

Dear Reviewer,

Thank you for the time you dedicated to reevaluating our manuscript and for your valuable suggestions. Your comments have been extremely helpful in improving the quality and clarity of our work.

The revised version of the manuscript, incorporating all the changes made in response to the reviewers' suggestions, is enclosed. The first review step is highlighted in yellow, the second in green

Below we provide a point-by-point response to the comments raised

- R1 — Age-specific incidence denominators and vaccination status. The authors acknowledge the limitation appropriately and have expanded the PREMAL description to note that vaccination status is not captured in the surveillance data. This is an acceptable response given the constraint that individual-level vaccination records cannot be linked. The revised Methods text provided in the cover letter is clear and proportionate. This comment is considered satisfactorily addressed, provided the expanded PREMAL description appears in the published manuscript exactly as quoted.

ANSWER: thank you

We confirm that the expanded PREMAL description will appear in the published manuscript exactly as quoted.

- R2 — Time-varying case ascertainment as a confounder of temporal trends. The authors accept the substantive point and report having strengthened the Discussion to acknowledge differential surveillance sensitivity across time. This is welcome.

ANSWER: thank you

- R3 — Misuse of the term "cumulative incidence." This remains the most significant unresolved issue. The authors state that they "calculated incidence rates, presented the annual rates individually, and reported the period incidence," and have added a mean annual incidence rate with confidence intervals. This is an improvement, but the response does not confirm whether the term "cumulative incidence" has been replaced throughout the manuscript, or whether the metric previously reported as "cumulative incidence 2019–2024" (18.19 per 100,000) continues to appear under that label. Inspection of the revised manuscript tables (Supplementary Table 1 and Tables 1–2) suggests that the column header "Cumulative incidence 2019–2024" is retained, as is the associated figure of 18.19 per 100,000 in the abstract. This is terminologically incorrect. A quantity derived by dividing total cases over a multi-year period by an average annual population count is a period incidence rate or, at best, a mean annual rate, not a cumulative incidence in any epidemiologically defensible sense. Cumulative incidence requires a defined cohort, event-free at baseline, followed for the duration of the period, and is bounded between 0 and 1. The retained terminology will mislead readers and is inconsistent with standard epidemiological usage. The authors must either rename this measure throughout (including in the abstract, tables, and supplementary material), or replace it entirely with the mean annual incidence rate now correctly calculated and presented. This revision is mandatory.

ANSWER: We thank the reviewer for this important observation. We fully agree with the suggestion and have revised the manuscript accordingly to avoid any potential misunderstanding for readers. Specifically, the measure has been consistently renamed throughout the manuscript—including the abstract, tables, and supplementary material.

- R4 — Missing hospitalization data and confidence intervals. The addition of confidence intervals across all tabulated estimates is appropriate and directly responsive to the reviewer's request; this element is satisfactorily addressed. The treatment of missing hospitalization data, however, remains inadequate. The authors assert that PREMAL is standardised across Local Health Authorities and that data collection is uniform, which they appear to offer as reassurance that the missing data are inconsequential. This conflates standardisation of data collection procedures with completeness of data reporting — two distinct properties of any surveillance system. Standardised forms do not preclude differential missingness if reporting behaviour varies systematically by year, LHA, age group, or clinical severity. I explicitly requested quantification of the proportion of missing values stratified by year and LHA, as well as a discussion of whether a missing-not-at-random mechanism is plausible. Neither has been provided. At minimum, the proportion of cases with "NA" hospitalization status should be reported in a supplementary table stratified by year and LHA, and the Discussion should acknowledge that if missing records are concentrated in 2024 (the year with the largest case load), apparent hospitalization rates for that year may be understated. A sensitivity analysis under a plausible worst-case assumption, for example, treating all NA cases as hospitalised, would further strengthen the analysis.

ANSWER: All reported cases include information on hospitalization; therefore, there are no NA values for this variable. We confirm what was previously stated regarding the PREMAL surveillance system.

Following your comment, we noticed that the Discussion section included the sentence: “Nevertheless, the high proportion of notifications with incomplete information on hospitalization status (‘NA’) limited the ability to conduct an in-depth analysis of disease severity and hospital burden.”

We agree that this wording could be misleading for the reader. As shown in Supplementary Table 2, there are no NA values for hospitalization. The missing data instead refer to clinical information and to the different variables introduced in PREMAL starting from 2024 (as described in the Materials and Methods section, under PREMAL).

For clarity, we have therefore rephrased the sentence in the manuscript.

"However, the high proportion of notifications with missing information on disease severity limited the ability to conduct an in-depth analysis of clinical severity and the associated hospital burden. Although this variable has been included in the surveillance system since 2024, the extent of missing data prevents a comprehensive assessment."

- R5 — Inferential language and statistical significance. The authors confirm that "significant differences" has been replaced with appropriate descriptive language throughout. This is satisfactorily addressed and requires no further action.

ANSWER: thank you

- The confidence intervals reported for several strata in Tables 1 and 2 include negative lower bounds (e.g., mean annual incidence for the 15–19 year group: –0.10 to 4.29; for the 20–24 year group: –0.38 to 0.58). Negative incidence rates are not epidemiologically interpretable. This arises from the application of normal approximation methods to Poisson counts that are too small for this approximation to be valid. Where case counts are sparse (fewer than five events per stratum per year), exact Poisson confidence intervals, or, if preferred, score-based intervals, should be substituted. The presence of negative lower bounds in published tables constitutes a reporting error that should be corrected prior to acceptance

ANSWER: Thank you very much for your valuable comment.

In response to your suggestion: we recalculated the confidence intervals. Specifically, for observations with counts  $\leq 5$ , we calculated the confidence intervals using the exact Poisson method. This methodological clarification has also been added to the Materials and Methods section of the manuscript.

We also have incorporated the confidence intervals (CIs) directly into the text of the manuscript. We sincerely appreciate your careful reading and constructive comments, which have helped us improve the methodological accuracy of the article