Llywodraeth Cymru
Welsh Government

Science Research Evidence

Retrospective analysis of short-term projections

22 September 2025

gov.wales

Retrospective analysis of short-term projections

Contents

Summary

Acute respiratory infections, including influenza, Respiratory Syncytial virus (RSV), and COVID-19, place a significant burden on hospitals across Wales, particularly during the winter season. Our modelling informs NHS resource planning. Therefore, accurately predicting patient admissions is crucial. We developed a statistical model to estimate admissions up to 14 days in advance. Following the winter of 2023/24, we assessed the model's performance by comparing its projections with the actual data. Our analysis indicates that the model tends to over-predict during peak periods, and that the projections are more reliable at the national level than at the local health board level.

Introduction

Acute respiratory infections significantly increase the demand on the NHS in Wales each year. In the 2023/24 season, there were a total of 14,110 admissions due to Influenza, 5,168 due to RSV, and 6,574 due to COVID-19.[1] Anticipating healthcare demand enables more effective allocation of scarce resources during periods of system pressure. The Science, Research and Evidence (SRE) division publishes an annual report estimating the pressures for the upcoming winter. However, as winter progresses, there is a need to adapt the model to real-time data and create short-term projections (STPs) to help with more accurate planning.

To do this, the modelling team in SRE collaborated with UK Health Security Agency (UKHSA) who developed a hierarchical generalised additive model (GAM) for the short-term forecasting of influenza hospital admissions in England for the winter of 2022/23 which they have been using when relevant.[2] [3] This model incorporated both spatial and temporal structures to forecast admissions two weeks in advance. The authors compared the performance of the GAM with other statistical models, such as the Autoregressive integrated moving average (ARIMA) and Prophet by Meta,[4] and found that the GAM provided more accurate projections. Given the strong

[1] Science Evidence Advice: winter modelling 2024 to 2025
[2] Forecasting influenza hospital admissions within English sub-regions using hierarchical generalised additive models | Communications Medicine (nature.com)
[3] Forecasting COVID-19, Influenza and RSV hospitalisations over winter 2023/24 in England | medRxiv
[4] Prophet | Forecasting at scale.

performance of the model, we collaborated closely with the UKHSA to tailor this model specifically to Wales. We created short-term admissions projections for COVID-19, Influenza and Respiratory Syncytial virus (RSV) throughout the winter of 2023/24.[5] The aim of the paper is to retrospectively evaluate how the short-term projections had performed in Wales by comparing the model estimates to the actual data.

Methods

Data Sources

- **Admissions Data:** Weekly new admissions due to influenza and RSV were obtained from the Public Health Wales (PHW) dashboard. [6] The daily COVID-19 admissions data was directly obtained by Public Health Wales and aggregated by health board.
- **Population Demographics:** Demographic information was sourced from Stats Wales. [7]
- **Geographic Information:** The geographic boundaries of health boards were retrieved from the Open Geography Portal. [8]

Model

The GAM model looked for spatial and temporal trends in admissions across 7 local health boards in Wales. It estimated the growth rate of hospital admissions and projected this forward in time, assuming the same growth rate would continue. The model employed flexible functions known as smooths, which are composed of many smaller basis functions that combine to give the function it's shape (k-values, see Appendix). The influenza and COVID-19 model analysed per-capita admissions as a smooth function of time at both the national and local health board levels. Additionally, shared penalties between adjacent health boards were represented

[5] For the purposes of this paper, we assume winter runs from 1 September 2023 to 31 March 2024
[6] ARI - Hospital admissions dashboard | Tableau Public
[7] National level population estimates by year, age and UK country (gov.wales)
[8] Open Geography Portal (statistics.gov.uk)

using a Markov random field. For RSV, the model was adjusted to include per-capita admissions as a smooth function of time and age groups and a shared penalty between age groups to account for the age-related component of RSV admissions. The model output was produced as admissions began to rise each winter. As different diseases show different temporal trends (RSV cases typically rise in late September while flu cases tend to rise in December), this resulted in model output being created using different projection dates for each disease. Due to low admission numbers, Powys Teaching Health Board was excluded from the model. The model output included the central estimate (the expected outcome) along with a prediction interval (PI), which accounted for the uncertainty in both the model parameters and the data (noise).

Spline tuning (model optimisation)

Before generating the model output, the number of basis functions for each smooth were optimised by spline tuning (as described below). 90 days of admissions data was divided into four 8-week segments where the model output could be compared with the historical data we already had. An iterative process on the model was implemented using a for loop with all possible sets of basis functions. The number of basis functions that minimised the prediction error (i.e. SMAPE, for more information see below) were selected as the optimal model parameters.

Model Evaluation Measures

Once the model outputs were generated, they were compared with the actual admissions data using the following evaluation measures:

- **SMAPE (Symmetric mean absolute percentage error):** an error metric that normalises the relative errors by dividing by both actual data and model estimate, (bounded between 0 and 200% (lower SMAPE values indicate optimal performance))

$$SMAPE = \frac{1}{n} \times \sum \frac{|\,forecast\ value\ - actual\ value\,|}{(|\,actual\ value\,| + |\,forecast\ value\,|)/2}$$

- **95% Interval coverage**: the percentage of actual data points that lie within 95% prediction interval of the model estimate, bounded between 0% and

100% (interval coverage is considered optimal when the measured coverage matches the nominal coverage i.e. 95%). [9]

Bias: Relative tendency to over- or under-predict (aspect of calibration), bounded between -1 and 1 (negative values indicate under prediction while positive values indicate overprediction). To determine if the model exhibited any bias, the bias scores for each model (at the national level) were tested against zero using a one-sample Wilcoxon test.

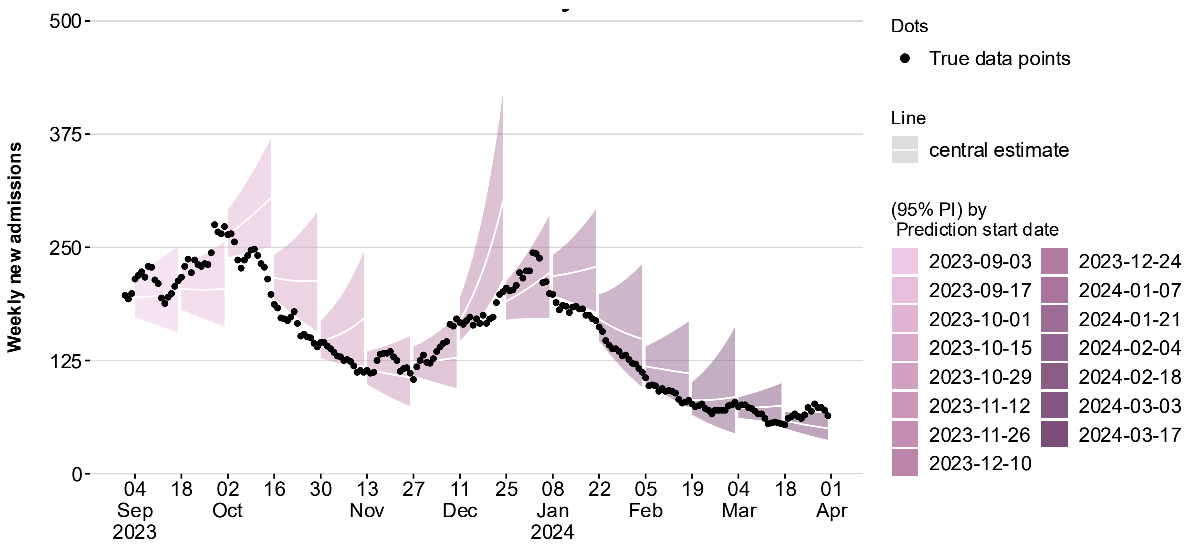
Results

COVID-19 Model

COVID-19 admissions showed two peaks in the winter of 2023/24, the first in late September and the second peak in January. The short-term projections for COVID-19 were created every two weeks from 3rd September 2023 until 31st March 2024 spanning 15 projection dates (Figure 1). The COVID-19 model performed better at the national level with lower SMAPE value of 16.86 (vs 31.22 at the local health board level, Table 1). A similar trend is observed in interval coverage, with national level estimates showing higher interval coverage than local health board level estimates (68.1% vs 40.32%, Table 1). Overall, the model exhibited a tendency to overpredict, with a bias score of 0.29 and 0.1 at the national and local health board level respectively. This effect is more pronounced during the peak/turning points since the model cannot accurately estimate the timing of peak (see projection dates 1st October 2023 and 10th Dec 2023).

[9] For a 95% interval coverage, the optimal performance would be when 95% of the data points (for example 19 out of 20) to fall within the interval. Similarly, for a 50% interval coverage, the optimal performance would be when 50% of the data points to lie between the upper and lower bounds.

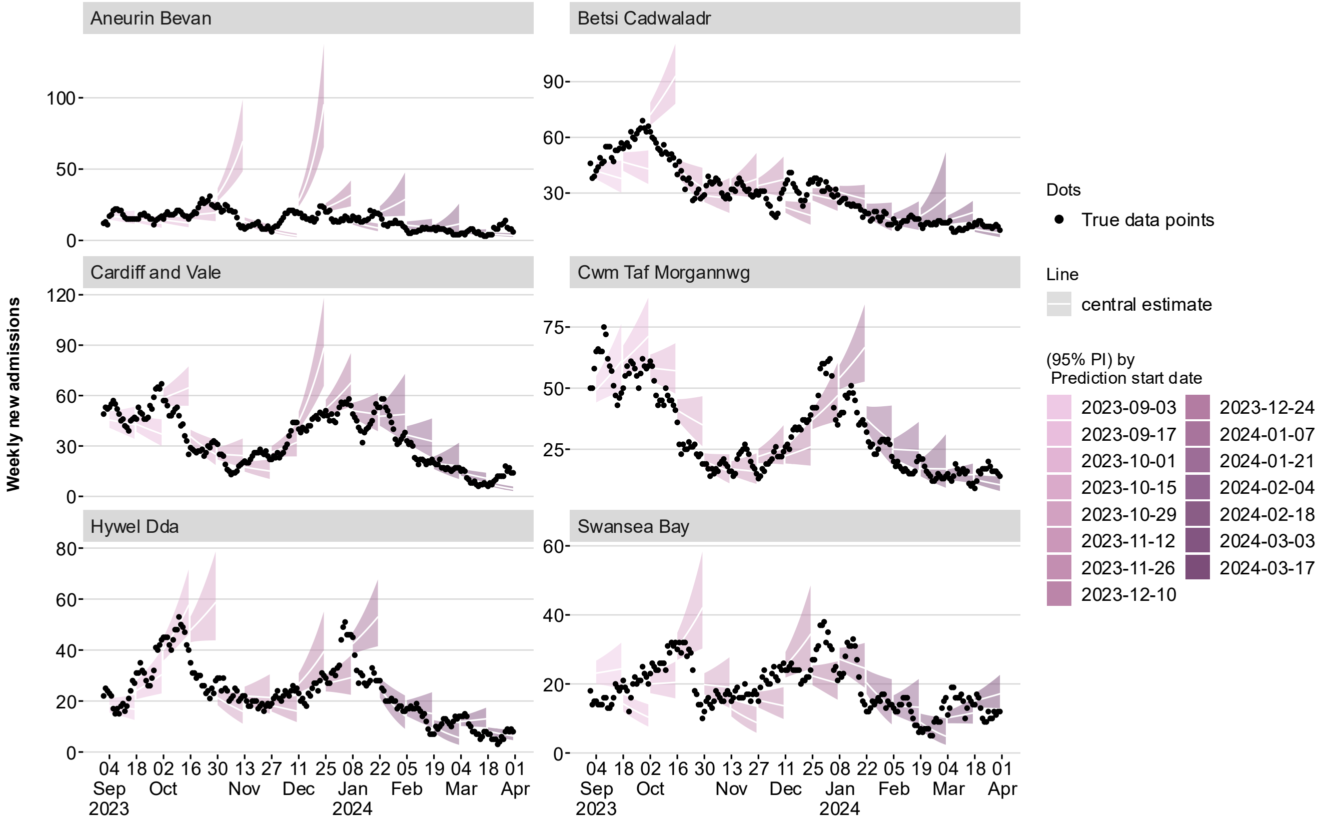
Figure 1: COVID-19 Admissions Projections between September 2023 and March 2024, Wales



Source: Public Health Wales and SRE Calculations

All performance metrics deteriorated significantly near or during peak dates, exhibiting the highest SMAPE (error), lowest interval coverage, and highest bias values (Figure 3).

Figure 2: Local Health Board COVID-19 Admissions Projections between September 2023 and March 2024, Wales



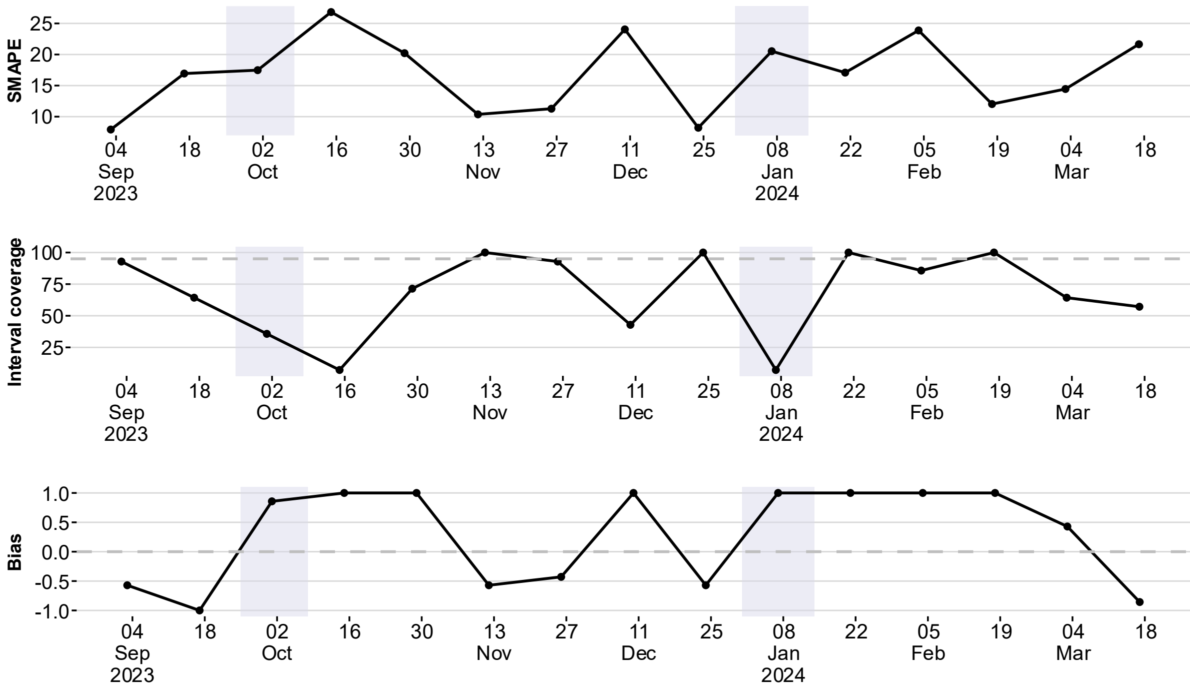
Source: Public Health Wales and SRE Calculations

Table 1: Performance metrics of the COVID-19 GAM model throughout the 2023/24 winter season

	SMAPE(%)	95% Interval coverage (%)	Bias
National	16.86	68.1	0.29
Local Health Board	31.22	40.32	0.10

Source: SRE calculations

Figure 3: Evaluation metrics of COVID-19 model estimates at the national level by projection date [Note 1][Note 2]



Source: SRE calculations

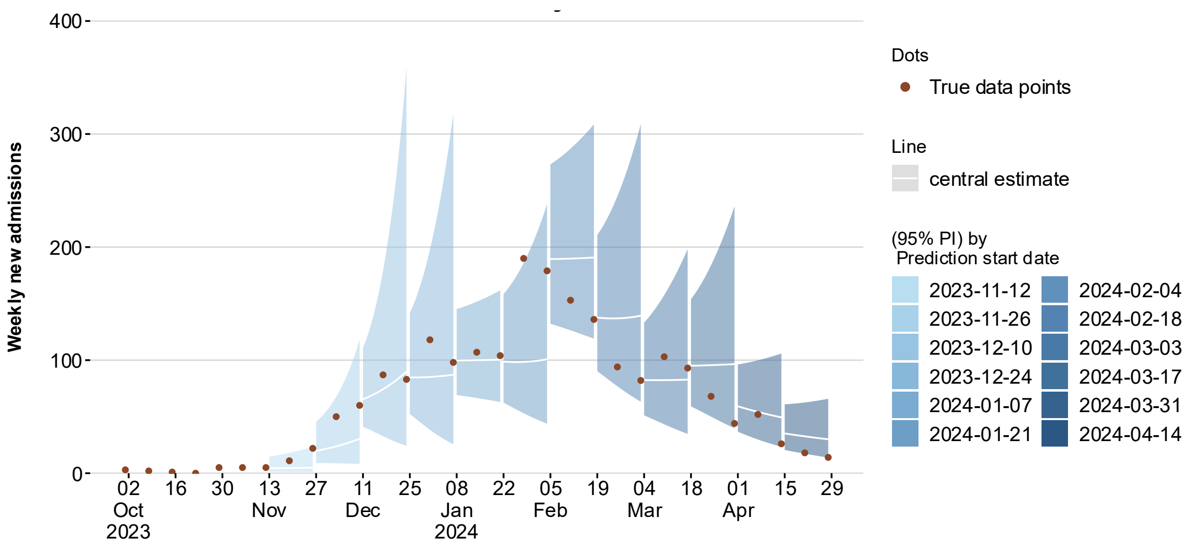
[Note 1]: Purple bars indicate the peaks in COVID-19 admissions during the 2023/24 winter season.

[Note 2]: Grey dashed lines indicate optimal performance of the model

Influenza Model

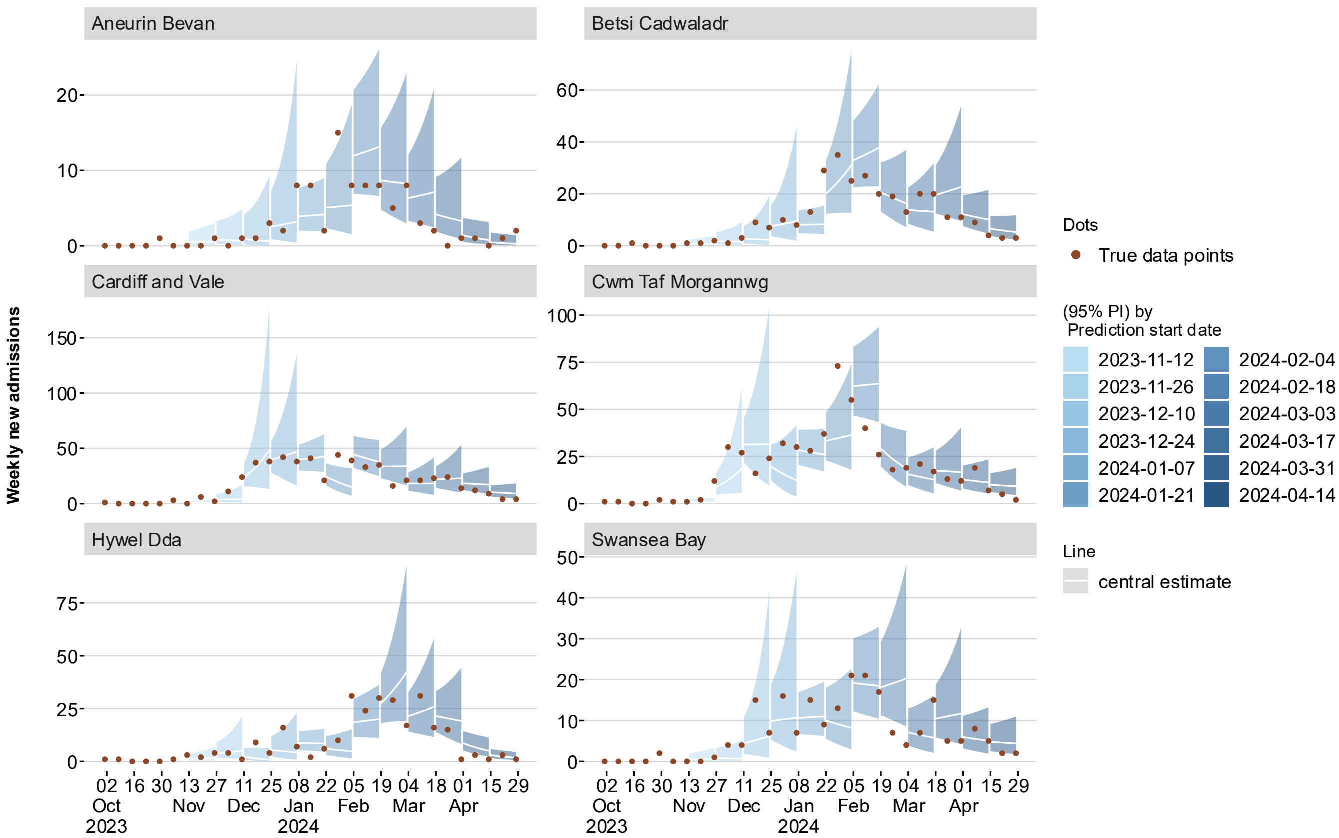
The short-term projections for Influenza were created every two weeks from 12th November 2023 until14th April 2024 spanning 12 projection dates. The influenza model performed better at the national level (SMAPE = 42.81%, interval coverage = 95.83%, Figure 4 and Table 2) than at the local health board level (SMAPE =65.55%, interval coverage= 70.83%, Figure 5 and Table 2). The model exhibited a tendency to underpredict slightly at the national level, with a bias score of -0.08. While the model exhibited a bias score of zero at the local health board level, it is important to note that the model under-predicted during the first half of the season (growth phase) and over-predicted during the second half of the season. Compared to the COVID-19 model output, the flu model has higher interval coverage but lower SMAPE.

Figure 4: Influenza Admissions Projections between November 2023 and April 2024, Wales



Source: Public Health Wales and SRE Calculations

Figure 5: Local Health Board Influenza Admissions Projections between November 2023 and April 2024, Wales



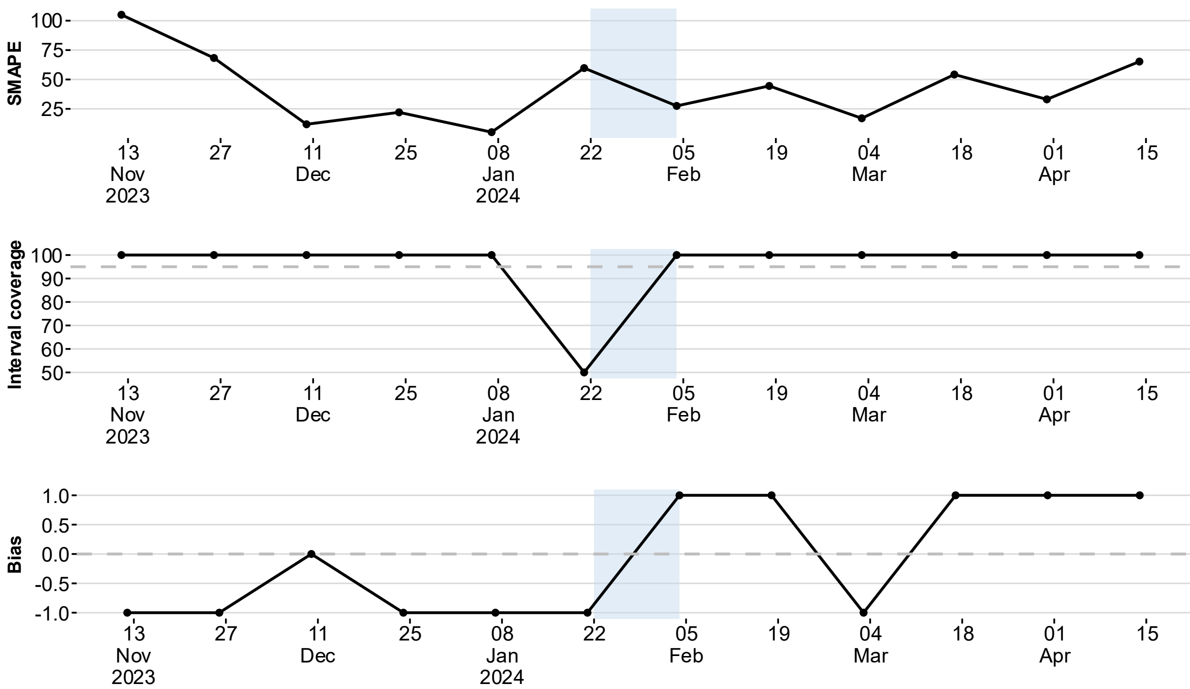
Source: Public Health Wales and SRE Calculations

Table 2: Performance metrics of Influenza GAM model throughout the 2023/24 winter season

	SMAPE (%)	95% Interval coverage (%)	Bias
National	42.81	95.83	-0.08
Local Health Board	65.55	70.83	0.0

Source: SRE calculations

Figure 6: Performance metrics of influenza model at the national level by projection date [Note 1]



Source: SRE calculations

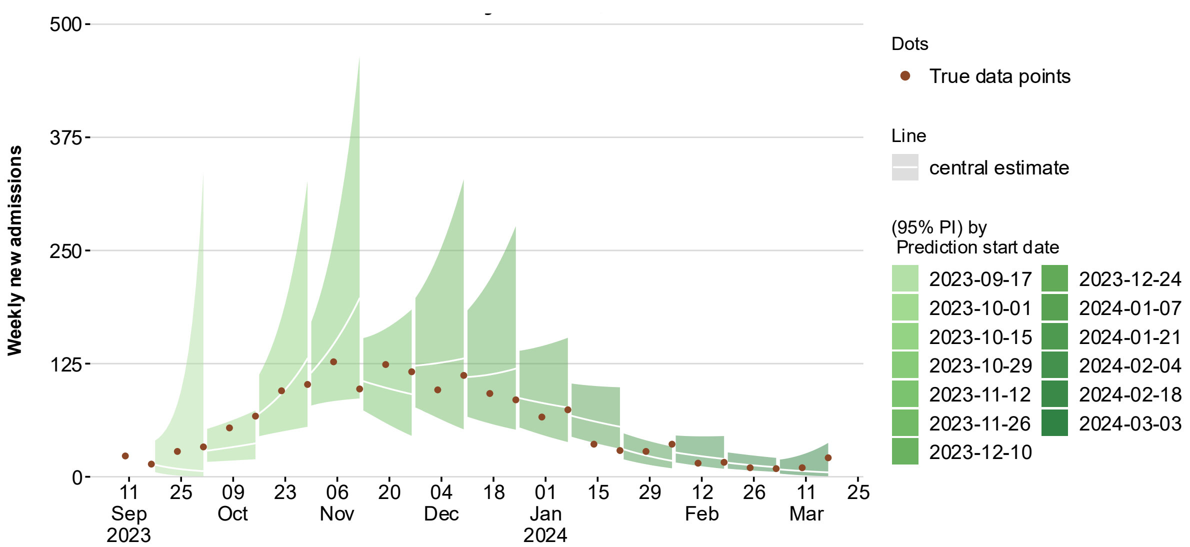
[Note 1]: Blue bar indicates the peak in Influenza admissions during the 2023/24 winter season.

[Note 2]: Grey dashed lines indicate optimal performance of the model

RSV Model

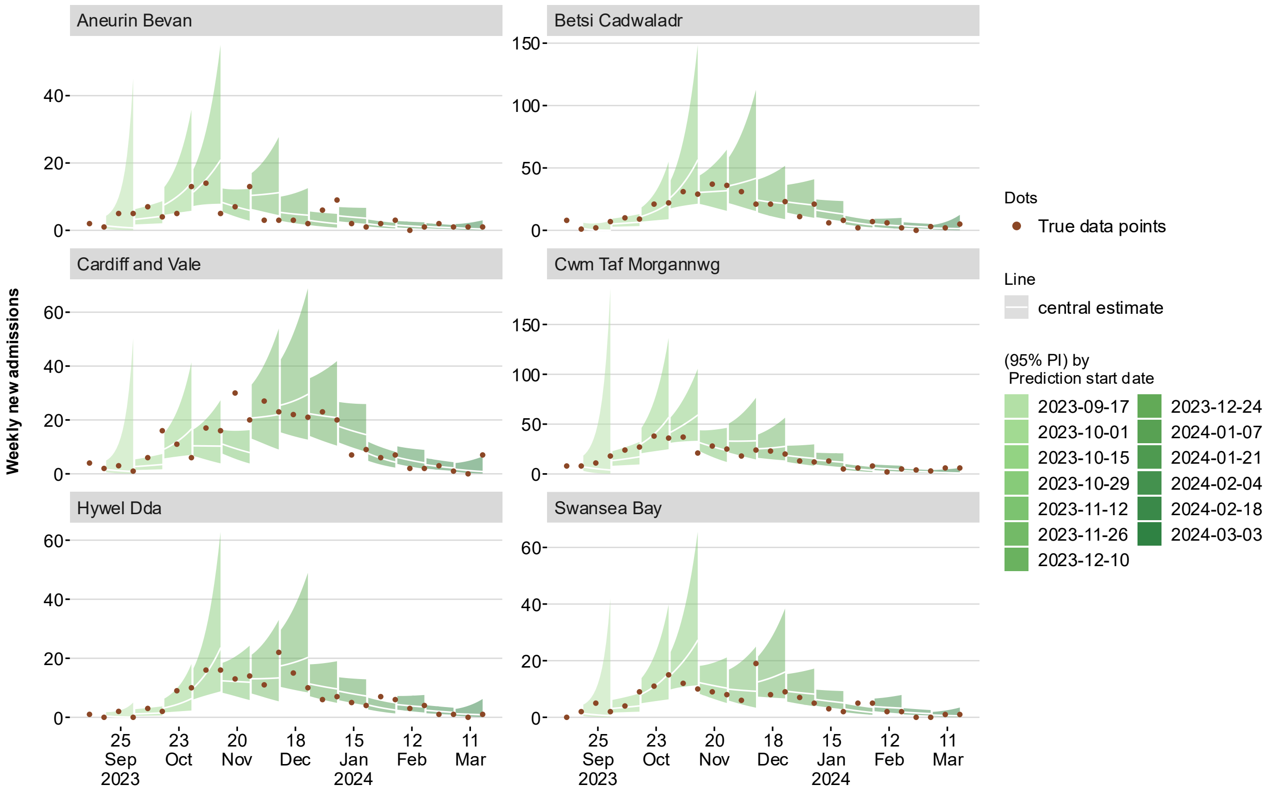
The short-term projections for RSV were created every two weeks from 17th September 2023 until 17th March 2024 spanning 13 projection dates. The RSV model included an age-specific component, which was not present in the COVID-19 and Influenza models. The RSV model performed better at the national level with a high interval coverage of 88.46% than at the local health board level (69.87%) and at the age level (58.97%). The model exhibited a tendency to overpredict at the national level and the age level, with a bias scores of 0.15 and 0.21 respectively (bias scores at local health board=-0.01).

Figure 7: RSV admissions projections between September 2023 and March 2024, Wales



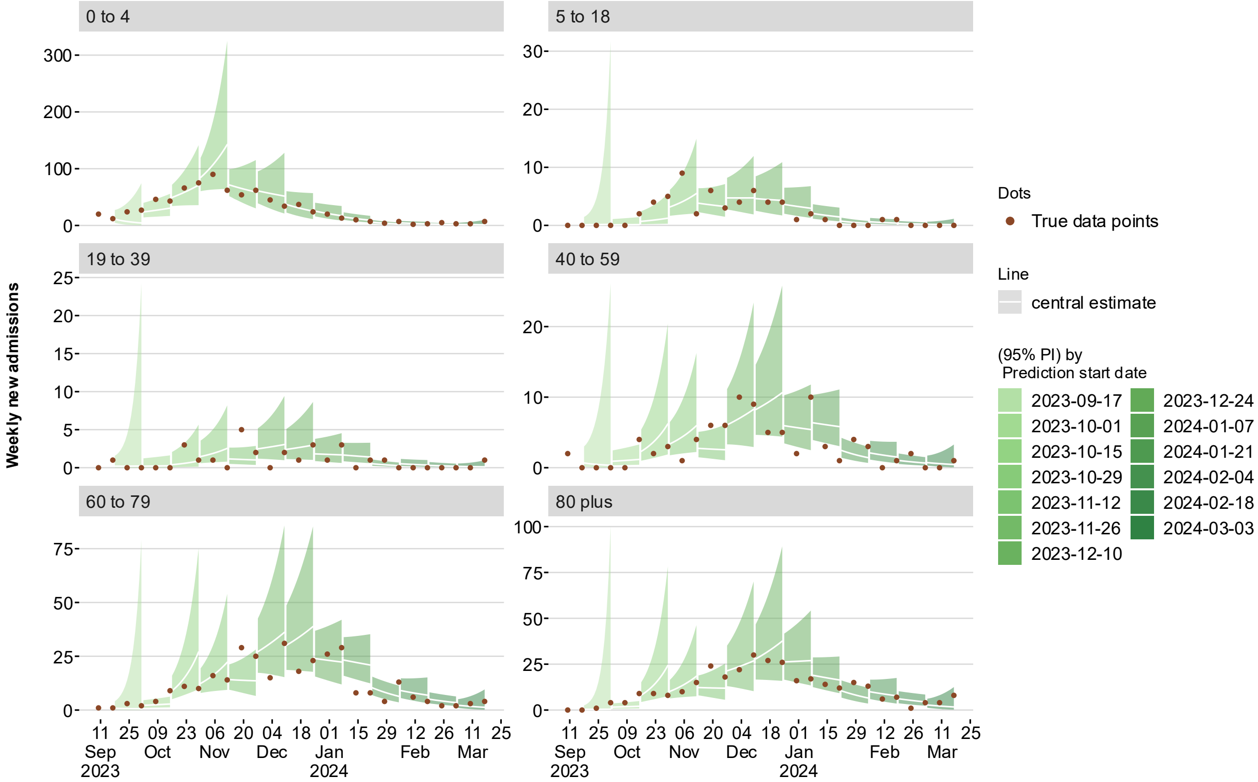
Source: Public Health Wales and SRE Calculations

Figure 8: Local health board RSV admissions projections between September 2023 and March 2024, Wales



Source: Public Health Wales and SRE Calculations

Figure 9: Age wise breakdown of RSV admissions projections between September 2023 and March 2024, Wales



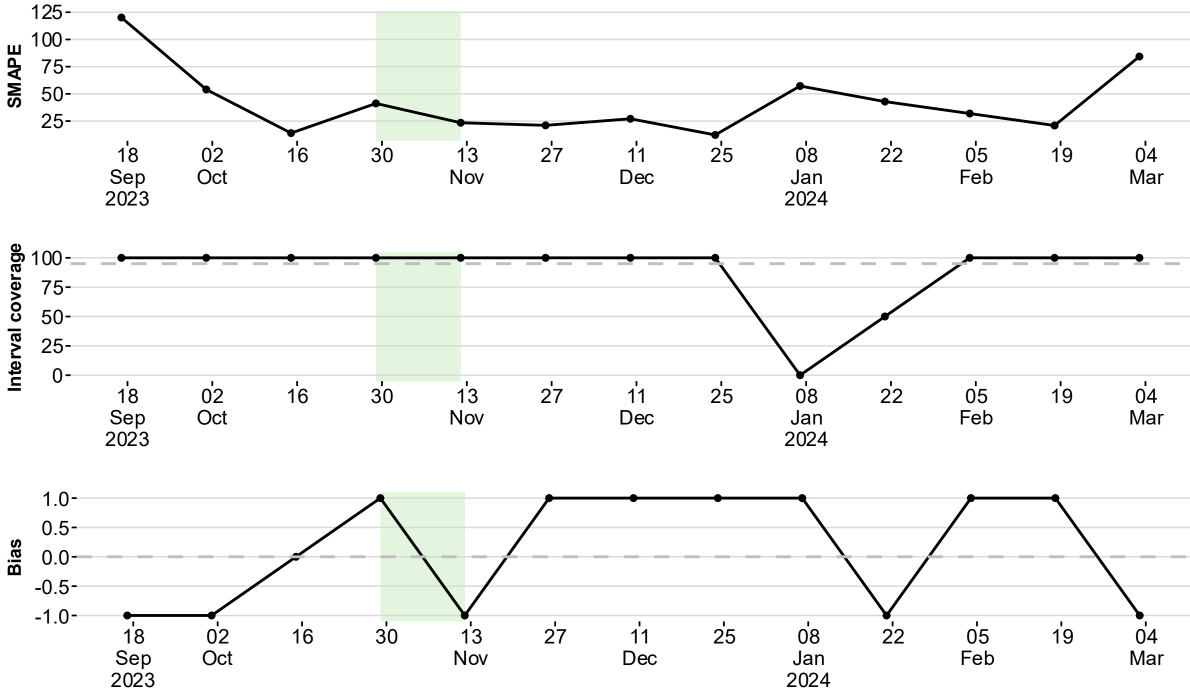
Source: Public Health Wales and SRE Calculations

Table 3: Performance metrics of the RSV GAM model throughout the 2023/24 winter season

Projection level	SMAPE (%)	95% Interval coverage (%)	Bias
National	42.34	88.46	0.15
Local Health Board	61.16	69.87	-0.01
Age Group	86.23	58.97	0.21

Source: SRE calculations

Figure 10: Performance metrics of RSV model at the national level by projection date [Note 1] [Note 2]



Source: SRE calculations

[Note 1]: Green bar indicates the peak in RSV admissions during the 2023/24 winter season.

[Note 2]: Grey dashed lines indicate optimal performance of the model

Discussion

The model's output, after spline tuning, aligned well with the historical ranges of hospital admissions. However, the highest errors occurred during/near peak admissions, as the model struggled to estimate these peaks accurately. The UKHSA has already compared the performance of Generalized Additive Models (GAMs) with other statistical techniques such as ARIMA and Prophet, finding that GAMs performed better. Although we did not conduct a comparative analysis with other models, we believe our model is robust as it makes more reliable assumptions about spatial relationships among local health boards in Wales and exponential growth in admissions (absent in the other models).

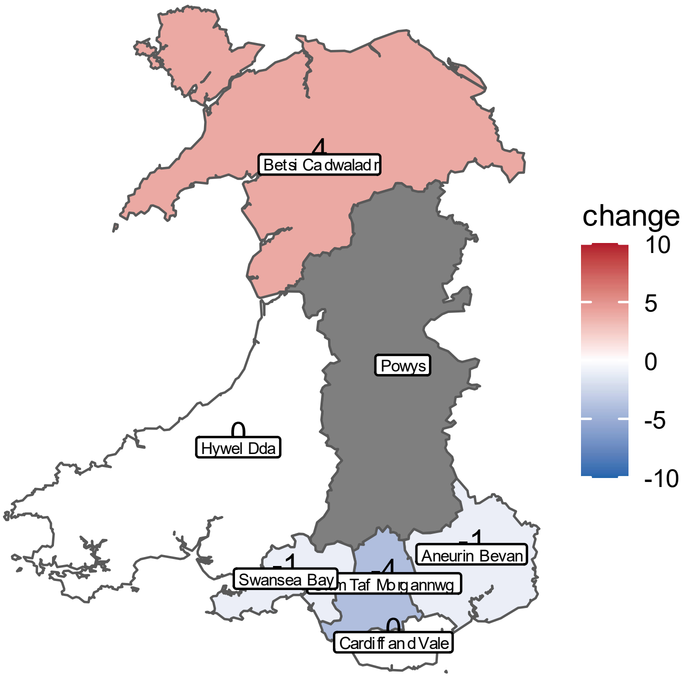
We evaluated the model's performance using three metrics: SMAPE, interval coverage, and bias. Among these, SMAPE is particularly useful as it measured the error between the model’s central estimate and the actual data points and provides a clear indication of the model's accuracy. However, defining a "good" or "acceptable" range for metrics such as absolute error or SMAPE can be challenging and context dependent. For example, SMAPE values close to 0% to indicate excellent performance with minimal error, while values approaching 200% suggest poor performance. Our SMAPE values ranged between 16.86% and 42.81% (at the national level), 31.22% and 65.5% (at the local health board level) and was 86.23% the age level. Since our SMAPE values never exceeded the mid-point, it can be argued that the models perform well. Moreover, the error metrics at the national level are comparable to that of the UKHSA model for England whose projections had mean absolute percentage errors of 27.3%, 30.9% and 15.7% for COVID-19, influenza, and RSV respectively at the national level.

Overall, the bias scores of all models at the national level were not significantly different from zero. We confirmed this by performing a one-sample Wilcoxon test for each model. The bias scores were close to 1 during or near admission peaks, suggesting that the model overestimated admissions during these periods. However, since the model is primarily used for planning, it might be more prudent to have a positive bias and overestimate admissions. The error metrics are more valuable for relatively ranking different models.

Overall, the STPs model performed better at the national level when compared to local health board level across all diseases. This is evidenced by higher interval coverage and lower SMAPE values. The COVID-19 model had lower SMAPE values than the flu and RSV models, indicating more accurate projections for this acute respiratory infection. This difference could be due to the data fed into the model, as daily admissions were available for COVID-19, whereas only weekly data was available for RSV and flu. Following discussions with PHW, we now receive daily admissions data for all three infections and will reevaluate the models for the upcoming year.

The model output was made accessible to users by creating heatmaps that highlight health boards experiencing increases or decreases in admissions, a feature that users have found particularly insightful and useful (See Figure 11). Currently, the model uses a frequentist approach to calculate prediction intervals. However, we are transitioning to Bayesian methods that will provide more realistic estimates of model uncertainty by using posterior samples to generate quantile-based forecasts. Once implemented, we will adopt improved evaluation metrics, such as weighted interval scores, which combine over- and under-prediction with sharpness into a single measure to assess the model's performance. Additionally, we are exploring the incorporation of other predictors, such as Google Trends data for flu searches and GP influenza-like illness consultation rates, to better understand and predict admission growth.

Figure 11: Example of choropleth highlighting change in the number of admissions over the next two weeks, by health board



Source: SRE Calculations

Conclusion

The model's outputs aligned well with the historical ranges of hospital admissions. It performed better at predicting national-level data compared to health board-level data for COVID-19, influenza, and RSV. Although the model tended to overestimate admissions during peak periods, this overestimation can be advantageous for planning purposes, ensuring that health boards have adequate staff and resources to manage winter pressures. We have explored sharing outputs using heat maps to improve accessibility of findings. Moving forward, we plan to refine the model by adopting Bayesian methods, which will provide more realistic estimates of model uncertainty, and by implementing improved evaluation metrics. Additionally, we are exploring the inclusion of other predictors, such as Google Trends data for flu searches and GP influenza-like illness consultation rates, to better understand and forecast admission trends.

Appendix

The optimised number of basis functions for each model are listed below (For more details on the optimisation process, see methods):

- k1: number of basis functions at the national level
- k2: number of basis functions at the local health board level
- k3 (only for RSV Model): number of basis functions at the age level

Table S1: Optimal number of basis functions for the COVID-19 GAM model by projection date

Projection date	k1	k2
2023-09-03	6	3
2023-09-17	6	3
2023-10-01	5	3
2023-10-15	6	4
2023-10-29	4	4
2023-11-12	3	4
2023-11-26	5	3
2023-12-10	3	4
2023-12-24	7	3
2024-01-07	7	3
2024-01-21	7	5
2024-02-04	3	5
2024-02-18	4	6
2024-03-03	4	3
2024-03-17	4	3

Source: SRE calculations

Table S2: Optimal number of basis functions for the influenza GAM model by projection date

Projection date	k1	k2
2023-11-12	3	4
2023-11-26	3	4
2023-12-10	4	5
2023-12-24	4	6
2024-01-07	3	3
2024-01-21	3	3
2024-02-04	4	3
2024-02-18	6	4
2024-03-03	3	6
2024-03-17	5	5
2024-03-31	3	3
2024-04-14	3	3

Source: SRE calculations

Table S3: Optimal number of basis functions for the RSV GAM model by projection date

Projection date	k1	k2	k3
2023-09-17	8	5	3
2023-10-01	4	4	5
2023-10-15	4	5	3
2023-10-29	6	4	5
2023-11-12	8	3	3
2023-11-26	8	3	3
2023-12-10	3	4	5
2023-12-24	3	4	4
2024-01-07	3	4	3
2024-01-21	3	3	3
2024-02-04	8	4	3
2024-02-18	3	4	4
2024-03-03	8	5	4

Source: SRE calculations