hammerstein S, Konig C, Dreisorner T, Frey A. Effects of COVID-19-related school closures on student achievement - a systematic review. Front Psychol. 2021;12:746289. doi:10.3389/fpsyg.2021.746289
Zierer K. Effects of pandemic-related school closures on pupils’ performance and learning in selected countries: a rapid review. Education Sciences. 2021;11(6). doi:10.3390/educsci11060252
Panagouli E, Stavridou A, Savvidi C, Kourti A, Psaltopoulou T, Sergentanis TN, et al. School performance among children and adolescents during COVID-19 pandemic: a systematic review. Children (Basel). 2021;8(12). doi:10.3390/children8121134

Chaabane S, Doraiswamy S, Chaabna K, Mamtani R, Cheema S. The impact of COVID-19 school closure on child and adolescent health: a rapid systematic review. Children (Basel). 2021;8(5). doi:10.3390/children8050415
Engzell P, Frey A, Verhagen MD. Learning loss due to school closures during the COVID-19 pandemic. Proc Natl Acad Sci U S A. 2021;118(17). doi:10.1073/pnas.2022376118
Viner R, Russell S, Saulle R, Croker H, Stansfield C, Packer J, et al. School closures during social lockdown and mental health, health behaviors, and well-being among children and adolescents during the first COVID-19 wave: a systematic review. JAMA Pediatr. 2022;176(4):400-9. doi:10.1001/jamapediatrics.2021.5840
Meherali S, Punjani N, Louie-Poon S, Abdul Rahim K, Das JK, Salam RA, et al. Mental health of children and adolescents amidst COVID-19 and past pandemics: a rapid systematic review. Int J Environ Res Public Health. 2021;18(7) doi:10.3390/ijerph18073432.
Viner R, Bonell C, Blakemore SJ, Hargreaves J, Panovska-Griffiths J. Schools should still be the last to close and first to open if there were any future lockdown. BMJ. 2022;376:o21. doi:10.1136/bmj.o21
Kourti A, Stavridou A, Panagouli E, Psaltopoulou T, Spiliopoulou C, Tsolia M, et al. Domestic violence during the COVID-19 pandemic: a systematic review. Trauma Violence Abuse. 2023;24(2):719-745. doi:10.1177/15248380211038690
Azevedo J, Hasan A, Goldemberg D, Iqbal S, Geven K. Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: a set of global estimates. World Bank. 2020. doi:10.1596/1813-9450-9284
Hanushek E, Woessman L. The economic impacts of learning losses. OECD Education Working Papers. 2020;225. doi:10.1787/21908d74-en
Fuchs-Schündeln N, Kuhn M, Tertilt M. The short-run macro implications of school and child-care closures. IZA DP. 2020;13353. doi:10.2139/ssrn.3565706

Regarding the manuscript in general, it is well written, the abstract and introduction are accessible, and the conclusions drawn are reasonable. Since the subject of school closures is well-studied, a more extensive discussion of the existing literature could have been provided, including not only results from modeling approaches but also empirical data from observational studies (for examples, see papers provided above and under the comment sections). The methodology and statistical tests themselves are appropriate, and description of error bars and probability values appear adequate.
My main concern is the input data used and in particular the contact data and matrices as well as the mode of inclusion of other and concurrent interventions into the statistical analysis.

1. Comments to the Authors

Thank you very much for the pleasure of reviewing this article. The article “the article “Retrospective evaluation of school-related measures on pre-vaccination
transmission dynamics of SARS-CoV-2” by Canforra and Escosio et al explores the effect of school closures in Portugal in the pre-vaccination period of the SARS-CoV-2 pandemic. The study uses a modeling approach based on seroprevalence and hospitalisations data from Portugal to investigate alternatives scenarios of not applying school closure during the initial phase and again applying school closures during the second phase of the pandemic and their respective effects on seroprevalence and hospitalisations in Portugal.

Comments

My main comments concern what are, in my opinion, oversimplifications of the pandemic dynamics used as basic assumptions as well as a the empirical data basis for the model. The majority of parameters used in the analysis are based on model estimates with only some of them taken from the literature, or empirical data as benchmarks for certain time points within the model sequence. Thus, the literature and empirical basis and all assumptions need to be sound in order to arrive at a reliable model parametrisation and dependable output.

Comment 1

The authors rely for their contact matrices and patterns, including the changes observed during the first lockdown, on the publications on Mistry and Backer et al.

These publications are methodologically sound, however, do not provide sufficiently granular data to address the specific context of schools and school-based contacts and their changes in response to interventions and over time. Both papers may provide age-specific contact rates and matrices. The resulting assumptions include, for example:

“during the second wave, transmission relevant school contacts were already reduced by approximately 80% relative to pre-pandemic levels (Supplementary Table S1)”

However, studies including data that is more granular and appropriate for this context have been published, including data to differentiate school, household, and other contact areas. More adequate data is available and should be relied upon for estimation of contact behavior, rates and changes in response to school closures. Additionally, contact pattern during vacations and prolonged school closures show a complex pattern that should not be neglected within the analysis. I would like to invite the authors to present a justification why more adequate contact data sources were not used for the model, or provide estimates that show that a more detailed contact data input does not alter the outcome of the analysis.

Examples include:

Specific for schools:

Mastrandrea R, Fournet J, Barrat A: Contact Patterns in a High School: A Comparison between Data Collected Using Wearable Sensors, Contact Diaries and Friendship Surveys. PLoS ONE 2015, 10(9):e0136497.

Fournet J, Barrat A: Contact patterns among high school students. PLoS ONE 2014, 9(9):e107878.

Mikolajczyk RT, Akmatov MK, Rastin S, Kretzschmar M: Social contacts of school children and the transmission of respiratory-spread pathogens. Epidemiology and Infection 2008, 136(6):813-822.

Social contact data more specific for Portugal:

Wong KLM, Gimma A, Coletti P, et al. Social contact patterns during the COVID-19 pandemic in 21 European countries - evidence from a two-year study. BMC Infect Dis. 2023;23(1):268. Published 2023 Apr 26. doi:10.1186/s12879-023-08214-y

Data on changes in contact behaviour specific for schools in response to measures:

Phuong HT, Bartz A, Jarynowski AK, et al. Changes in social contact patterns in Germany during the SARS-CoV-2 pandemic - an analysis based on the COVIMOD study. BMC Infect Dis. 2025;25(1):588. Published 2025 Apr 23. doi:10.1186/s12879-025-10917-3

Rodiah I, Vanella P, Kuhlmann A, et al. Age-specific contribution of contacts to transmission of SARS-CoV-2 in Germany. Eur J Epidemiol. 2023;38(1):39-58. doi:10.1007/s10654-022-00938-6 (already cited in paper and includes contact matrices)

Data on changes in contact behaviour during school vacation:

Eames KT, Tilston NL, Edmunds WJ. The impact of school holidays on the social mixing patterns of school children. Epidemics. 2011;3(2):103-108. doi:10.1016/j.epidem.2011.03.003

Robert Koch Institut. Wirksamkeit und Wirkung von anti-epidemischen Maßnahmen auf die COVID-19-Pandemie in Deutschland (StopptCOVID-Studie). Berlin: Robert Koch Institut; 2023.

Comment 2

I would have appreciated a definition of “school closure” in terms of the real-life impact on school-based contacts.
We found in our own work that school closures are not usually associated with a complete depletion of school-related contacts as they are not absolute, e.g. there is a continuation of contacts for special needs children or children of key workers in place throughout most school closures in OECD countries. Thus, these are rather one end of the spectrum of measures to reduce presence of school children but in practice never reach 0. In particular the continuation of supervision of special needs students means that within this population contacts levels remain rather constant even in times of “school closures” as deemed by policy.
The term ζakl as used, e.g., in Equation 2, is denoted with the explanation that this would only include non-school contacts. Since we found that during periods of “school closures”, depending on their exact application, attendance rates could remain as high as 5-12%, a justification of this approach or a definition of “school closure” closer to the real-life situation in Portugal during these periods would be appreciated.

Heinsohn T, Lange B, Vanella P, et al. Infection and transmission risks of COVID-19 in schools and their contribution to population infections in Germany: A retrospective observational study using nationwide and regional health and education agency notification data. PLoS Med. 2022;19(12):e1003913. Published 2022 Dec 20. doi:10.1371/journal.pmed.1003913

Comment 3

Building on the previous two comments: The authors use as their first scenario the omission of school closures from the pandemic response. Within their model, in-school contacts remain unchanged (Equation 3). This is a rather theoretical scenario. The authors state that it is reasonable to assume that measures as were observed after the summer vacations were unlikely to be implemented immediately as schools remained open. This is a reasonable assumption, however, a complete continuation of within-school contacts at the prepandemic level is equally unlikely.

This assumption is justified by the abrupt onset of the pandemic in early 2020, which likely left insufficient time to implement effective within-school measures. As a result, the marginal benefit of closing schools was greater during this period. We confirmed this mechanism in sensitivity analyses, which tested the effect of applying a mitigation scaling factor to school contacts in Scenario 1. Introducing a reduction in school-based contacts attenuated the benefit of school closures, supporting the interpretation above.

If we look at the countries that refrained from imposing school closures during the first wave in Europe, namely Sweden and Iceland (and Belarus), we can see that open schools did not equal open schools at the pre-pandemic level. In Sweden, where primary schools remained open, the minimum package consisted of stay-at-home orders for symptomatic students, hand hygiene and intensified surface cleaning, adaptions to schedules to avoid crowding, recommendations to hold classes outdoors, and cancelling of e.g. choirs and assemblies. In Iceland, cohorting and capping of class sizes, distancing for teachers, hand and other hygiene measures, regular symptom checks, and cancellations of high-risk congregations were instituted. Further measures were optional and left to the schools. Within the sensitivity analysis, this is considered (Scenario 1.2).
However, from a real-world perspective, it seems more likely that a scenario like Scenario 1.2 should have been part of the main analysis of this paper and a scenario without any reduction in school contacts would have formed a sensitivity analysis. This would have avoided the need to include the assumption that other measures would not have been relaxed until the summer vacation as part of the Scenario 1 to make it more plausible.
I would thus recommend instead an adjusted Scenario 1.2 as the main focus of the analysis, or otherwise would require further justification of choosing Scenario 1 as the main output.

Comment 4

The authors assume for their model different susceptibilities across age groups. The relative susceptibilities assumed for 3 age groups are based on one contact tracing study contacted in China during the early pandemic:

The relative susceptibilities of the 0-20 years age group, $\beta_{1,2,3}$, and of the 20-60 years age group, $\beta_{4,5,6,7}$, had log-normal priors with medians 0.23 and 0.64, respectively, and a scale parameter of 0.5 [39]. The relative susceptibility of the 60+ age group was set to $\beta_{8,9,10} \equiv 1$

The estimates of relative susceptibility, in particular for children, seem to be rather extreme estimates compared to other literature including high-quality meta-analysis, e.g. Viner et al. I would like to see a justification of this parameter and the specific study used since the literature on this is rather extensive and the chosen value appears to be rather an outlier compared to other examples.

Viner RM, Mytton OT, Bonell C, et al. Susceptibility to SARS-CoV-2 Infection Among Children and Adolescents Compared With Adults: A Systematic Review and Meta-analysis JAMA Pediatr. 2021;175(2):143-156. doi:10.1001/jamapediatrics.2020.4573

Comment 5

This comment concerns the stated assumption that hospitalised cases did not contribute to pandemic dynamics:

“Second, hospitalized individuals were assumed to be non-infectious, as they constituted a small proportion of the total infectious population and were subject to various personal protective measures in hospitals.”

In fact, during the first wave estimate range around 20% of recorded hospitalisations for SARS-CoV-2 were nosocomial in origin (data mainly from UK, for Portugal I could only identify one article - see below). Even though their contribution to total infections and seropravelence was estimated to be low, their contribution to hospitalisations (one central outcome of the paper) was not negligible. I would like to see a justification or sensitivity analysis to show that an inclusion of hospitalised cases towards further hospitalisations would not alter the outcomes of the analysis.

Knight GM, Pham TM, Stimson J, et al. The contribution of hospital-acquired infections to the COVID-19 epidemic in England in the first half of 2020. BMC Infect Dis. 2022;22(1):556. Published 2022 Jun 18. doi:10.1186/s12879-022-07490-4

Bhattacharya A, Collin SM, Stimson J, et al. Healthcare-associated COVID-19 in England: A national data linkage study. J Infect. 2021;83(5):565-572. doi:10.1016/j.jinf.2021.08.039

For Portugal:
https://www.paginaum.pt/2022/01/12/morreram-mais-de-1-300-pessoas-que-apanharam-covid-19-nos-hospitais-durante-a-primeira-fase-da-pandemia

Comment 6

It is laudable that the authors included seroprevalence data as a benchmark for their model calibration, despite the large uncertainty around seroprevalence estimates within the cited “First National Serological Survey (ISNCOVID-19)” conducted in Portugal between 21 May 2020 and 8 July 2020. The authors state:

Therefore, we also incorporated seroprevalence data, which was sampled around day tsero = 93 post 26 February 2020. We neglected the waning of immunity and assumed the probability of a positive serological test for a randomly selected individual in age group k to be pk ≡ (1 – Sk(tsero)/Nk).

t(sero) would in this case fall on the 29th of May 2020, which is skewed toward the beginning of the sampling period. Taking into account the diagnostic test performance for detection of seroprevalence assumed by the serological survey at 10 days after symptom onset, we reach the beginning of June as the cut-off date. Due to the low levels of incidence during the summer period in Portugal, I would have liked to see a sensitivity analysis where the seroprevalence data is assumed not to present just the first wave until the summer vacation, but also an analysis where this data is assumed to cover also the summer vacation for school children to see whether this would change effects of school closures.

Comment 7

The authors describe a delay within their model regarding the uptake and thus effect of changes:

The rates Ki, governing the speed at which policy changes took effect, were given Exp(1) priors, implying that the uptake of measures took around 6 days

It is not quite clear how this assumption is justified. Rather, a different effect has been observed especially for the early phases of the pandemic, when uptake of measures was often linked to their public announcement rather than their date of coming into effect.
Germany’s largest analysis of pandemic measures found that the effect of measures on R is measurable 2 days prior to the measure actually coming into effect due to anticipatory behaviour changes in the population – an effect more pronounced towards the early phases of the pandemic.

Robert Koch Institut. Wirksamkeit und Wirkung von anti-epidemischen Maßnahmen auf die COVID-19-Pandemie in Deutschland (StopptCOVID-Studie). Berlin: Robert Koch Institut; 2023.

A similar effect is described for other countries, including Spain:
Santamaría L, Hortal J. Chasing the ghost of infection past: identifying thresholds of change during the COVID-19 infection in Spain. Epidemiology and Infection. 2020;148:e282. doi:10.1017/S0950268820002782

Comment 8

The concurrence of introduction of measures during the pandemic often poses challenges to the differentiation of their effects. It is not quite clear how, within the setup of the study, where all interventions other than school closures were subsumed into a single parameter denoting “non-school contacts”, this complexity would have been captured and the problem at hand resolved. I would like to see a more detailed overview of the measures introduced in Portugal during the course of the study period (this overview could be placed in the supplementary material) to discern the relative political importance of school measures at the time. Furthermore, I would like to see a further discussion of the issue of concurrence and differentiation of effect, and how it was resolved in the design of the model to isolate the specific effect of school closures from other concurrent measures both within the first lockdown as well as within the second pandemic wave as measures increasingly stringent measures were introduced.

Reviewer #2

(Remarks to the Author)
The authors infer the impact of school closures in Portugal and use this to explore counterfactual scenarios. The methodology is robust and the conclusions supported by the analysis. Also all steps necessary to ensure reproducibility seem to have been taken. I have only few minor remarks:

* p.6 line 4: is 1-zeta the reduction, or zeta?, same question of zeta_s, try to be clearer here.
* Equation 6 and 7, write the variable for which you model the probability in parentheses after the P: e.g. P_{hosp}(h_{k,i}) = ... and P_sero(l_k) = ... Otherwise one has to guess it.
* In equation 6 and 7 and 8, don't use P, but L_{hosp} or l_p or f, it isn't a probability, but a probability density/likelihood. It gives the wrong impression and might lead to misunderstandings if a P is used here. Probabilities are the integrals of the the probability densities.

* Report how well the chains converged, e.g. via R-hat.
* Fig. 2+3: Add a legend to explain what the blue and orange lines/regions are, so that one doesn't have to read the caption.
* Results section, epidemiological impact: reduce the number of significant digits you report, if the credible interval is much larger, it gives a misleading impression of precision.

Reviewer #3

(Remarks to the Author)
The manuscript by Canfora et al., presents a retrospective study on the impact of school closure and reopening on SARS-CoV-2 transmission in Portugal in the first year of the pandemic, prior to mass vaccination.

In this work, the authors use a previously developed age-stratified compartmental model fitted to hospitalization and seroprevalence data to reconstruct SARS-CoV-2 transmissibility conditions during the spring and autumn of 2020. By producing counterfactual scenarios, they assess the impact of school closures on hospitalization burden, concluding that decisions on schools are influenced both by school-based mitigation measures and the intensity of transmission in the community captured by the time-varying reproductive number.
Age-specific contributions to transmission are also estimated through elasticity analysis, offering insights into the role of different age groups. However, these results do not appear to be fully integrated into the narrative (for example, they are not explicitly mentioned in the abstract) despite representing a potentially valuable contribution to the study.

Although the manuscript is built on a validated modeling approach - with limitations appropriately acknowledged and a comprehensive sensitivity analysis - I do have comments regarding some assumptions and scenario choices. In conclusion, I consider this work suitable for publication with some revisions. Below, I provide a point-by-point list.

INTRODUCTION

1. Relevant transmission modeling works on school and community interventions to safely reopening schools after the first lockdown - such as those in France - are not explicitly mentioned. I would suggest incorporating a more in-depth review of the literature to better contextualize the complexity of decision-making around school reopening and closing during the pandemic.

METHODS

1. Could you please clarify why you chose to stratify the population into ten age classes instead of five as those in the seroprevalence data? This would have allowed to sum over the same age groups in the loglikelihood computation without aggregating a posteriori. Additionally, I wonder whether larger age classes might improve the model fit, especially considering that hospitalization data for children in the [0–5) and [5–10) age groups appear to be sparser than for other age categories.

2. Could you specify how the goodness of fit was assessed?

3. Related to 1., you calibrate age-specific susceptibility yet this does not vary among individuals aged 0 to 19 years. This assumption might reduce the need for a finer age resolution. At the same time, adolescents may exhibit more similar epidemiological characteristics to adults. Have you consulted the recent literature on age-specific susceptibility?

4. In Scenario 1, you assume that school contact rates would have remained at pre-pandemic levels, despite the presence of strong restrictions in other settings. Retrospectively, it might be more informative to explore alternative scenarios: What would have happened if the same mitigation measures of scenario 2 had been applied during the first lockdown (currently in the SI)? What kind of interventions could have allowed schools to remain open while keeping hospitalizations at reasonable levels? This would better support pandemic preparedness.

5. The relaxation of restrictions following the first lockdown is not fully detailed in the manuscript. I suppose this refers to general measures such as mask-wearing, physical distancing, and refraining from handshakes, as mentioned in the section on Time-varying contact patterns. Does this also include remote working or other measures? Similarly, the protective measures implemented in schools - such as reduced classroom capacity, increased outdoor activities, staggered schedules, and enhanced cleaning protocols - are briefly mentioned in the introduction but not described elsewhere. Were these protective measures consistent across all school levels, or did they vary?

6. I would recommend adding a brief explanation of the elasticity matrix and its interpretation, especially for readers who may be less familiar with this concept.

7. The notation for the regimes in the text is not consistent (first labels from i) to v), then numbers ranging from 0 to 5). Standardizing the notation may facilitate the reading.

8. Figure 1 includes elements - such as the ‘midpoint’ or scenario 1 and 2 - that have not yet been introduced in the text when it is first cited. Providing this context beforehand may improve comprehension.

9. In Figure S3, the average number of contacts per day for children aged 5-10 years is 0.8 for the pre-pandemic period, which seems low. If I am interpreting the figure correctly, shouldn't this value correspond to the red line shown in Figure S1F for the 5-10 age group?

RESULTS

1. Before reaching the results section, I did not realize that the restrictions from the first lockdown were not relaxed in scenario 1. This should be more clearly stated also in the Methods section describing the scenarios.

2. The frequent changes in age group throughout the results can be confusing. Additionally, I found a bit inconsistent that results put in the Supplementary Information were extensively discussed in the main text. While I understand that aggregating into two age classes reflects how decisions were made by policymakers at the time, I would encourage reconsidering this choice if needed.

3. It would be helpful to include a legend in Figure 2, as it is currently missing. Clarifying the color-coding used for the lines would help reading the plot. I would also mention in the caption that you are showing absolute number and not new daily hospital admission per 100,000.

4. The section or figure titles could be revised to more clearly reflect the main findings. For example, the title of the first section could be more informative by explicitly stating the epidemiological impact of school closures on daily hospitalizations. Similarly, for impact of scenarios on hospitalizations in Figure 2. Additionally, without the use of descriptive keywords, reader has to repeatedly recall the specific definitions of Scenario 1 and Scenario 2. Introducing concise labels or keywords may improve readability.

5. 'Impact of school closures: Scenario 1 versus Scenario 2' seems more a discussion point than a result.

DISCUSSION

1. The authors cite as a strength the fact that age-specific mixing patterns were inferred directly from the data rather than imposed a priori. However, this approach cannot be validated against empirical contact data due to their lack, limiting the ability to assess the accuracy of the inferred patterns.

GENERAL

The main text would benefit from some shortening to better highlight the most important points (the current word count appears to exceed 7,000, if I'm not mistaken).

Equations that have already been presented in the previous work could be moved to the Supplementary Information. Including a diagram to illustrate disease progression and rates - distinguishing between parameters informed by the literature and those calibrated through model fitting - could be an alternative.

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
Thank you for your revision and answers. From my side, it can be published as is.

Reviewer #3

(Remarks to the Author)
I thank the authors for addressing the points I previously raised. However, I believe that the writing would still benefit from some refinement.

Abstract
The results in the abstract are currently presented beginning with the elasticity analysis, whereas in the main text this section appears after the scenario analysis. I suggest to align the order for greater consistency.

Introduction
The additions to the introduction do not sufficiently highlight the differences across the cited studies. Among the references included, it remains unclear how each study differs in its approach and main findings relative to the authors' research questions. For instance, reference [24] also introduces household risk, while references [20], [21], and [22] focus more on within-school transmission in later phases of the pandemic rather than at its onset. In contrast, reference [23] is more

explorative and do not use epidemiological data for calibration. Providing a clearer description of the specific contributions of each cited study would help readers better situate the current work within the existing literature.

Safe protocols for reopening schools were also assessed here: Di Domenico, L., Pullano, G., Sabbatini, C.E. et al. Modelling safe protocols for reopening schools during the COVID-19 pandemic in France. Nat Commun 12, 1073 (2021). https://doi.org/10.1038/s41467-021-21249-6

Finally, I feel that the complexities involved in decision-making around school closures and reopenings during pandemics do not currently emerge from the text. Several factors can play a role in these decisions, including ensuring learning continuity, preventing social isolation among children, and reducing barriers to education. These important aspects are not mentioned and their consideration would provide a more comprehensive perspective.

Discussion
In my previous point-by-point review, I asked what types of interventions could have allowed schools to remain open while keeping hospitalizations at reasonable levels. Could the authors please comment on this aspect to better support their conclusion regarding school closures during future pandemics?
The current response stating that the choice of interventions falls within the domain of policymakers is understandable. However, it would be valuable to include a brief discussion on the effectiveness and feasibility of full school closures.

Reviewer #4

(Remarks to the Author)
Comments on Reviewer 1's requests:

A. General methodological transparency (Reviewer 1 – General comment, Comment 1–2)

Several of Reviewer 1’s general methodological requests remain insufficiently addressed:

1. Assumptions transparency
- The reviewer requested a structured overview (e.g. a table) listing:
- Key model assumptions,
- Their empirical or literature support,
- Sensitivity analyses performed for each assumption, and
- The impact on main outcomes (with 95% credible intervals).
While assumptions are discussed throughout the text and supplement, this information is not yet consolidated in a transparent, reviewer-readable format, nor is it clearly stated which assumptions were revised in response to review.

2. Collateral effects of school closures
- The manuscript still lacks a dedicated “policy trade-offs” discussion summarizing evidence on learning loss, mental health, equity, and macroeconomic impacts of school closures. Reviewer 1 explicitly requested such a discussion, supported by the extensive literature already provided, and framed in the Portuguese context.
At present, epidemiological effects are discussed largely in isolation, which limits policy interpretability.

3. Portugal-specific contact data mapping
- The reviewer requested a table summarizing:
- Available Portugal-specific contact data (sources, dates),
- Overlap with the study period,
- How these data were used (calibration, validation, exclusion). This mapping remains missing, and it is still unclear how Portugal-specific contact evidence constrains or validates the inferred contact dynamics.

4. Quantification of new sensitivity analyses
- Several newly added sensitivities (holidays, mitigation sweeps, partial closures) are described qualitatively, but Reviewer 1 explicitly requested numerical effects (e.g. changes in peak and cumulative hospitalizations and age-specific seroprevalence) with 95% CrIs, ideally summarized in a compact table in the main text.

5. Justification of the “~80% reduction” in school contacts
The claim of strong school contact reduction is not yet supported by:
- A plot of the inferred school contact multiplier over time with uncertainty, or
- Explicit triangulation against external school-specific evidence to justify plausibility.

6. Identifiability and attribution diagnostics
The reviewer requested explicit discussion and diagnostics (e.g. posterior correlations, simulation-recovery) clarifying:
- What aspects of school vs. non-school effects are identifiable,
- Where attribution remains uncertain.
These diagnostics are still absent.

B. Definition of “school closure” (Reviewer 1 – Comment 3)

Reviewer 1's concern regarding the definition of "school closure" remains partially unresolved.

Specifically:
- The manuscript continues to use "school closure" as implying zero school-based contacts in the core model, while later analyses reintroduce residual contacts through mitigation factors.
- This creates ambiguity in how the term should be interpreted and how parameters (e.g. ζ_{akl}) map to real-world contact patterns.
- The reviewer explicitly requested either:
- A definition closer to actual implementation in Portugal, or
- A clear justification that the zero-contact case is intended as an analytical extreme rather than a literal representation.

To resolve this comment, a brief but explicit clarification in the main text (Methods or Discussion) is still needed, explaining:
- That "school closure" is an analytical reference scenario representing maximal suppression of structured school contacts,
- That real-world closures involved non-zero residual attendance,
- Why a zero or near-zero assumption is nevertheless used (e.g. as an upper-bound counterfactual in the absence of detailed Portugal-specific contact data),
- How readers should interpret results in light of the mitigation-spectrum analyses.

C. Concurrence of measures and attribution (Reviewer 1 – Comment 8)

While the added overview of measures in Portugal is appreciated, the core conceptual concern remains only partially addressed:
- The manuscript still lacks a sufficiently clear explanation of how the model design manages concurrence of interventions when all non-school measures are aggregated into a single parameter.
- It remains unclear:
- Under what assumptions can the effect of school closures be interpreted as distinct,
- What cannot be disentangled given this aggregation,
- How these limitations affect causal interpretation.

Reviewer 1 explicitly requested a more detailed discussion of this issue, which is still needed in the Discussion section or in supplementary information document.

4. Minor residual points (not blocking, but noted)
- Comment 4 (age-specific susceptibility priors): Largely addressed; a brief justification of the choice of prior study or a statement on robustness to alternative priors would fully close the loop.
- Comment 5 (hospitalized individuals assumed non-infectious): Adequately justified; although a sensitivity analysis was requested, the explanation provided is likely sufficient.
- Comment 6 (seroprevalence timing): Adequately addressed; a short quantitative summary of sensitivity results would strengthen completeness.
- Comment 7 (uptake delays, Ki): Fully addressed.

Version 2:

Reviewer comments:

Reviewer #3

(Remarks to the Author)
Although the authors have addressed the suggestions raised during the previous round of review, the main message of the paper remains somewhat unclear to me. In particular, it is not evident whether the authors are suggesting that school closures should still be considered a viable option to curb transmission during future pandemics.

For example, the manuscript states: "Our results suggest that integrated strategies combining school-based and broader community interventions are essential to effectively reduce transmission during future pandemics." It is unclear what is meant by "school-based interventions" in this context. Are the authors referring specifically to school closures, which are the focus of the analysis, or to alternative school-based mitigation measures? If the latter, these interventions are not evaluated in the present study and therefore should introduced and discussed before.

Apart from this point, I believe the authors have addressed the comments.

Reviewer #4

(Remarks to the Author)
All my comments have been adequately addressed.

Response letter

Reviewer 1

Dear Editor,

Thank you very much for the honour and pleasure for allowing me to review the article "Retrospective evaluation of school-related measures on pre-vaccination transmission dynamics of SARS-CoV-2" by Canforra and Escosio et al. Please find below my recommendation and comments.

As a limitation, I would like to state that while I have completed substantial work on the epidemiology of SARS-CoV-2 in schools as well as assessments of interventions targeting this contact domain, I am an expert in epidemiology rather than an expert in modeling. Thus, my review will focus primarily on the assumptions, use of data, and epidemiological logic of the paper whereas I would recommend an expert of modelling methodology as a second reviewer to reliably assess this methodological aspect of the paper.

Regarding the novelty of the finding, the effect and use of school closures is well-studied and the findings of the review do not provide a completely novel insight for the field. It is true that this analysis is the first of its kind for Portugal and the approach of retrofitting is not commonly used. However, given the scarcity of empirical data for retrofitting the model to the Portuguese context, other countries with more available and specific data on contacts (and their change over time) and more extensive seroprevalence data could have served as a more appropriate basis to answer the central study questions.
Additionally, as the authors state themselves within the discussion, a mere analysis of the epidemiological effects of school closures within the current discourse is of less interest if it does not also include a measure of the collateral effects. In particular a discussion of scenario 2 would have benefitted from a consideration of the unintended consequences of such prolonged school closures complementing the epidemiological analysis.

See for example:

European Center for Disease Prevention and Control. COVID-19 in children and the role of school settings in transmission - second update. 8 July 2021. Stockholm: ECDC; 2021. doi:10.2900/48630

Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, Gavenciak T, et al. Inferring the effectiveness of government interventions against COVID-19. Science. 2021;371(6531):eabd9338. doi:10.1126/science.abd9338

Walsh S, Chowdhury A, Braithwaite V, Russell S, Birch JM, Ward JL, et al. Do school closures and school reopenings affect community transmission of COVID-19? A systematic review of observational studies. BMJ Open. 2021;11(8):e053371. doi:10.1136/bmjopen-2021-053371

Hume S, Brown SR, Mahtani KR. School closures during COVID-19: an overview of systematic reviews. BMJ Evidence-Based Medicine. 2022;28(3):164-174. doi:10.1136/bmjebm-2021-111849

Hammerstein S, Konig C, Dreisorner T, Frey A. Effects of COVID-19-related school closures on student achievement - a systematic review. Front Psychol. 2021;12:746289. doi:10.3389/fpsyg.2021.746289

Zierer K. Effects of pandemic-related school closures on pupils' performance and learning in selected countries: a rapid review. Education Sciences. 2021;11(6). doi:10.3390/educsci11060252

Panagouli E, Stavridou A, Savvidi C, Kourti A, Psaltopoulou T, Sergentanis TN, et al. School performance among children and adolescents during COVID-19 pandemic: a systematic review. Children (Basel). 2021;8(12). doi:10.3390/children8121134

Chaabane S, Doraiswamy S, Chaabna K, Mamtani R, Cheema S. The impact of COVID-19 school closure on child and adolescent health: a rapid systematic review. Children (Basel). 2021;8(5). doi:10.3390/children8050415

Engzell P, Frey A, Verhagen MD. Learning loss due to school closures during the COVID-19 pandemic. Proc Natl Acad Sci U S A. 2021;118(17). doi:10.1073/pnas.2022376118

Viner R, Russell S, Saulle R, Croker H, Stansfield C, Packer J, et al. School closures during social lockdown and mental health, health behaviors, and well-being among children and adolescents during the first COVID-19 wave: a systematic review. JAMA Pediatr. 2022;176(4):400-9. doi:10.1001/jamapediatrics.2021.5840

Meherali S, Punjani N, Louie-Poon S, Abdul Rahim K, Das JK, Salam RA, et al. Mental health of children and adolescents amidst COVID-19 and past pandemics: a rapid systematic review. Int J Environ Res Public Health. 2021;18(7) doi:10.3390/ijerph18073432.

Viner R, Bonell C, Blakemore SJ, Hargreaves J, Panovska-Griffiths J. Schools should still be the last to close and first to open if there were any future lockdown. BMJ. 2022;376:o21. doi:10.1136/bmj.o21

Kourti A, Stavridou A, Panagouli E, Psaltopoulou T, Spiliopoulou C, Tsolia M, et al. Domestic violence during the COVID-19 pandemic: a systematic review. Trauma Violence Abuse. 2023;24(2):719-745. doi:10.1177/15248380211038690

Azevedo J, Hasan A, Goldemberg D, Iqbal S, Geven K. Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: a set of global estimates. World Bank. 2020. doi:10.1596/1813-9450-9284

Hanushek E, Woessman L. The economic impacts of learning losses. OECD Education Working Papers. 2020;225. doi:10.1787/21908d74-en

Fuchs-Schündeln N, Kuhn M, Tertilt M. The short-run macro implications of school and child-care closures. IZA DP. 2020;13353. doi:10.2139/ssrn.3565706

Regarding the manuscript in general, it is well written, the abstract and introduction are accessible, and the conclusions drawn are reasonable. Since the subject of school closures is well-studied, a more extensive discussion of the existing literature could have been provided, including not only results from modeling approaches but also empirical data from observational studies (for examples, see papers provided above and under the comment sections). The methodology and statistical tests themselves are appropriate, and description of error bars and probability values appear adequate.
My main concern is the input data used and in particular the contact data and matrices as well as the mode of inclusion of other and concurrent interventions into the statistical analysis.

1. Comments to the Authors

Thank you very much for the pleasure of reviewing this article. The article "the article "Retrospective evaluation of school-related measures on pre-vaccination transmission dynamics of SARS-CoV-2" by Canforra and Escosio et al explores the effect of school closures in Portugal in the pre-vaccination period of the SARS-CoV-2 pandemic. The study uses a modeling approach based on seroprevalence and hospitalisations data from Portugal to investigate alternatives scenarios of not applying school closure during the initial phase and again applying school closures during the second phase of the pandemic and their respective effects on seroprevalence and hospitalisations in Portugal.

Comments

My main comments concern what are, in my opinion, oversimplifications of the pandemic dynamics used as basic assumptions as well as a the empirical data basis for the model. The majority of parameters used in the analysis are based on model estimates with only some of them taken from the literature, or empirical data as benchmarks for certain time points within the model sequence. Thus, the literature and empirical basis and all assumptions need to be sound in order to arrive at a reliable model parametrisation and dependable output.

Response:
Thank you for your thorough review and valuable feedback. We agree that it is important for our assumptions to be sound and well-justified, as they affect the reliability of the model. In response, we have reviewed our assumptions and the data sources used, ensuring they are clearly outlined and supported by existing literature. We have also performed specification checks and added partial school closure scenarios, which we believe strengthens the reliability of our findings. We also highlighted the additional limitations of our study based on your report and those of the other referees.

Comment 1

The authors rely for their contact matrices and patterns, including the changes observed during the first lockdown, on the publications on Mistry and Backer et al.

These publications are methodologically sound, however, do not provide sufficiently granular data to address the specific context of schools and school-based contacts and their changes in response to interventions and over time. Both papers may provide age-specific contact rates and matrices. The resulting assumptions include, for example:

"during the second wave, transmission relevant school contacts were already reduced by approximately 80% relative to pre-pandemic levels (Supplementary Table S1)"

However, studies including data that is more granular and appropriate for this context have been published, including data to differentiate school, household, and other contact areas. More adequate data is available and should be relied upon for estimation of contact behavior, rates and changes in response to school closures. Additionally, contact pattern during vacations and prolonged school closures show a complex pattern that should not be neglected within the analysis. I would like to invite the authors to present a justification why more adequate contact data sources were not used for the model, or provide estimates that show that a more detailed contact data input does not alter the outcome of the analysis.

Examples include:

Specific for schools:

Mastrandrea R, Fournet J, Barrat A: Contact Patterns in a High School: A Comparison between Data Collected Using Wearable Sensors, Contact Diaries and Friendship Surveys. PLoS ONE 2015, 10(9):e0136497.

Fournet J, Barrat A: Contact patterns among high school students. PLoS ONE 2014, 9(9):e107878.

Mikolajczyk RT, Akmatov MK, Rastin S, Kretzschmar M: Social contacts of school children and the transmission of respiratory-spread pathogens. Epidemiology and Infection 2008, 136(6):813-822.

Social contact data more specific for Portugal:

Wong KLM, Gimma A, Coletti P, et al. Social contact patterns during the COVID-19 pandemic in 21 European countries - evidence from a two-year study. BMC Infect Dis. 2023;23(1):268. Published 2023 Apr 26. doi:10.1186/s12879-023-08214-y

Data on changes in contact behaviour specific for schools in response to measures:

Phuong HT, Bartz A, Jarynowski AK, et al. Changes in social contact patterns in Germany during the SARS-CoV-2 pandemic - an analysis based on the COVIMOD study. BMC Infect Dis. 2025;25(1):588. Published 2025 Apr 23. doi:10.1186/s12879-025-10917-3

Rodiah I, Vanella P, Kuhlmann A, et al. Age-specific contribution of contacts to transmission of SARS-CoV-2 in Germany. Eur J Epidemiol. 2023;38(1):39-58. doi:10.1007/s10654-022-00938-6 (already cited in paper and includes contact matrices)

Data on changes in contact behaviour during school vacation:

Eames KT, Tilston NL, Edmunds WJ. The impact of school holidays on the social mixing patterns of school children. Epidemics. 2011;3(2):103-108. doi:10.1016/j.epidem.2011.03.003

Robert Koch Institut. Wirksamkeit und Wirkung von anti-epidemischen Maßnahmen auf die COVID-19-Pandemie in Deutschland (StopptCOVID-Studie). Berlin: Robert Koch Institut; 2023.

Response:
Thank you for this important comment on our use of social contact data. We agree that pandemic contacts were complex, shaped both by policy interventions in and outside schools, and by behavioral changes.

Our estimated contacts are qualitatively consistent with Rodiah et al. [EJE, 2023], specifically regarding children and adults up to age 34 carrying the majority of contacts during the pandemic. We note that direct quantitative comparison with non-Portuguese studies, such as the German contact patterns, is difficult due to unaligned intervention measures outside of the first lockdown.

We are not aware of extensive additional data for Portugal that would allow us to model the contacts in their entire complexity at the household and societal level for the entire period of our study – Wong et al. [BMC Infect. Dis., 2023] covers only a short period. Therefore, we have quantified this complexity by: (1) assuming additivity of school and non-school contacts; (2) allowing non-school contacts to transition from a more stringent regime to a less stringent one and vice versa, with these transitions estimated within the model via logistic functions in time. The latter allows for gradual and multiple non-pharmaceutical interventions, for behavioral changes such as reducing contacts in anticipation of a measure being implemented, and for some changes during the holiday period. Additionally, we included a school mitigation factor to take into account other measures such as mask wearing and social distancing to better quantify differential changes in non-school contacts from one regime to the next.

For model fitting, evidence that we used sufficient data sources to model the March 2020–January 2021 timeframe is the internal model validity and ability to reproduce all age-group hospitalization waves over a year (Supplementary Figures S10-S11).

However, to address your concerns for the counterfactuals, specifically for how the school contacts may have been shaped by other non-pharmaceutical or behavioral interventions in the counterfactual scenarios, we added new results.

First, we addressed the impact of holidays on our counterfactual scenarios via a sensitivity analysis which confirmed the main findings on school closure effects in Scenario 2 were not substantially changed (Scenario 2.1; Supplementary Figure S32). This finding aligns with Earnes et al. [Epidemics 2011], who observed that children's contacts halved during holidays (our Relax 3 in Supplementary Figure S5), without drastically affecting the overall results (Figures 3 and 4 versus Supplementary Figure S32).

Second, we now added in the main text new results that explore the scenarios by varying school mitigation levels (both increasing and decreasing levels in Figure 5). Instead of assuming schools are fully open or closed, Supplementary Section 13 further explores intermediate situations which are discussed in the response to your next comment.

However, we also acknowledge that inferring the dynamic contact structure without additional empirical data on contacts is a limitation of our work.

Main text, Discussion, Page 13, Lines 328-330:
"[...] This approach ensured model validity, as it successfully reproduces observed dynamics; however, due to the lack of empirical contact data, the inferred contacts cannot be formally validated, which represents a limitation of our study."

Comment 2

I would have appreciated a definition of “school closure” in terms of the real-life impact on school-based contacts.
We found in our own work that school closures are not usually associated with a complete depletion of school-related contacts as they are not absolute, e.g. there is a continuation of contacts for special needs children or children of key workers in place throughout most school closures in OECD countries. Thus, these are rather one end of the spectrum of measures to reduce presence of school children but in practice never reach 0. In particular the continuation of supervision of special needs students means that within this population contacts levels remain rather constant even in times of “school closures” as deemed by policy.
The term ζakl as used, e.g., in Equation 2, is denoted with the explanation that this would only include non-school contacts.
Since we found that during periods of “school closures”, depending on their exact application, attendance rates could remain as high as 5-12%, a justification of this approach or a definition of “school closure” closer to the real-life situation in Portugal during these periods would be appreciated.

Heinsohn T, Lange B, Vanella P, et al. Infection and transmission risks of COVID-19 in schools and their contribution to population infections in Germany: A retrospective observational study using nationwide and regional health and education agency notification data. PLoS Med. 2022;19(12):e1003913. Published 2022 Dec 20. doi:10.1371/journal.pmed.1003913

Response:
Thank you for this insightful comment. In our analysis, when we refer to "school closures," we consider a complete suspension of all school-based contacts. However, we acknowledge that in reality, as you pointed out, school closures may not fully eliminate school-related contacts, especially for special needs children or children of key workers. In a new section, we capture these spectrum of measures by varying the school mitigation not just for school closures but also for school openings. The analysis is explained in the “Sensitivity analyses” subsection (Pages 7-8, Lines 197-201) and its results are discussed in the “Robustness of results” subsection (Page 11, Lines 274-282) of the main text.

Comment 3

Building on the previous two comments: The authors use as their first scenario the omission of school closures from the pandemic response. Within their model, in-school contacts remain unchanged (Equation 3). This is a rather theoretical scenario. The authors state that it is reasonable to assume that measures as were observed after the summer vacations were unlikely to be implemented immediately as schools remained open. This is a reasonable assumption, however, a complete continuation of within-school contacts at the prepandemic level is equally unlikely.

This assumption is justified by the abrupt onset of the pandemic in early 2020, which likely left insufficient time to implement effective within-school measures. As a result, the marginal benefit of closing schools was greater during this period. We confirmed this mechanism in sensitivity analyses, which tested the effect of applying a mitigation scaling factor to school contacts in Scenario 1. Introducing a reduction in school-based contacts attenuated the benefit of school closures, supporting the interpretation above.

If we look at the countries that refrained from imposing school closures during the first wave in Europe, namely Sweden and Iceland (and Belarus), we can see that open schools did not equal open schools at the pre-pandemic level. In Sweden, where primary schools remained open, the minimum package consisted of stay-at-home orders for symptomatic students, hand hygiene and intensified surface cleaning, adaptions to schedules to avoid crowding, recommendations to hold classes outdoors, and cancelling of e.g. choirs and assemblies. In Iceland, cohorting and capping of class sizes, distancing for teachers, hand and other hygiene measures, regular symptom checks, and cancellations of high-risk congregations were instituted. Further measures were optional and left to the schools. Within the sensitivity analysis, this is considered (Scenario 1.2).
However, from a real-world perspective, it seems more likely that a scenario like Scenario 1.2 should have been part of the main analysis of this paper and a scenario without any reduction in school contacts would have formed a sensitivity analysis. This would have avoided the need to include the assumption that other measures would not have been relaxed until the summer vacation as part of the Scenario 1 to make it more plausible.
I would thus recommend instead an adjusted Scenario 1.2 as the main focus of the analysis, or otherwise would require further justification of choosing Scenario 1 as the main output.

Response:
Thank you for this valuable comment. We agree with the points raised regarding the realism of Scenario 1. In response, we have revised our Scenario 1 to better reflect real-world dynamics. Now, Scenario 1 incorporates the reopening of schools during the first lockdown, while keeping the rest of the measures consistent with the baseline scenario. This includes the assumption that other measures, such as relaxation of restrictions, continued in a similar manner as before. This adjustment aims to provide a more realistic and meaningful comparison to other scenarios, as suggested, and we

believe it strengthens the robustness of our analysis. This is detailed further in the "Counterfactual scenarios vs baseline" (Page 5, Lines 122-124) and "Time-varying contact patterns" (Page 6, Lines 151-154) subsubsections of the main text. The previous Scenario 1 is still included in the sensitivity analyses (Supplementary Section 15, Scenario 1.1).

Comment 4

The authors assume for their model different susceptibilities across age groups. The relative susceptibilities assumed for 3 age groups are based on one contact tracing study contacted in China during the early pandemic:

The relative susceptibilities of the 0-20 years age group, β1,2,3, and of the 20-60 years age group,
β4,5,6,7, had log-normal priors with medians 0.23 and 0.64, respectively, and a scale parameter of 0.5 [39]. The relative susceptibility of the 60+ age group was set to β8,9,10 ≡ 1

The estimates of relative susceptibility, in particular for children, seem to be rather extreme estimates compared to other literature including high-quality meta-analysis, e.g. Viner et al.

I would like to see a justification of this parameter and the specific study used since the literature on this is rather extensive and the chosen value appears to be rather an outlier compared to other examples.

Viner RM, Mytton OT, Bonell C, et al. Susceptibility to SARS-CoV-2 Infection Among Children and Adolescents Compared With Adults: A Systematic Review and Meta-analysis JAMA Pediatr. 2021;175(2):143-156. doi:10.1001/jamapediatrics.2020.4573

Response:
Thank you for raising this important point and the opportunity to clarify our approach. We do not use fixed values for age-specific susceptibilities. Rather, the values quoted in the manuscript represent prior distributions used in the Bayesian evidence synthesis, not fixed assumptions. To avoid confusion, we have revised the methods to state *"had log-normal prior distributions"* instead of *"had log-normal priors"* (Supplementary material, Page 10, Line 644). In the revised text, we have further explained the following.

Supplementary material, Data and methods, Page 7, Lines 173-175:
"[...] Posterior medians and 95% credible interval (CrI) of the estimated model parameters are provided in Supplementary Table S1 and the corresponding posterior distributions are shown in Supplementary Figures S6-S8. [...]"

The posterior median susceptibility of the 0-20 years age group was 0.48 [95% CrI 0.31 - 0.68] (Supplementary Table S1 and Supplementary Figure S8). This estimate is consistent with the systematic review and meta-analysis by Viner et al. [JAMA Pediatr. 2021], which reported that children and adolescents have lower susceptibility to SARS-CoV-2, with an odds ratio of 0.56 for being an infected contact compared with adults. We now highlight this concordance in the "Epidemiological impact" subsection of the main text.

Main text, Results, Page 8, Lines 225-228:

“[...] However, in relative terms, youth experienced the largest increase in hospitalizations compared to baseline. This occurred despite their susceptibility being estimated at approximately half that of adults over 20 years (Supplementary Table S1), consistent with findings of the systematic review and meta-analysis in Viner et al. [54]. [...]”

Comment 5

This comment concerns the stated assumption that hospitalised cases did not contribute to pandemic dynamics:

“Second, hospitalized individuals were assumed to be non-infectious, as they constituted a small proportion of the total infectious population and were subject to various personal protective measures in hospitals.”

In fact, during the first wave estimate range around 20% of recorded hospitalisations for SARS-CoV-2 were nosocomial in origin (data mainly from UK, for Portugal I could only identify one article - see below). Even though their contribution to total infections and seropravelence was estimated to be low, their contribution to hospitalisations (one central outcome of the paper) was not negligible. I would like to see a justification or sensitivity analysis to show that an inclusion of hospitalised cases towards further hospitalisations would not alter the outcomes of the analysis.

Knight GM, Pham TM, Stimson J, et al. The contribution of hospital-acquired infections to the COVID-19 epidemic in England in the first half of 2020. BMC Infect Dis. 2022;22(1):556. Published 2022 Jun 18. doi:10.1186/s12879-022-07490-4

Bhattacharya A, Collin SM, Stimson J, et al. Healthcare-associated COVID-19 in England: A national data linkage study. J Infect. 2021;83(5):565-572. doi:10.1016/j.jinf.2021.08.039

For Portugal:
https://www.paginaum.pt/2022/01/12/morreram-mais-de-1-300-pessoas-que-apanharam-covid-19-nos-hospitais-durante-a-primeira-fase-da-pandemia

Response:
Thank you for this important comment and for highlighting relevant literature on nosocomial transmission [Knight et al.; Bhattacharya et al.], which we cited in the revised manuscript.

We would like to clarify how hospitalizations and nosocomial transmission are represented in our model. First, the hospitalization data used for model fitting cover all new COVID-19 diagnoses in hospitalized patients, including nosocomial cases, i.e., infections acquired in healthcare settings. Second, the susceptible, exposed, infectious, and recovered compartments represent the entire population, including individuals hospitalized for non-COVID conditions, hospital staff and visitors. Thus, nosocomial transmission originating from infectious staff, visitors, or patients hospitalized for other conditions is implicitly captured.

By contrast, individuals in the hospitalized-with-COVID compartment (H) were assumed to be non-infectious. This reflects both their small relative contribution to the total infectious population and the stricter infection control measures applied once a

COVID-19 diagnosis was established. For example, by 28 May 2020 there were 5120 [95% CrI 4845–5390] hospitalizations with COVID-19 in Portugal, compared to an estimated 358493 [95% CrI 273805–435534] total infections (3.60% [95% CrI 2.70–4.30] seroprevalence in a population of 10286300). Thus, individuals hospitalized with COVID-19 represented approximately 1% of all infectious individuals at that time. In our model, their infectivity was therefore approximated as negligible.

We emphasize that our statement "hospitalized individuals were assumed to be non-infectious" refers specifically to patients already hospitalized with a COVID-19 diagnosis, not to the origin of hospitalizations (community vs. nosocomial). Our focus is on school vs. non-school transmission at the population level, and we therefore did not stratify further by within-hospital dynamics, which were outside the scope of our study.

To make this clearer, we have revised the text as follows:

Main text, Data and methods, Page 3, Lines 70-73:
"[...] The national age-specific COVID-19 hospital admission data from 2 March 2020 to 15 January 2021 were provided by the Central Administration of the Health System and the Shared Services of the Ministry of Health. The hospitalization data cover all new COVID-19 diagnoses in hospitalized patients, including nosocomial cases, i.e., infections acquired in healthcare settings. [...]"

Main text, Data and methods, Page 4, Lines 99-105:
"[...] Second, we assumed that individuals hospitalized with COVID-19 were non-infectious, as they represented a small proportion of the total infectious population and were subject to strict infection control measures once their COVID-19 diagnosis was established. Third, since the model was fitted to hospitalization incidence, intra-hospital dynamics and disease-related deaths were not explicitly modeled. Nonetheless, nosocomial transmission of SARS-CoV-2 [48,49] is implicitly captured, as the susceptible, exposed, infectious, and recovered compartments include individuals who may be hospitalized for non-COVID conditions. [...]"

Comment 6

It is laudable that the authors included seroprevalence data as a benchmark for their model calibration, despite the large uncertainty around seroprevalence estimates within the cited "First National Serological Survey (ISNCOVID-19)" conducted in Portugal between 21 May 2020 and 8 July 2020. The authors state:

Therefore, we also incorporated seroprevalence data, which was sampled around day tsero = 93 post 26 February 2020. We neglected the waning of immunity and assumed the probability of a positive serological test for a randomly selected individual in age group k to be pk ≡ (1 − Sk(tsero)/Nk).

t(sero) would in this case fall on the 29th of May 2020, which is skewed toward the beginning of the sampling period. Taking into account the diagnostic test performance for detection of seroprevalence assumed by the serological survey at 10 days after symptom onset, we reach the beginning of June as the cut-off date. Due to the low levels of incidence during the summer period in Portugal, I would have liked to see a sensitivity analysis where the seroprevalence data is assumed not to present just the first wave until the summer vacation, but also an analysis where this data is assumed to cover also the summer vacation for school children to see whether this would change effects of school closures.

Response:

Thank you for this thoughtful suggestion. To address this, we performed a sensitivity analysis by fitting the data on two different dates: the last day of the survey period and the last day before the second pandemic wave. The model fits were similar to the original ones, and the effects on transmission scenarios were qualitatively similar across these dates, indicating that the findings are robust to the timing of the seroprevalence data. The details of the methods and the results are in the "Sensitivity analyses" subsection (Page 8, Lines 202-205) and "Robustness of results" subsection (Page 11, Lines 283-285), respectively.

Comment 7

The authors describe a delay within their model regarding the uptake and thus effect of changes:

The rates Ki, governing the speed at which policy changes took effect, were given Exp(1) priors, implying that the uptake of measures took around 6 days

It is not quite clear how this assumption is justified. Rather, a different effect has been observed especially for the early phases of the pandemic, when uptake of measures was often linked to their public announcement rather than their date of coming into effect.
Germany's largest analysis of pandemic measures found that the effect of measures on R is measurable 2 days prior to the measure actually coming into effect due to anticipatory behaviour changes in the population – an effect more pronounced towards the early phases of the pandemic.

Robert Koch Institut. Wirksamkeit und Wirkung von anti-epidemischen Maßnahmen auf die COVID-19-Pandemie in Deutschland (StopptCOVID-Studie). Berlin: Robert Koch Institut; 2023.

A similar effect is described for other countries, including Spain:
Santamaría L, Hortal J. Chasing the ghost of infection past: identifying thresholds of change during the COVID-19 infection in Spain. Epidemiology and Infection. 2020;148:e282. doi:10.1017/S0950268820002782

Response:
Thank you for this insightful comment and for drawing our attention to important evidence on anticipatory behavioral changes [Robert Koch Institut, 2023; Santamaría et al., 2020]. To clarify our modelling choice and implications, we have revised the methods section as follows.

Main Text, Data and methods, Page 5, Lines 111-119:
"[...] In our model, we do not explicitly specify the dates when measures were announced or formally enacted. Instead, after an initial period without control measures, we estimate mid-point transitions between five empirically defined behavioral regimes throughout 2020: 0) the first lockdown, during which schools were closed in late March; 1) a relaxation of restrictions following the first lockdown, beginning in late May; 2) further easing of measures, including the reopening of schools after the summer holidays in late August; 3) the second lockdown in early November; and 4) a subsequent relaxation of measures in late December. The timing and speed of transitions between these regimes were estimated from fitting the model to the data. Supplementary Figure S1 shows that the estimated transition mid-points are in line with the measures implemented in Portugal (see also Table S2 for a description of these measures). [...]"

We apologize to the Reviewer for the confusion on K_i. The prior rates K_i, mentioned by the Reviewer, represent prior assumptions about the speed of these transitions. We assigned Exponential(1) prior distributions to these rates because they are weakly informative. The 6-day transition time is simply the result of calculating the change in uptake from 5% to 95% at the prior's mean rate ($E[K_i]= 1$) i.e., $2 \log(1/0.05 - 1) = 5.88 \approx 6$ days. This, however, is not a hard prior constraint. The prior's true implication is a broad 95% credible interval for the uptake time, spanning [1.6,176] days. Using this vague prior ensured that our posterior estimates were driven by the observed data. The final rates (Supplementary Figure S6) are consistent with observations reported in the literature [Robert Koch Institut, 2023; Santamaría et al., 2020].

Supplementary material, Parameter estimation, Page 10, Lines 652-655:
"[...] Finally, the rates K_i, determining how quickly behavioral changes occurred, were given Exp(1) priors. While its mean $E(K_i)=1$ corresponds to a 6-day transition time for 5% to 95% uptake, its broad 95% credible interval of [1.6, 176] days confirms it is weakly informative, thus allowing the posterior estimates to be driven by the observed data and aligning with transition ranges reported in the literature [35,36]."

We would like to add that the way how we model transitions between different behavioral regimes is novel in the literature (see our references [Rozhnova et al., 2021; Viana et al., 2021]) and it has certain advantages, because it estimates the overall effect of measures rather than looking at various factors such as when measures are announced, when they are enacted, how people adjust behavior in between. This method is especially convenient when there are few contact matrices available.

Comment 8

The concurrence of introduction of measures during the pandemic often poses challenges to the differentiation of their effects. It is not quite clear how, within the setup of the study, where all interventions other than school closures were subsumed into a single parameter denoting "non-school contacts", this complexity would have been captured and the problem at hand resolved. I would like to see a more detailed overview of the measures introduced in Portugal during the course of the study period (this overview could be placed in the supplementary material) to discern the relative political importance of school measures at the time. Furthermore, I would like to see a further discussion of the issue of concurrence and differentiation of effect, and how it was resolved in the design of the model to isolate the specific effect of school closures from other concurrent measures both within the first lockdown as well as within the second pandemic wave as measures increasingly stringent measures were introduced.

Response:
Thank you for these thoughtful comments. A detailed overview of the measures introduced in Portugal during the study period is now provided in the "Measures introduced in Portugal" section of the Supplementary material. This overview offers context on the specific political and public health measures implemented during the pandemic.

Regarding the differentiation of effects from various concurrent measures, our objective was not to isolate the effects of individual interventions, but rather to assess the impact of school closures specifically. As the contact data do not allow to differentiate the effects of school and non-school measures at a finer level, we ask for the referee's understanding in not pursuing this differentiation further.

Our focus has been on quantifying the impact of school closures. To explore this further, we have added new results that vary the extent of school closures (as a percentage) to show how different levels of school closure influence the transmission dynamics and the overall outcomes. The analysis is explained in the "Sensitivity analyses" subsection (Pages 7-8, Lines 197-201) and its results are discussed in the "Robustness of results" subsection (Page 11, Lines 274-282) of the main text.

Reviewer 2

The authors infer the impact of school closures in Portugal and use this to explore counterfactual scenarios. The methodology is robust and the conclusions supported by the analysis. Also all steps necessary to ensure reproducibility seem to have been taken. I have only few minor remarks:

Response:
Thank you for the positive assessment of our work. We are pleased that our approach and reproducibility measures were well received.

** p.6 line 4: is 1-zeta the reduction, or zeta?, same question of zeta_s, try to be clearer here.*

Response:
Thank you for this helpful comment. We agree that the use of 1-ζ could be confusing. To clarify, ζ represents the mitigation factor for contacts during the first lockdown, and its complement, 1-ζ, represents the reduction in contact transmission. To avoid confusion, we have revised the manuscript to consistently use ζ.

Main text, Data and methods, Page 6, Lines 138-141:
"[...] With the parameter $0 \le \zeta \le 1$, we also quantified the mitigation factor for transmission-relevant contacts during the first lockdown due to protective measures in the general population not explicitly modeled, such as mask-wearing, refraining from shaking hands, physical distancing, and remote working. [...]"

** Equation 6 and 7, write the variable for which you model the probability in parentheses after the P: e.g. P_{hosp}(h_{k,i}) = ... and P_sero(l_k) = ... Otherwise one has to guess it.*

Response:
Thank you for this helpful suggestion. We have changed Equations 6 and 7 (now in the Supplementary material, Page 10, Lines 664 and 676) to indicate in parentheses after P which parameters and variables the probabilities are dependent on, and have updated Equation 8 (now in the Supplementary material, Page 11, Line 678).

** In equation 6 and 7 and 8, don't use P, but L_{hosp} or l_p or f, it isn't a probability, but a probability density/likelihood. It gives the wrong impression and might lead to misunderstandings if a P is used here. Probabilities are the integrals of the the probability densities.*

Response:
Thank you for this comment. We agree that distinguishing between probabilities and probability densities is important. In this case, we deliberately used the symbol P because $h_{k,i}$ and ℓ_k are discrete observations. They correspond to the integer number of hospitalizations observed in a day, and the number of individuals testing seropositive in a cohort of size L_k. Therefore, P_{hosp} and P_{sero} are genuine probabilities, not densities. However, due to the fact that these objects define the likelihood function, the original text incorrectly referred to them as *"likelihood"*. In the revised text, we have corrected this terminology to avoid confusion (Supplementary material, Pages 10-11, Lines 662, 674, and 677).

** Report how well the chains converged, e.g. via R-hat.*

Response:
In our previous paper [Viana et al., 2021], we reported: *"Convergence was assessed with the Gelman-Rubin Rhat-statistic, which was close to 1 for all parameters."* To make this a bit more precise, we re-fitted the model so that we could run Stan diagnostics. As fitting this model is rather time-consuming, we ran 6 parallel chains, each with 300 warmup samples and 500 iterations. This resulted in an $\widehat{R} < 1.01$ for all parameters, which we now indicate in the "Dynamical and observation model" subsection of the Supplementary material.

Supplementary material, Parameter estimation, Page 11, Lines 680-681:
"[...] MCMC convergence was assessed with the $\widehat{R}$ statistic (using the R posterior package). The value of $\widehat{R}$ was below 1.01 for all parameters. For other details, see Viana et al. [30]."

** Fig. 2+3: Add a legend to explain what the blue and orange lines/regions are, so that one doesn't have to read the caption.*

Response:
Thank you for this suggestion. We have added legends to Figures 3 and 4 (previously Figures 2 and 3) in the main text to clarify the meaning of the blue and orange lines/regions. For consistency, legends have also been added to the new Figure 5 and the figures in the Supplementary material.

** Results section, epidemiological impact: reduce the number of significant digits you report, if the credible interval is much larger, it gives a misleading impression of precision.*

Response:

We apologize for any confusion. The “,” in the text was used as a thousands separator (e.g., for hospitalizations), which may have been misinterpreted. We have now removed this separator. In addition, as suggested, we have reduced the number of significant digits reported for seroprevalence and percent changes in hospitalizations in the “Epidemiological impact” subsection.

Reviewer 3

The manuscript by Canfora et al., presents a retrospective study on the impact of school closure and reopening on SARS-CoV-2 transmission in Portugal in the first year of the pandemic, prior to mass vaccination.

In this work, the authors use a previously developed age-stratified compartmental model fitted to hospitalization and seroprevalence data to reconstruct SARS-CoV-2 transmissibility conditions during the spring and autumn of 2020. By producing counterfactual scenarios, they assess the impact of school closures on hospitalization burden, concluding that decisions on schools are influenced both by school-based mitigation measures and the intensity of transmission in the community captured by the time-varying reproductive number.

Response:
Thank you for the accurate summary and for highlighting the key aspects of our study, including the use of an age-stratified model and the influence of both school-based measures and community transmission on school policy decisions.

Age-specific contributions to transmission are also estimated through elasticity analysis, offering insights into the role of different age groups. However, these results do not appear to be fully integrated into the narrative (for example, they are not explicitly mentioned in the abstract) despite representing a potentially valuable contribution to the study.

Response:
Thank you for emphasizing the importance of the elasticity analysis. We have revised the abstract to include the key age-specific findings related to elasticities, ensuring these results are fully integrated into the manuscript's narrative.

Main text, Abstract, Page 2, Lines 9-13:
"[…] Elasticity analyses were used to quantify the contributions of different age groups to transmission. Our results show that the pandemic in 2020 was primarily driven by adults. We found that keeping schools open during the first lockdown would have led to hospitalizations rising far beyond observed levels, unless schools contacts would reach at most 50% of their pre-pandemic levels. […]"

Although the manuscript is built on a validated modeling approach - with limitations appropriately acknowledged and a comprehensive sensitivity analysis - I do have comments regarding some assumptions and scenario choices. In conclusion, I consider this work suitable for publication with some revisions. Below, I provide a point-by-point list.

Response:
Thank you for recognizing the contributions of modeling approach and sensitivity analyses. Below, we address the comments on assumptions and scenario choices point by point.

Introduction

1. Relevant transmission modeling works on school and community interventions to safely reopening schools after the first lockdown - such as those in France - are not explicitly mentioned. I would suggest incorporating a more in-depth review of the literature to better contextualize the complexity of decision-making around school reopening and closing during the pandemic.

Response:
Thank you for this valuable suggestion. We have expanded our introduction to provide a more comprehensive overview of international modeling studies on the impact of school-based measures and community interventions aimed at safely reopening schools after the first lockdown. In particular, we have incorporated around fifteen additional references, including works from France, Germany, the Netherlands, Belgium, China, Canada, the USA, the UK, and Australia. These studies adopt various approaches, including agent-based, metapopulation and deterministic modelling, enabling better understanding of the complexities involved in decision-making around school closures and reopenings during the pandemic. We believe that these additions strengthen the basis of our work, situating the Portuguese case within a richer international context.

Main text, Introduction, Page 2-3, Lines 40-45:
"To understand this balance, it is necessary to look beyond the school environment itself and consider the wider societal impact of these measures. In the international literature, many studies evaluate interventions implemented within schools, but rarely quantify their impact on broader epidemiological outcomes at the national level [19-25]. For example, research in France examined protocols for the safe reopening of schools, often using agent-based models to capture detailed in-school transmission dynamics [26-28]. While such approaches provide valuable insights into mitigation within the school setting, they are limited in their ability to capture national level transmission."

Methods

1. Could you please clarify why you chose to stratify the population into ten age classes instead of five as those in the seroprevalence data? This would have allowed to sum over the same age groups in the loglikelihood computation without aggregating a posteriori. Additionally, I wonder whether larger age classes might improve the model fit, especially considering that hospitalization data for children in the [0–5) and [5–10) age groups appear to be sparser than for other age categories.

Response:
Thank you for this comment. The choice of age classes represents a balance between computational feasibility and the need to accurately model age-dependent contacts. Aggregating smaller age classes into larger ones can result in a loss of detail regarding how children interact with each other, as well as with teachers and parents. Since the main focus of our study was the effect of school closures, we ensured that such contacts were modeled in sufficient detail. For example, school-age children have frequent contacts with individuals aged 30–40, but not with those aged 20–30. Reducing the number of age classes while maintaining detail in younger age groups would still misalign the classes with the available seroprevalence data. Moreover, the oldest age groups are important due to their higher hospitalization rates, so aggregating these classes could reduce our ability to assess the impact of school closures on hospitalizations. We have added the following sentence to the Data section to clarify this choice.

Main Text, Data and methods, Page 4, Lines 91-93:
"[...] These specific age groups were chosen (i) to ensure a sufficiently fine age stratification to capture the contacts of school-aged children with adults; (ii) to align with age-dependent contact patterns, risk of hospitalization, and available seroprevalence data. [...]"

2. Could you specify how the goodness of fit was assessed?

Response:
Goodness of fit was assessed in two ways. First, we used posterior predictive checks that confirmed that simulated data from the model posterior replicates original data (Supplementary Figure S10-S11). Second, we assessed MCMC convergence with the $\widehat{R}$ statistic. We added a sentence (Lines 680-681) and referred to relevant figures showing age-dependent fits, as well as projections for alternative scenarios.

Supplementary material, Parameter estimation, Page 11, Lines 680-681:
"[...] MCMC convergence was assessed with the $\widehat{R}$ statistic (using the R posterior package). The value of $\widehat{R}$ was below 1.01 for all parameters. For other details, see Viana et al. [30]."

3. Related to 1., you calibrate age-specific susceptibility yet this does not vary among individuals aged 0 to 19 years. This assumption might reduce the need for a finer age resolution. At the same time, adolescents may exhibit more similar epidemiological characteristics to adults. Have you consulted the recent literature on age-specific susceptibility?

Response:
Thank you for this comment. As noted in our response to Reviewer 1, Comment 4, our estimates are consistent with the systematic review and meta-analysis by Viner et al. [JAMA Pediatr. 2021], which indicates that children have a relative susceptibility of approximately 0.5, while adults and the elderly have similar susceptibility. We adopted this coarse-grained resolution to avoid potential identifiability issues in the model, ensuring robust and stable parameter estimation and replication of data with the model estimates.

4. In Scenario 1, you assume that school contact rates would have remained at pre-pandemic levels, despite the presence of strong restrictions in other settings. Retrospectively, it might be more informative to explore alternative scenarios: What would have happened if the same mitigation measures of scenario 2 had been applied during the first lockdown (currently in the SI)? What kind of interventions could have allowed schools to remain open while keeping hospitalizations at reasonable levels? This would better support pandemic preparedness.

Response:
Thank you for this insightful comment. We added analyses where we explored varying school opening mitigation levels in order to assess their effect on cumulative hospitalizations, seroprevalence, reproduction number, and elasticities (Figure 5 and Supplementary Figure S19). As expected, the resulting values lie between those of the baseline and scenarios, showing clear increasing or decreasing trends depending on the level of school opening or closure.

With respect to the question of what kind of interventions could have allowed schools to remain open while keeping hospitalizations at reasonable levels, our dataset does not allow us to disentangle non-school contacts into specific categories. This limitation prevents us from reliably modeling such hypothetical interventions. For this reason, our focus remains on evaluating the consequences of the measures that were actually implemented, while the choice of which interventions should have been adopted falls within the domain of policymakers rather than our retrospective modeling approach.

5. The relaxation of restrictions following the first lockdown is not fully detailed in the manuscript. I suppose this refers to general measures such as mask-wearing, physical distancing, and refraining from handshakes, as mentioned in the section on Time-varying contact patterns. Does this also include remote working or other measures? Similarly, the protective measures implemented in schools - such as reduced classroom capacity, increased outdoor activities, staggered schedules, and enhanced cleaning protocols - are briefly mentioned in the introduction but not described elsewhere. Were these protective measures consistent across all school levels, or did they vary?

Response:
Thank you for pointing this out. The parameter ζ did also include other measures such as remote working. We revised the parameter explanation to mention this (Page 6, Lines 138-141).

Regarding school protective measures, many interventions were applied across all school levels but adapted to the specific needs of each. For example, outdoor activities were particularly promoted in primary schools where they could be more easily implemented, while secondary schools emphasized mask-wearing, staggered schedules, and reduced classroom capacity. In our model formulation, we assume that the estimated parameters u_j's for each regime would capture these intermediate behaviors, ranging between pre-pandemic and full lockdown. Additionally, these interventions are cited in the "Measures introduced in Portugal" section of the Supplementary material.

6. I would recommend adding a brief explanation of the elasticity matrix and its interpretation, especially for readers who may be less familiar with this concept.

Response:
Thank you very much for this suggestion. We have restructured the "Reproduction number and elasticity analysis" subsection of the "Data and methods" section to help readers understand the concept without the complexity of the analysis. To further aid understanding, we have emphasized the elasticities using the adult-youth structure. All the other details are moved to the Supplementary material.

Main text, Data and methods, Reproduction number and elasticity analysis, Page 7, Lines 183-188:
"While R(t) quantifies the transmission potential at a given time, it is the elasticity analysis that provides insight into the drivers of changes in R(t) [52,53]. Using this analysis, we calculated cumulative elasticities, $e_k(t)$, which can be interpreted as the relative contribution of each age group k=1,...,n to R(t). In our analysis, we further aggregate cumulative elasticities into those of youth (0-20 years), $e_{[0,20)}(t)$, and adults (20+ years), $e_{20+}(t)=1-e_{[0,20)}(t)$. Further details on the reproduction number calculation and elasticity analyses are given in Supplementary Material, Section 8."

7. The notation for the regimes in the text is not consistent (first labels from i) to v), then numbers ranging from 0 to 5). Standardizing the notation may facilitate the reading.

Response:
Thank you for pointing this out. We agree that the notation was inconsistent and could be confusing for the reader. In the revised version, we have standardized the notation for the regimes throughout the text, figures, and tables. We now consistently use the numbers ranging from 0 to 4. In particular, this is in the description of regimes in the "Pandemic timeline" subsubsection (Page 5, Lines 113-116) with a mentioned indexing notation.

Main text, Data and methods, Page 5, Lines 119-120:
"[...] The behavioral regimes were denoted by the index j=0,1,2,3,4 in the analysis that follows."

8. Figure 1 includes elements - such as the 'midpoint' or scenario 1 and 2 - that have not yet been introduced in the text when it is first cited. Providing this context beforehand may improve comprehension.

Response:
Thank you very much for this feedback. We carefully considered the suggestion and clarified the text to better guide the reader. To further improve comprehension, the "midpoint" estimations are now well-described in the "Pandemic timeline" subsection of the main text, where the figure is first cited. Similarly, the Scenario 1 and Scenario 2 annotations are detailed in the succeeding "Counterfactual scenarios vs baseline" subsection. The figure itself remains essential to illustrate the division between the two periods and the select counterfactual measures.

Main text, Data and methods, Page 5, Lines 111-113:
"[...] In our model, we do not explicitly specify the dates when measures were announced or formally enacted. Instead, after an initial period without control measures, we estimate mid-point transitions between five empirically defined behavioral regimes throughout 2020: [...]"

Main text, Data and methods, Page 5, Lines 116-119:
"[...] The timing and speed of transitions between these regimes were estimated from fitting the model to the data. Supplementary Figure S1 shows that the estimated transition mid-points are in line with the measures implemented in Portugal (see also Table S2 for a description of these measures). [...]"

Main text, Data and methods, Page 5, Lines 122 and 127:
"We retrospectively evaluated two counterfactual scenarios (Figure 2): [...], both respectively shown by the yellow-shaded regions in Figure 2."

9. In Figure S3, the average number of contacts per day for children aged 5-10 years is 0.8 for the pre-pandemic period, which seems low. If I am interpreting the figure correctly, shouldn't this value correspond to the red line shown in Figure S1F for the 5-10 age group?

Response:
Thank you for your careful review and for pointing out this discrepancy. We agree that the value for children aged 5-10 in Supplementary Figure S2F (previously Supplementary Figure S1F) is correct. We have since recalculated the average contacts and have generated a new Supplementary Figure S3 (previously Supplementary Figure S2) that correctly shows the average number of contacts, which now agrees with the correct values as you noted. The succeeding figures for each of the ten age groups, Supplementary Figures S4 and S5, are also readjusted.

Results

1. Before reaching the results section, I did not realize that the restrictions from the first lockdown were not relaxed in scenario 1. This should be more clearly stated also in the Methods section describing the scenarios.

Response:
Thank you for this observation. We reformulated Scenario 1 which is now adjusted to account for real-world pandemic dynamics, following your suggestions and that of Reviewer 1. The previous version of Scenario 1 is still included as a sensitivity analysis.

Main text, Data and methods, Page 6, Lines 151-154:
"To formulate the counterfactual scenarios in Scenario 1 covering the period from 26 February to 15 July 2020, we assumed that schools stayed open during the first lockdown at full capacity - with pre-pandemic contact levels - until the first relaxation of measures. Analysis of contact structures with varying school mitigation levels or prolonged non-school lockdown measures are presented in the Sensitivity analyses section below."

2. The frequent changes in age group throughout the results can be confusing. Additionally, I found a bit inconsistent that results put in the Supplementary Information were extensively discussed in the main text. While I understand that aggregating into two age classes reflects how decisions were made by policymakers at the time, I would encourage reconsidering this choice if needed.

Response:
Thank you for this valuable comment. We have carefully revised the manuscript to ensure consistency and clarity. In particular, we standardized the terminology to consistently use "youth" and "adults" throughout the main text and figures. While we refer to some intermediate age groups when necessary in the main text, their specific results are presented in the Supplementary material.

3. It would be helpful to include a legend in Figure 2, as it is currently missing. Clarifying the color-coding used for the lines would help reading the plot. I would also mention in the caption that you are showing absolute number and not new daily hospital admission per 100,000.

Response:
Thank you for this helpful comment. We have addressed it by adding a legend to Figure 3 (previously Figure 2) to clarify the color-coding, and by revising the caption to clearly state that absolute numbers are shown. Legends are also shown for other figures such

as Figure 4 (previously Figure 3), Figure 5 and most of the figures in the Supplementary material.

4. The section or figure titles could be revised to more clearly reflect the main findings. For example, the title of the first section could be more informative by explicitly stating the epidemiological impact of school closures on daily hospitalizations. Similarly, for impact of scenarios on hospitalizations in Figure 2. Additionally, without the use of descriptive keywords, reader has to repeatedly recall the specific definitions of Scenario 1 and Scenario 2. Introducing concise labels or keywords may improve readability.

Response:
Thank you for the suggestion. We have revised the section and figure titles to make them more informative and reflective of the main findings. This is reflected in Page 8, Line 214; Page 9, Line 234; Page 10, Line 247; and Page 11, Line 255.

5. 'Impact of school closures: Scenario 1 versus Scenario 2' seems more a discussion point than a result.

Response:
We appreciate the Reviewer's helpful observation regarding the placement of this subsection. We have revised the "Results" section to entirely remove this content, as its main point is fully addressed within the "Discussion" section (Pages 12-13, Lines 300-311).

Discussion

1. The authors cite as a strength the fact that age-specific mixing patterns were inferred directly from the data rather than imposed a priori. However, this approach cannot be validated against empirical contact data due to their lack, limiting the ability to assess the accuracy of the inferred patterns.

Response:
Thank you for this comment. We acknowledge this limitation. In many modelling studies, age-specific contact patterns are often taken directly from available data or adjusted by simple scaling factors. In our study, these patterns are inferred directly from the data, which we consider a strength because the model is able to reproduce the observed dynamics accurately. This successful reproduction provides confidence in the validity of the model, even though, we agree that, due to the lack of empirical contact data, we cannot formally validate the inferred patterns, and this remains a limitation of our approach. We also underlined this in our discussion.

Main text, Discussion, Page 13, Lines 328-330:
"[...] This approach ensured model validity, as it successfully reproduces observed dynamics; however, due to the lack of empirical contact data, the inferred contacts cannot be formally validated, which represents a limitation of our study."

General

The main text would benefit from some shortening to better highlight the most important points (the current word count appears to exceed 7,000, if I'm not mistaken).

Response:
We appreciate the Reviewer's suggestion. We carefully revised the manuscript and shortened the main text, which now counts 5494 words.

Equations that have already been presented in the previous work could be moved to the Supplementary Information. Including a diagram to illustrate disease progression and rates - distinguishing between parameters informed by the literature and those calibrated through model fitting - could be an alternative.

Response:
Thank you for this helpful suggestion. In the revised version, we have moved the equations to the Supplementary material and replaced them with a diagram illustrating the disease progression and rates (Figure 1). In addition, we already provided a comprehensive table listing all parameters in the Supplementary material (Table S1), which clearly specifies whether they were derived from the literature or estimated through model fitting.

Response to reviewers’ comments

Reviewer 2

Thank you for your revision and answers. From my side, it can be published as is.

Response:
We are grateful to the reviewer for the positive assessment of our manuscript and for the helpful comments provided during the review process. We are pleased that the revisions have satisfactorily addressed all the issues raised.

Reviewer 3

I thank the authors for addressing the points I previously raised. However, I believe that the writing would still benefit from some refinement.

Abstract

The results in the abstract are currently presented beginning with the elasticity analysis, whereas in the main text this section appears after the scenario analysis. I suggest to align the order for greater consistency.

Response:
Thank you for this suggestion. We agree that aligning the order of results in the abstract with their presentation in the main text improves the logical flow. We have rearranged the abstract to present the scenario analyses first, followed by the elasticity analysis results.

Abstract, Page 2, Lines 14-19:
"We found that keeping schools open during the first lockdown would have led to hospitalizations rising far beyond observed levels, unless school contacts reached at most 50% of their pre-pandemic levels. However, closing schools after the summer holidays, when adult-driven community transmission was already high, would not have been sufficient to prevent the autumn wave of hospitalizations in 2020, though it could have mitigated its impact. Elasticity analyses quantified the age group contributions to transmission, showing that adults primarily drove the 2020 pandemic."

Introduction

The additions to the introduction do not sufficiently highlight the differences across the cited studies. Among the references included, it remains unclear how each study differs in its approach and main findings relative to the authors' research questions. For instance, reference [24] also introduces household risk, while references [20], [21], and [22] focus more on within-school transmission in later phases of the pandemic rather than at its onset. In contrast, reference [23] is more explorative and do not use epidemiological data for calibration. Providing a clearer description of the specific contributions of each cited study would help readers better situate the current work within the existing literature.

Safe protocols for reopening schools were also assessed here: Di Domenico, L., Pullano, G., Sabbatini, C.E. et al. Modelling safe protocols for reopening schools during the COVID-19 pandemic in France. Nat Commun 12, 1073 (2021). https://doi.org/10.1038/s41467-021-21249-6

Response:
Thank you for this important comment. We agree that the previous version of the Introduction did not clearly distinguish between the different approaches and contributions of the cited studies.

To address this, we have revised the Introduction to explicitly structure the literature according to three main analytical approaches:

1. community-level studies evaluating the impact of school closures and reopenings on epidemiological outcomes,
2. mechanistic modeling studies incorporating age-structured contact patterns across schools, households, and the broader community, and
3. studies focusing on within-school transmission dynamics or mitigation strategies for safe reopening.

Within this framework, we now clarify the specific focus and limitations of each group of studies. For example, we highlight that some studies examine household transmission risk (e.g., Lessler et al.), others focus on within-school transmission during later phases of the pandemic (e.g., Lasser et al., Leng et al., Colosi et al.), while others are exploratory and not calibrated to epidemiological data (e.g., McGee et al.). Following the reviewer's suggestion, we have also included the study by Di Domenico et al. (2021), positioning it among modeling studies evaluating school reopening strategies.

To further improve clarity and transparency, we have added Supplementary Table 4, also reproduced below, summarizing the main characteristics of the cited studies, including their methodological approach, setting, and primary focus. The revised text is reported below.

Main text, Introduction, Page 3, Lines 50-52 and 58-62:

“[...] Moreover, while reducing school-based contacts likely curtailed in-school transmission, this effect may have been offset by increased household interactions, as highlighted by studies explicitly examining household transmission risk [18,19]. [...]”

[...] In the international literature, studies differ in their focus: some evaluate transmission dynamics within school settings, particularly during later phases of the pandemic [20-23], while others explore mitigation strategies for safe reopening, often using agent-based or exploratory models that are not always calibrated to epidemiological data [24-28]. While such approaches provide valuable insights into in-school transmission and intervention effectiveness, they are limited in their ability to capture national-level transmission patterns.”

#	Authors	Title	Journal, year, and country	About
7	Ken Rice et al.	Effect of school closures on mortality from coronavirus disease 2019: old and new predictions	The BMJ, 2020, UK	The study focuses on the United Kingdom (~70 million population), using data available as of March 2020, with simulations spanning 800 days from January 1, 2020, and intervention periods of approximately 3 months (91 days). Independent replication and validation of the Imperial College London individual-based model (IBMIC) using CovidSim to simulate virus spread and assess the impact of non-pharmaceutical interventions on mortality.
8	Ganna Rozhnova, et al.	Model-based evaluation of school- and non-school-related measures to control the COVID-19 pandemic	Nature Communications, 2021, The Netherlands	Mathematical modeling study using an age-structured deterministic SEIR model, fitted in a Bayesian framework to hospitalization and seroprevalence data from the first wave (Feb–Apr 2020), with scenarios for August and November 2020. The analysis focuses on contact patterns, distinguishing between school settings (primary and secondary schools) and non-school settings (households and community contacts).
9	Sebastian Walsh, et al.	Do school closures and school reopenings affect community transmission of COVID-19? A systematic review of observational studies	BMJ, 2021, multiple countries	Systematic review of observational (mainly ecological) studies assessing the impact of school closures and reopenings on SARS-CoV-2 transmission, including studies published in 2020–2021 with data up to January 2021. The analysis focuses on school settings (nurseries, primary and secondary schools) and their effects on community transmission outcomes (infection rates, hospitalizations, and mortality).
10	Hannah Littlecott, et al.	Measures implemented in the school setting to contain the COVID-19 pandemic	National Library of Medicine, 2024, multiple countries	Cochrane systematic review (update of a rapid review) including only real-world empirical studies, explicitly excluding modeling studies, with evidence up to February 2022. The analysis focuses on school settings (K–12) and their impact on students, staff, households, and the broader community.
11	Jasmina	Determining the	The	Stochastic agent-based modeling study

	Panovska-Griffiths, et al.	optimal strategy for reopening schools, the impact of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave in the UK: a modelling study	Lancet, 2020, UK	(using Covasim) to assess the effectiveness of testing, tracing, and isolation strategies in the context of school reopening, calibrated on UK data from January–June 2020 with projections through December 2021. The analysis is set in the United Kingdom and incorporates contact structures across households, schools, workplaces, and community settings.
12	Arden Baxter, et al.	Evaluating scenarios for school reopening under COVID19	BMC Public Health, 2022, USA (Georgia)	Agent-based simulation study projecting the impact of different school reopening strategies, covering the period from February 18 to November 24, 2020. Conducted in Georgia (USA), the analysis models interactions across households, K–12 schools, workplaces, and community settings.
13	Romain Ragonnet, et al.	Estimating the impact of school closures on the COVID-19 dynamics in 74 countries: A modelling analysis	PLoS Medicine, 2025, 74 countries	Semi-mechanistic mathematical modeling study using simulations to compare observed scenarios with counterfactual “schools open” scenarios over 2020–2022. Conducted across 74 countries, modeling contact structures in households, schools, workplaces, and community settings.
14	Benjamin Lee, et al.	Modeling the impact of school reopening on sars-cov-2 transmission using contact structure data from Shanghai.	BMC Public Health, 2020, China (Shanghai)	Age-structured mathematical modeling study (SIR/SEIR-type) assessing the impact of school reopening on the basic reproduction number (R_0), based on contact data from Shanghai (early 2020). The analysis focuses on age-specific contact patterns across schools, households, and community/social settings.
15	Steven Abrams, et al.	Modelling the early phase of the Belgian COVID-19 epidemic using a stochastic compartmental model and studying its implied future trajectories	Epidemics, 2021, Belgium	Age-structured stochastic SEIR modeling study, calibrated in a Bayesian (MCMC) framework using hospitalization, mortality, and seroprevalence data from Belgium (March–May 2020), with projections through December 2020. The analysis models interactions across households, schools, workplaces, transport, and leisure/community settings.

16	Mark M Dekker, et al.	Reducing societal impacts of SARS-CoV-2 interventions through subnational implementation	eLife, 2023, The Netherlands	Stochastic agent-based SEIR modeling study comparing the epidemiological and social impact of national versus subnational interventions during the first COVID-19 wave in the Netherlands (Feb 27–June 1, 2020). The analysis models interactions across 380 municipalities, distinguishing contacts in home, school, work, and community/other settings.
17	Pei Yuan, et al.	School and community reopening during the covid-19 pandemic: a mathematical modelling study.	Royal Society Open Science, 2022, Toronto	Deterministic compartmental modeling study (S-E-A-I-R-D) structured by age, household, and context, using data from July 31 to November 23, 2020 to assess the impact of school reopening. Conducted in Toronto (Canada), the analysis models interactions across households, schools, and community settings (including in-person work and remote work).
18	Kiesha Prem, et al.	The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study	National Library of Medicine, 2020, China (Wuhan)	Age-structured deterministic SEIR modeling study simulating the impact of physical distancing measures on virus transmission, with projections over one year (late 2019–end of 2020). The analysis uses contact matrices across households, schools, workplaces, and community (other) settings.
19	Lasser	Assessing the impact of SARS-CoV-2 prevention measures in Austrian schools using agent-based simulations and cluster tracing data.	Nature Communications, 2022, Austria	Agent-based epidemiological modeling study calibrated with real cluster-tracing data (616 clusters, weeks 36–45 of 2020) to assess the effectiveness of NPIs, with simulations adapted to the Delta variant scenario. The analysis focuses on Austrian school settings (primary, lower and upper secondary, with/without after-school care) and households, modeling contacts among students, teachers, and relatives.
20	Leng	Quantifying pupil-to-pupil SARS-CoV-2 transmission and the impact of lateral flow testing in English secondary schools.	Nature Communications, 2022, England	Stochastic individual-based modeling study, Bayesian-calibrated using PCR/LFT testing data and school absence records, covering August 31, 2020 to May 23, 2021. The analysis focuses on secondary schools in England, modeling interactions among students (close contacts, year groups) and case importation from the local community.

21	Elisabetta Colosi, et al.	Proactive vs. reactive COVID-19 screening in schools: lessons from experimental protocols in France during the Delta and Omicron waves	medRxiv, 2025, France	Stochastic agent-based modeling study using empirical contact data to compare weekly screening versus reactive strategies during the Delta (Nov–Dec 2021) and Omicron (Jan–Feb 2022) waves. The analysis focuses on primary school settings in France (ages 6–10), modeling within-class transmission and community-introduced cases.
22	McGee	Model-driven mitigation measures for reopening schools during the covid-19 pandemic.	PNAS, 2021, USA	Stochastic network-based SEIR modeling study simulating transmission dynamics and the effectiveness of mitigation measures over a 150-day school semester (parameters reflecting 2021 variants). The analysis focuses on primary and secondary schools, modeling interactions among students, teachers, and staff within classes, social groups, and households.
23	Lessler	Household covid-19 risk and in-person schooling	Science, 2021, USA	Statistical analysis of a large online survey (COVID-19 Symptom Survey via Facebook), using sampling weights and regression models to estimate adjusted odds ratios during Nov–Dec 2020 and Jan–Feb 2021. Conducted across the United States, the analysis focuses on households with in-person schooling (K–12) and teachers, assessing associated infection risks.
24	Heinsohn	Infection and transmission risks of COVID-19 in schools and their contribution to population infections in Germany: A retrospective observational study using nationwide and regional health and education agency notification data	PLOS Medicine, 2022, Germany	Retrospective observational study combining descriptive infection risk analysis, multiple linear regression to assess control measures, and SEIRS compartmental modeling, covering March 2020 to April 2022 (wild-type to Omicron). Conducted in Germany (national and regional data), the analysis focuses on schools (students and teachers), households, and community settings.
25	Elisabetta Colosi, et al.	Modelling COVID-19 in school settings to	Anaesthesia Critical	Stochastic agent-based modeling study using empirical contact data to assess the impact of different control protocols over 90-day

		evaluate prevention and control protocols	Care & Pain Medicine, 2022, France	simulations under a Delta-wave scenario. The analysis focuses on a primary school setting (ages 6–11) in Lyon, France, modeling within-school contact networks and community-introduced cases.
26	Elisabetta Colosi, et al.	Screening and vaccination against COVID-19 to minimise school closure: a modelling study	The Lancet, 2022, France	Stochastic agent-based modeling study integrating empirical contact data with cost–benefit analysis of intervention protocols, calibrated on Alpha wave data (France, Mar–Jun 2021) and extended to Delta and Omicron scenarios. The analysis focuses on primary and secondary schools in France, modeling detailed contact networks between students and teachers and community-introduced cases.
27	Benjamin Faucher, et al.	Agent-based modelling of reactive vaccination of workplaces and schools against COVID-19	Nature Communications, 2022, France	Stochastic agent-based modeling study comparing reactive versus mass vaccination strategies, calibrated on Delta variant dynamics (autumn 2021) with simulations over a two-month period. The analysis is set in Metz (France), modeling interactions across schools (12+), universities, workplaces, households, transport, and community settings.
28	Laura Di Domenico, et al.	Modelling safe protocols for reopening schools during the COVID-19 pandemic in France	Nature Communications, 2021, France	Retrospective (ex-post) data-driven analysis calibrated on the first wave (up to April 2020), with projections covering the post-lockdown period from May 11 to August 2020. Conducted in Île-de-France (France), the analysis models contacts across households, schools (from preschool to high school), workplaces, transport, and leisure/community settings.
29	Thi Mui Pham, et al.	Seasonal patterns of sars-cov-2 transmission in secondary schools: a modelling study.	medRxiv, 2022, The Netherlands	Agent-based modeling study simulating Omicron transmission dynamics and evaluating the effectiveness of control interventions over a 30-month period starting January 3, 2022. The analysis focuses on secondary school settings (Netherlands), modeling interactions between students and teachers, including in-school contacts, leisure contacts, and community introductions.

Finally, I feel that the complexities involved in decision-making around school closures and reopenings during pandemics do not currently emerge from the text. Several factors can play a role in these decisions, including ensuring learning continuity, preventing social isolation among children, and reducing barriers to education. These important aspects are not mentioned and their consideration would provide a more comprehensive perspective.

Response:

Thank you for this insightful comment. We agree that the complexity of decision-making around school closures and reopenings extends beyond purely epidemiological considerations.

To address this point, we have revised the Introduction and Discussions to explicitly acknowledge key societal factors involved in these decisions, including learning continuity, social isolation among children, and inequalities in access to education.

Main text, Introduction, Page 3, Lines 52-57:

"[...] Consequently, decisions regarding school closures required a careful balance between epidemiological benefits and their substantial educational, social, and economic costs, including challenges related to maintaining learning continuity, mitigating social isolation among children, and ensuring equitable access to education [12].

To understand this balance, it is necessary to look beyond the school environment itself and consider the wider societal impact of these measures, which extend to multiple dimensions of well-being and access to essential services. [...]"

Main text, Discussion, Page 11, Lines 348-352:

"[...] However, the benefits of school closures on preventing such surges should be weighed against larger societal costs such as learning losses [58,65-68], deterioration in mental health among youth [69-71], and long-term macroeconomic consequences [72-74]. Globally, these costs fell disproportionately on children from lower socioeconomic backgrounds [75]. However, in Portugal, studies on this specific divide remain limited, making this disparity a concern given documented pandemic-era impacts on learning and mental health among youth [76-79]."

Discussion

In my previous point-by-point review, I asked what types of interventions could have allowed schools to remain open while keeping hospitalizations at reasonable levels. Could the authors please comment on this aspect to better support their conclusion regarding school closures during future pandemics?
The current response stating that the choice of interventions falls within the domain of policymakers is understandable. However, it would be valuable to include a brief discussion on the effectiveness and feasibility of full school closures.

Response:
Thank you for this important comment. We agree that our previous response did not sufficiently address the practical implications of our findings.

In the revised manuscript, we have expanded the Discussion to provide a clearer interpretation of the types of interventions that could support keeping schools open while maintaining hospitalizations at manageable levels. In particular, we now highlight the role of transmission outside the school setting, especially among adults, as a key factor modulating the impact of school reopening, and discuss how reductions in non-school contacts may mitigate the associated epidemiological burden.

We have also included a brief discussion on the effectiveness and feasibility of full school closures, emphasizing that their impact is context-dependent and may be more pronounced during early high-transmission phases, while becoming less effective when community transmission remains elevated.

Main text, Discussion, Page 11, Lines 353-362:
“Although our model is not designed to explicitly disentangle specific non-school interventions without additional data, our results provide indirect insights into alternative strategies that could support keeping schools open. In particular, transmission outside the school setting -- especially among adults — emerged in our analysis as a key determinant of the overall impact of school reopening. Our sensitivity analyses indicate that reductions in adult non-school contacts can substantially mitigate the increase in infections and hospitalizations associated with higher levels of school activity, whereas increases tend to amplify it [80]. This is consistent with previous studies showing that school-based measures alone are insufficient without concurrent reductions in community transmission [8, 18]. From a policy perspective, this suggests that interventions targeting adult mixing (e.g., remote work or reduced social interactions), potentially combined with partial school measures, may represent a feasible complement or alternative to full school closures, depending on the context.”

Reviewer 4

Comments on Reviewer 1's requests:

A. General methodological transparency (Reviewer 1 – General comment, Comment 1–2)

Several of Reviewer 1’s general methodological requests remain insufficiently addressed:

1. Assumptions transparency

- The reviewer requested a structured overview (e.g. a table) listing:
- Key model assumptions,
- Their empirical or literature support,
- Sensitivity analyses performed for each assumption, and
- The impact on main outcomes (with 95% credible intervals).
While assumptions are discussed throughout the text and supplement, this information is not yet consolidated in a transparent, reviewer-readable format, nor is it clearly stated which assumptions were revised in response to review.

Response:
Thank you for this suggestion to improve the transparency of our methods. We have created a consolidated overview of our key model assumptions, their empirical justifications, and the corresponding sensitivity analyses in the Supplementary Table 1. Where applicable, we also reference the newly added data tables (Supplementary Tables 6-11), which show the quantitative impacts on our main outcomes and sensitivity analyses.

2. Collateral effects of school closures

- The manuscript still lacks a dedicated "policy trade-offs" discussion summarizing evidence on learning loss, mental health, equity, and macroeconomic impacts of school closures. Reviewer 1 explicitly requested such a discussion, supported by the extensive literature already provided, and framed in the Portuguese context.
At present, epidemiological effects are discussed largely in isolation, which limits policy interpretability.

Response:
Thank you for this valuable suggestion. We agree that a clearer discussion of policy trade-offs is essential to improve the interpretability of our findings. In response, we have revised the Discussion to explicitly address the broader societal consequences of school closures, incorporating evidence on learning loss, mental health impacts, educational inequalities, and long-term macroeconomic consequences, supported by the literature and framed within the Portuguese context.

Main text, Discussion, Page 11, Lines 348-352:
"[...] However, the benefits of school closures on preventing such surges should be weighed against larger societal costs such as learning losses [58,65-68], deterioration in mental health among youth [69-71], and long-term macroeconomic consequences [72-74]. Globally, these costs fell disproportionately on children from lower socioeconomic backgrounds [75]. However, in Portugal, studies on this specific divide remain limited, making this disparity a concern given documented pandemic-era impacts on learning and mental health among youth [76-79]."

3. Portugal-specific contact data mapping

- The reviewer requested a table summarizing:
- Available Portugal-specific contact data (sources, dates),
- Overlap with the study period,
- How these data were used (calibration, validation, exclusion). This mapping remains missing, and it is still unclear how Portugal-specific contact evidence constrains or validates the inferred contact dynamics.

Response:
Thank you for this request. For the pre-pandemic contact matrices, the only viable options are the inferred synthetic matrices of Mistry et al. and Prem et al., which represent comparable values, though the latter provides lower age resolution. Portugal was not among the countries included in POLYMOD (Mossong et al., 2008), meaning no empirically measured pre-pandemic contact matrices exist for Portugal.

For pandemic-era contacts, no Portugal-specific data fully covering our study period were available. The only Portuguese pandemic contact survey identified is from CoMix (Wong et al., 2023; Verelst et al. 2021), but it overlaps only with the final weeks of our study period (December 22, 2020 - January 14, 2021; corresponding to their Waves A1-A2) and does not include individuals under 20 years, the age group central to our analysis. It was therefore unsuitable for calibration or validation.

In the absence of Portuguese pandemic contact data, we used Backer et al. (2021) to derive the ratio of lockdown-to-pre-pandemic contacts across age groups, which was then applied to the Portuguese baseline matrix, following Viana et al. (2021). We acknowledge this as a limitation of our study and have stated this explicitly in the Discussion. A structured summary of the closest datasets available and used or not-used in our study is provided in Supplementary Table S4.

Discussion, Page 11, Lines 370-372:
"[...] This approach ensured model validity, as it successfully reproduces observed dynamics; however, due to the lack of empirical contact data, the inferred contacts cannot be formally validated, which represents a limitation of our study."

4. Quantification of new sensitivity analyses

- Several newly added sensitivities (holidays, mitigation sweeps, partial closures) are described qualitatively, but Reviewer 1 explicitly requested numerical effects (e.g. changes in peak and cumulative hospitalizations and age-specific seroprevalence) with 95% CrIs, ideally summarized in a compact table in the main text.

Response:

Thank you for this suggestion. We have calculated the requested numerical effects (cumulative and peak hospitalizations, end-date seroprevalence, contacts, reproduction number, and youth cumulative elasticities) for all sensitivity analyses. To maintain the conciseness of the main text while providing the requested tables, we placed these in the Supplementary Materials (Tables S9 and S10). We have also updated the main text to direct readers to these supplementary tables, alongside other tables for the other sensitivity analyses.

Main text, Results, Robustness of results, Page 10, Lines 325-326:

"A comprehensive summary of all quantitative results across all evaluated sensitivities is provided in Supplementary Tables 6-11."

5. Justification of the "~80% reduction" in school contacts

The claim of strong school contact reduction is not yet supported by:
- A plot of the inferred school contact multiplier over time with uncertainty, or
- Explicit triangulation against external school-specific evidence to justify plausibility.

Response:
Thank you for pointing this out. The statement Reviewer 1 quoted, *"during the second wave, transmission relevant school contacts were already reduced by approximately 80% relative to pre-pandemic levels (Supplementary Table S1)"*, has been removed since we revised the main analysis to instead use regime-specific parameters (u_j's). The analysis that it refers to has been moved to the sensitivity analyses.

In the revised main analysis, the school contact multiplier for each period is directly represented by the inferred regime-specific parameter u_j. Because these multipliers correspond directly to these parameters, we provide their full posterior distributions in Supplementary Figure 6 to explicitly demonstrate the uncertainty requested. The numerical estimates with 95% credible intervals are summarized in Supplementary Table 2, and the resulting contact patterns over time are visualized in Supplementary Figures 3-5.

6. Identifiability and attribution diagnostics

The reviewer requested explicit discussion and diagnostics (e.g. posterior correlations, simulation-recovery) clarifying:
- What aspects of school vs. non-school effects are identifiable,
- Where attribution remains uncertain.
These diagnostics are still absent.

Response:
Thank you for this comment. We have added a figure of the posterior correlation matrix (Supplementary Figure 9). We also address the identifiability and attribution below.

Regarding structural identification, the contact split is fundamentally constrained by two aspects of the model:

1. The decomposition $b_{kl} = s_{kl} + (b_{kl} - s_{kl})$ is fully determined by Mistry et al., meaning the structure of school vs. non-school contacts is fixed rather than estimated.
2. The estimated parameters u_j's scale total contacts and are shared across both components, meaning posterior uncertainty propagates to both environments. The mitigation factor ζ is only applied to the lockdown (non-school) contacts.

Our posterior correlation matrix provides evidence of empirical identifiability, showing that most parameters are weakly correlated. Importantly, the mitigation factor ζ and the contact parameters u_j's are weakly correlated, which suggests that the first lockdown contact behavior is relatively separable from non-school contact behavior in subsequent regimes. The more noticeable correlations arise from expected pairings among transmission and hospitalization parameters, none of which directly undermine the school vs. non-school attribution.

To complement this, we re-fitted the model to run Stan diagnostics using 6 parallel chains with 300 warmup and 500 sampling iterations. This resulted in $\hat{R} < 1.01$ for all parameters, while the minimum Bulk ESS exceeded 500 and the minimum Tail ESS exceeded 600 across all parameters, which is sufficient for reliable estimation of posteriors. These diagnostics confirm that the sampler thoroughly mixed and converged on a stable, well-defined posterior distribution. We provide these diagnostics here for the reviewers' reference, omitting from the manuscript for brevity and to focus on epidemiological outcomes.

Despite the parameter estimation being computationally robust, we note that residual uncertainty remains from two structural assumptions not captured by the posterior alone:

1. whether the pre-pandemic contact matrices, which are derived from the established methodology of Mistry et al., perfectly map to real-world Portuguese contact behavior, and
2. whether non-school behaviors would remain unchanged under counterfactual school policies.

We therefore deliberately presented our counterfactual results across a spectrum of school mitigation levels in the original manuscript (Supplementary Section 14).

B. Definition of "school closure" (Reviewer 1 – Comment 3)

Reviewer 1's concern regarding the definition of "school closure" remains partially unresolved.

Specifically:
- The manuscript continues to use "school closure" as implying zero school-based contacts in the core model, while later analyses reintroduce residual contacts through mitigation factors.
- This creates ambiguity in how the term should be interpreted and how parameters (e.g. ζ_a□□) map to real-world contact patterns.
- The reviewer explicitly requested either:
- A definition closer to actual implementation in Portugal, or
- A clear justification that the zero-contact case is intended as an analytical extreme rather than a literal representation.

To resolve this comment, a brief but explicit clarification in the main text (Methods or Discussion) is still needed, explaining:
- That "school closure" is an analytical reference scenario representing maximal suppression of structured school contacts,
- That real-world closures involved non-zero residual attendance,
- Why a zero or near-zero assumption is nevertheless used (e.g. as an upper-bound counterfactual in the absence of detailed Portugal-specific contact data),
- How readers should interpret results in light of the mitigation-spectrum analyses.

Response:
Thank you for this comment. To resolve this, we have added explicit clarifications in both the Methods and Discussion sections to state that the zero-contact assumption is intended as an analytical extreme of maximal suppression, which serves as an upper-bound counterfactual rather than a literal description of what occurred in Portugal. We acknowledge that real-world closures involved non-zero residual attendance (Heinsohn et al., 2022), so Scenario 2 likely overestimates the epidemiological impact of the policy as implemented. This scenario is retained because no sufficiently granular Portugal-specific contact data on residual attendance exist for the study period (Supplementary Table 5). However, we also presented results on a spectrum of intermediate scenarios (Figure 5 and Supplementary Section 14).

Main text, Methods, Time-varying contact patterns, Page 6, Lines 184-186:
"[...] Regarding the third assumption, this scenario of zero school-based contacts is intended as an analytical extreme of maximal suppression, which serves as an upper-bound counterfactual since real-world closures involve non-zero residual attendance. […]"

Main text, Discussions, Page 10, Lines 336-339:
"[...] Because real-world closures involved non-zero residual attendance, our main

results for full closure should be interpreted as upper-bound estimates of the epidemiological impact of school closures. The intermediate scenarios addressed in our sensitivity analyses, which vary partial school mitigation levels, reflect the more nuanced real-world conditions. […]"

C. Concurrence of measures and attribution (Reviewer 1 – Comment 8)

While the added overview of measures in Portugal is appreciated, the core conceptual concern remains only partially addressed:
- The manuscript still lacks a sufficiently clear explanation of how the model design manages concurrence of interventions when all non-school measures are aggregated into a single parameter.
- It remains unclear:
- Under what assumptions can the effect of school closures be interpreted as distinct,
- What cannot be disentangled given this aggregation,
- How these limitations affect causal interpretation.

Reviewer 1 explicitly requested a more detailed discussion of this issue, which is still needed in the Discussion section or in supplementary information document.

Response:
Thank you for this comment. First, we would like to clarify that non-school measures are not aggregated into a single parameter. Instead, they are implicitly captured by a_{kl}, the inferred lockdown contacts. On top of this, the parameter ζ captures the mitigation factor on transmission per contact due to protective measures not explicitly modeled. For the subsequent regimes, the parameter $(1-u_j)$ describes the period-specific proportion of time an individual behaves as during the pandemic. These three jointly describe non-school measures for each period, transitioning smoothly via estimated logistic functions. Second, for the counterfactual scenarios, total contacts are separated into additive school and non-school components in which only the school contacts are varied to represent school opening and closure. This is so that differences to the baseline are solely arising due to changes in school contacts.

Given these two model design choices, we additionally assume that non-school behavior would have been the same regardless of the school policies implemented, which may not hold in practice. With that, our results should be interpreted as quantifying the marginal contribution of school contacts to transmission, acknowledging that the concurrent contributions of individual non-school measures cannot be disentangled. Indeed, disentangling the effects of specific non-school interventions is not the goal of our work, nor would it be possible without additional, highly granular behavioral data. We have made this explicit in the main text and the supplement.

Main text, Methods, Contact rates, Counterfactual scenarios vs baseline, Page 5, Lines 151-152:
"[…] Note that in both scenarios, non-school contacts evolve as in reality so that scenario differences are solely attributed to changes in school contacts. […]"

Main text, Discussion, Page 11, Lines 353-354:
"Although our model is not designed to explicitly disentangle specific non-school

interventions without additional data, our results provide indirect insights into alternative strategies that could support keeping schools open. […]"

Supplementary materials, Time-varying contact patterns, On non-school contacts, Page 15, Lines 773-779:

"For the counterfactual scenarios, total contacts are separated into additive school and non-school components, with only school contacts varied to represent opening and closure, while non-school contacts evolve as in the baseline. This, however, assumes that non-school behavior would have been identical regardless of the school policy adopted, which may not hold in practice, as school-related measures in Portugal were consistently implemented alongside broader non-school interventions (Supplementary Section 2). Our subsequent results should therefore be interpreted in light of this assumption, acknowledging that the independent contributions of individual non-school measures cannot be disentangled."

4. Minor residual points (not blocking, but noted)

- Comment 4 (age-specific susceptibility priors): Largely addressed; a brief justification of the choice of prior study or a statement on robustness to alternative priors would fully close the loop.

Response:
Thank you for your comment. We chose Jing et al. (2020) as the basis of our prior distributions for the age-specific susceptibilities, as it provides age-specific estimates obtained from a reliable methodology, and it also covers the pre-vaccination period. Yet, the prior remains broad and weakly informative since we used a log-normal scale parameter of 0.5 to ensure that posterior estimates are primarily data-driven. Moreover, the posterior estimates we got are consistent with those reported by Viner et al. (2021), which supports robustness to the specific prior choice. We have added a brief clarification to this effect in the Supplementary material.

Supplementary materials, Parameter estimation, Prior distributions, Page 17, Lines 785-787:
"*The relative susceptibilities of the 0-20 years age group, $\beta_{1,2,3}$, and of the 20-60 years age group, $\beta_{4,5,6,7}$, had log-normal prior distributions with medians 0.23 and 0.64, respectively, and a scale parameter of 0.5 [2], ensuring a weakly informative prior.*"

- Comment 5 (hospitalized individuals assumed non-infectious): Adequately justified; although a sensitivity analysis was requested, the explanation provided is likely sufficient.

Response:
Thank you for your review of this assumption. We appreciate your confirmation that the justification provided is sufficient.

- Comment 6 (seroprevalence timing): Adequately addressed; a short quantitative summary of sensitivity results would strengthen completeness.

Response:
Thank you for your comment. Two tables of model outcomes for the two counterfactual periods are added to summarize the results of the sensitivity analyses (Supplementary Tables 8-9).

- Comment 7 (uptake delays, Ki): Fully addressed.

Response:
Thank you for acknowledging that this point is now fully resolved.

Response to reviewers' comments

Reviewer 3

Although the authors have addressed the suggestions raised during the previous round of review, the main message of the paper remains somewhat unclear to me. In particular, it is not evident whether the authors are suggesting that school closures should still be considered a viable option to curb transmission during future pandemics.

For example, the manuscript states: "Our results suggest that integrated strategies combining school-based and broader community interventions are essential to effectively reduce transmission during future pandemics." It is unclear what is meant by "school-based interventions" in this context. Are the authors referring specifically to school closures, which are the focus of the analysis, or to alternative school-based mitigation measures? If the latter, these interventions are not evaluated in the present study and therefore should introduced and discussed before.

Apart from this point, I believe the authors have addressed the comments.

Response:
Thank you for this comment, which gave us an opportunity to clarify both our terminology and our central message.

First, on whether school closures should still be considered a viable option in future pandemics: yes and we now say so explicitly. The final paragraph of the Introduction (Lines 80-83) and the reworded closing sentences of the Abstract (Lines 21-24) and Conclusions (Lines 416-421) state that school closures remain a viable option, but are most effective when integrated with broader community interventions, particularly in later phases of high transmission outside schools, where closures alone did not bring $R(t)$ below 1.

Second, on what we mean by *"school-based interventions"*: we refer to reductions in school-related contacts, from partial mitigation to full closure, as evaluated through the parameter ξ, not to distinct in-school measures such as mask-wearing, ventilation, or testing. We now define this in the Methods (Lines 155-157).

Third, on whether we mean closures or alternative school-based measures: we mean closures (and, more broadly, reductions in school contacts). Our conclusions do not rest on the two extremes alone, i.e., intermediate levels are quantified in our sensitivity analyses (Figure 5), so the full spectrum is accounted for.

We hope these revisions resolve the identified ambiguity.

Reviewer 4

All my comments have been adequately addressed.

Response:
Thank you for your positive assessment of our manuscript and helpful feedback. We are pleased that our revisions have satisfactorily addressed all raised concerns.