

Supplementary Information

Multimodal AI for Bacterial vs Viral Upper Respiratory Tract Infection Multicentre Development and Internal Validation.

Sho Okiyama, Yuji Ariyasu, Hiroshi Yoshihara, Atsushi Fukuda, Masao Iwagami

Contents:

Supplementary Figure 1. Example pharyngeal image.

Supplementary Figure 2. SHAP summary plot.

Supplementary Table 1. Baseline characteristics.

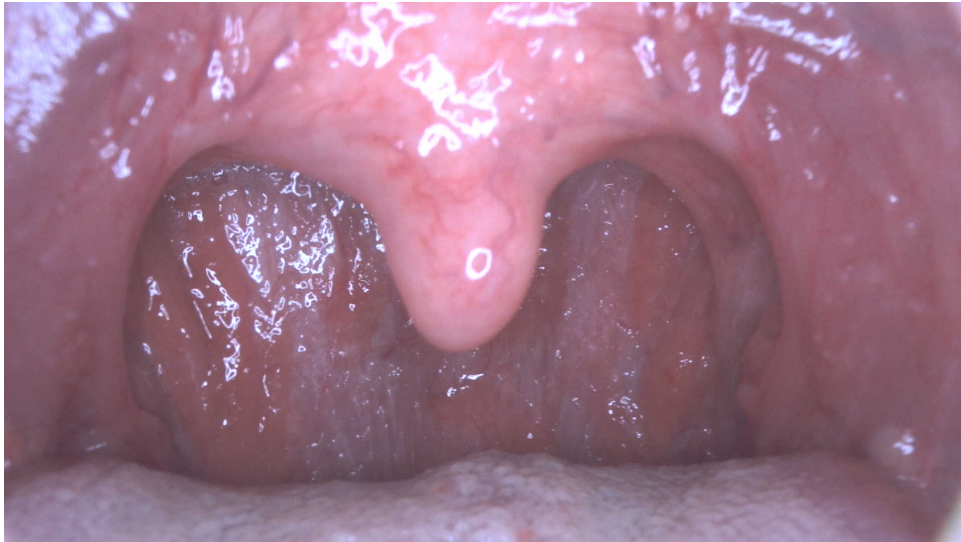
Supplementary Table 2. Unified definitions of clinical scores.

Supplementary Table 3. Antibiotic-use reduction simulation.

Supplementary Figure 1

Example pharyngeal image.

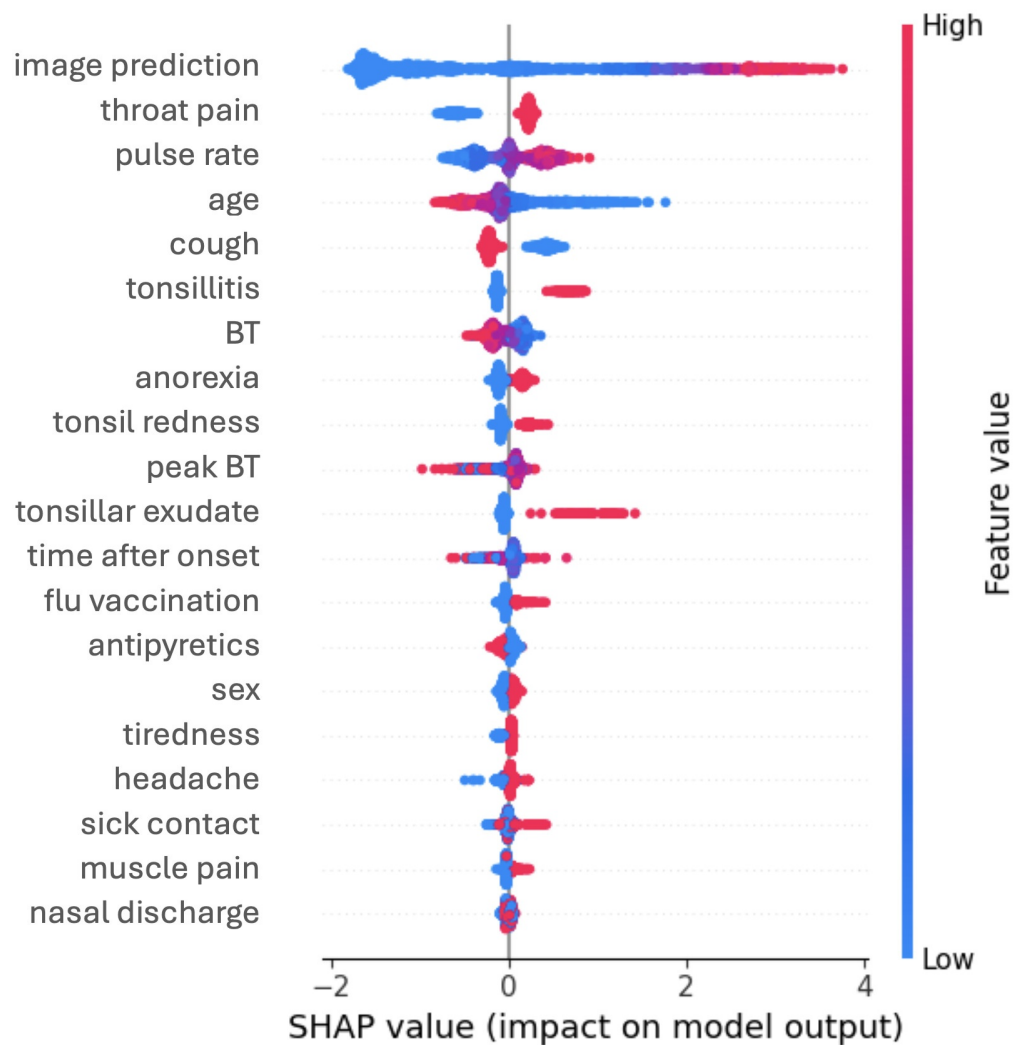
Representative image of the posterior pharyngeal wall captured by the pharyngeal camera system used in this study. Each visit contributed up to 18 images; the single best image was automatically selected using a dedicated model that prioritised focus and wide capture of the posterior pharyngeal wall. Patient-identifying information was not present in the raw images.



Supplementary Figure 2

SHAP summary plot for the fusion model.

SHAP summary plot showing the direction and magnitude of each feature's contribution to the fusion model's prediction of GAS likelihood. Features are ranked by mean absolute SHAP value. Red points indicate higher feature values, blue points lower feature values; horizontal position reflects the impact on the model output.



BT: body temperature

Supplementary Table 1

Baseline characteristics of the study population.

Characteristics	All (n=6666)	GAS positive (n=812)	GAS negative (n=5854)
Age (years), mean (SD)	32.2 (18.9)	20.0 (14.3)	33.9 (18.9)
Time from onset (hours), mean (SD)	27.0 (22.4)	26.4 (21.0)	27.1 (22.6)
Highest BT before visit (°C), mean (SD)	38.5 (0.8)	38.6 (0.7)	38.5 (0.8)
BT at visit (°C), mean (SD)	37.7 (0.9)	37.7 (0.9)	37.7 (0.9)
Pulse rate (bpm), mean (SD)	95.4 (19.2)	104.0 (19.8)	94.2 (18.8)
Male, n (%)	3292 (49.4%)	396 (48.8%)	2896 (49.5%)
Female, n (%)	3374 (50.6%)	416 (51.2%)	2958 (50.5%)
Close contact, n (%)	2316 (34.7%)	302 (37.2%)	2014 (34.4%)
Recent influenza vaccination, n (%)	1743 (26.1%)	194 (23.9%)	1549 (26.5%)
Use of antipyretics, n (%)	2899 (43.5%)	371 (45.7%)	2528 (43.2%)
Tiredness, n (%)	4904 (73.6%)	609 (75.0%)	4295 (73.4%)
Appetite loss, n (%)	2674 (40.1%)	436 (53.7%)	2238 (38.2%)
Chill, n (%)	3360 (50.4%)	442 (54.4%)	2918 (49.8%)
Sweating, n (%)	1467 (22.0%)	223 (27.5%)	1244 (21.3%)
Joint pain, n (%)	2911 (43.7%)	307 (37.8%)	2604 (44.5%)
Muscle pain, n (%)	2020 (30.3%)	245 (30.2%)	1775 (30.3%)
Headache, n (%)	4170 (62.6%)	545 (67.1%)	3625 (61.9%)
Nasal discharge/congestion, n (%)	3966 (59.5%)	448 (55.2%)	3518 (60.1%)
Cough, n (%)	4477 (67.2%)	420 (51.7%)	4057 (69.3%)
Sore throat, n (%)	4840 (72.6%)	639 (78.7%)	4201 (71.8%)
Digestive symptoms, n (%)	1009 (15.1%)	154 (19.0%)	855 (14.6%)
Tonsillitis, n (%)	1046 (15.7%)	381 (46.9%)	665 (11.4%)
Tonsillar exudate, n (%)	369 (5.5%)	178 (21.9%)	191 (3.3%)
Tonsillar redness, n (%)	1812 (27.2%)	523 (64.4%)	1289 (22.0%)

BT: body temperature. SD: standard deviation.

Supplementary Table 2

Unified definitions of clinical scores used for comparison.

Score	Components	Age adjustment	Range	Positive threshold (for specificity ≥ 0.90 comparison)	Temperature rule
Centor-3	Tonsillar exudate = 1; fever ≥ 38.0 °C = 1; absence of cough = 1	None	0–3	≥ 3	Either BT at visit ≥ 38.0 °C or peak BT after onset ≥ 38.0 °C
McIsaac-Proxy	Sum of Centor-3 components (as above)	Age 6–14: +1; 15–44: 0; ≥ 45 : -1	-1 to 4	≥ 3	Same as above

BT: body temperature. The Centor and McIsaac scores shown are proxy implementations adapted from the original definitions to match the variables available in this study; no swollen or tender anterior cervical lymph nodes were recorded systematically, so the Centor component for lymphadenopathy was omitted, yielding a 3-component 'Centor-3'. The 'McIsaac-Proxy' applies age adjustments to this 3-component score.

Supplementary Table 3

Antibiotic-use reduction simulation at the operational threshold (specificity ≥ 0.90 ; per 1,000 visits; baseline prescribing rate 30%).

GAS prevalence	Total prescriptions under AI	Reduction vs 30% (count)	Reduction vs 30% (%)	Unnecessary prescriptions (FP)	Missed cases (FN)	PPV	NPV	FPR	FNR
5%	133	167	55.7%	95	12	28.6%	98.6%	10.0%	24.0%
10%	166	134	44.7%	90	24	45.8%	97.1%	10.0%	24.0%
15%	199	101	33.6%	85	36	57.3%	95.5%	10.0%	24.0%
20%	232	68	22.6%	80	48	65.6%	93.8%	10.0%	24.0%

FP: false positive; FN: false negative; PPV: positive predictive value; NPV: negative predictive value; FPR: false positive rate; FNR: false negative rate. Baseline prescribing rate of 30% is a representative value consistent with published reports that ambulatory URTI/acute respiratory infection antibiotic prescribing is roughly 20–40% under conventional care. Model operating point fixed at the threshold satisfying specificity ≥ 0.90 , which corresponds to sensitivity 0.761 and specificity 0.900 on the held-out folds.