



Supplementary material

The supplementary material was provided by the authors and some information may not have been peer reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by the Hong Kong Academy of Medicine and the Hong Kong Medical Association. The Hong Kong Academy of Medicine and the Hong Kong Medical Association disclaim all liability and responsibility arising from any reliance placed on the content.

Supplement to: P He, Y Liang, Y Zou, et al. Development and optimisation strategies for a nomogram-based predictive model of malignancy risk in thyroid nodules. *Hong Kong Med J* 2026;Epub 30 Jan 2026. <https://doi.org/10.12809/hkmj2512718>.

SUPPLEMENTARY TABLE 1. Comparison of apparent and bias-corrected model performance based on bootstrap internal validation in the derivation cohort (n=909)

Performance metric	Original apparent performance	Bias-corrected performance	Difference (apparent-corrected)
C-statistic (AUC)	0.730	0.728	0.002
95% CI for AUC	(0.697-0.762)	(0.692-0.758)	N/A
Sensitivity	63.2%	62.5%	0.7%
Specificity	74.9%	74.1%	0.8%
Overall accuracy	67.7%	66.9%	0.8%
Positive predictive value	72.5%	71.8%	0.7%
Negative predictive value	65.9%	65.3%	0.6%

Abbreviations: 95% CI = 95% confidence interval; AUC = area under the receiver operating characteristic curve; N/A = not applicable

SUPPLEMENTARY TABLE 2. Expanded predictor distribution and univariate logistic regression odds ratios for malignancy (n=909)^{* †}

	Optimisation (n=551)	Unchanged (n=358)	Reference[§]	OR (95% CI)[‡]	P value
Sex					
Female	384 (69.7%)	269 (75.1%)	Ref = Female	1.00 (Ref)	N/A
Male	167 (30.3%)	89 (24.9%)	Reference	1.31 (0.97-1.77)	0.08
Age, y					
≤40	257 (46.6%)	162 (45.3%)	Ref >60 years	1.51 (0.95-2.40)	0.07
40-60	249 (45.2%)	153 (42.7%)	Ref >60 years	1.55 (0.97-2.47)	0.06
>60	45 (8.2%)	43 (12.0%)	Reference	1.00 (Ref)	N/A
Nodule size, mm					
≤10	297 (53.9%)	186 (51.9%)	Ref >30 mm	1.42 (0.89-2.27)	0.14
10-30	221 (40.1%)	122 (34.1%)	Ref >30 mm	1.58 (0.97-2.56)	0.06
>30	33 (6.0%)	50 (14.0%)	Reference	1.00 (Ref)	N/A
No. of nodules					
Multiple	161 (29.2%)	94 (26.3%)	Ref = Single	1.30 (0.96-1.76)	0.08
Single	390 (70.8%)	264 (73.7%)	Reference	1.00 (Ref)	N/A
Composition					
Solid	542 (98.3%)	301 (84.1%)	Ref = Mixed	10.12 (1.12-91.34)	0.04
Cystic	1 (0.2%)	8 (2.2%)	Ref = Mixed	1.30 (0.14-11.89)	0.81
Mixed	8 (1.5%)	49 (13.7%)	Reference	1.00 (Ref)	N/A
Echogenicity					

Isoechoic or hyperechoic	42 (7.6%)	29 (8.1%)	Reference	1.00 (Ref)	N/A
Hypoechoic	351 (63.7%)	215 (60.1%)	Ref = Isoechoic or hyperechoic	1.12 (0.86-1.86)	0.64
Very hypoechoic	158 (28.7%)	114 (31.8%)	Ref = Isoechoic or hyperechoic	0.95 (0.56-1.62)	0.87
Calcification type					
None	238 (43.2%)	164 (45.8%)	Reference	1.00 (Ref)	N/A
Micro	280 (50.8%)	178 (49.7%)	Ref = None	0.70 (0.37-1.32)	0.27
Coarse	33 (6.0%)	16 (4.5%)	Ref = None	0.76 (0.40-1.42)	0.39
Blurred margin					
Yes	389 (70.6%)	229 (64.0%)	Ref = Margin blur	0.73 (0.55-0.98)	0.04
No	162 (29.4%)	129 (36.0%)	Reference	1.00 (Ref)	N/A
Aspect ratio >1					
Yes	165 (29.9%)	131 (36.6%)	Ref <1	1.35 (1.01-1.79)	0.03
No	386 (70.1%)	227 (63.4%)	Reference	1.00 (Ref)	N/A
Abnormal cervical LN					
Yes	136 (24.7%)	30 (8.4%)	Ref = No	0.27 (0.18-0.42)	<0.01
No	415 (75.3%)	328 (91.6%)	Reference	1.00 (Ref)	N/A
C-TIRADS category					
3	6 (1.1%)	56 (15.6%)	Ref = 5	0.15 (0.05-0.40)	<0.01
4A	80 (14.5%)	27 (7.5%)	Ref = 5	4.14 (2.11-8.13)	<0.01
4B	114 (20.7%)	48 (13.4%)	Ref = 5	3.32 (1.79-6.14)	<0.01
4C	326 (59.2%)	192 (53.6%)	Ref = 5	2.37 (1.38-4.09)	<0.01
5	25 (4.5%)	35 (9.8%)	Reference	1.00 (Ref)	N/A

Abbreviations: 95% CI = 95% confidence interval; C-TIRADS = Chinese Thyroid Imaging Reporting and Data System; LN = lymph node; N/A = not applicable;

OR = odds ratio; Ref = Reference

* Data are presented as No. (%), unless otherwise specified. Percentages may not total 100% due to rounding. Missing data accounted for less than 2% of all predictors and were handled using complete-case analysis

† The dependent variable was pathological diagnosis (1 = malignant, 0 = benign). Multicategory predictors (age, nodule size, composition, echogenicity, calcification, and C-TIRADS) were analysed using dummy coding, with the category listed as 'Reference' serving as the baseline. Binary predictors (sex, number of nodules, margin, taller-than-wide shape, and suspicious lymph nodes) were included directly without dummy variables

‡ Reference categories were selected based on the lowest malignancy risk or the most clinically relevant baseline

§ Odds ratios were derived from univariate binary logistic regression models. An OR >1 indicates an increased risk of malignancy relative to the reference category, whereas an OR <1 indicates a reduced risk

|| Wald tests of the individual regression coefficients

SUPPLEMENTARY TABLE 3. Adjusted odds ratios from the multivariable binary logistic regression model for predicting Chinese Thyroid Imaging Reporting and Data System optimisation*

Predictor	Adjusted OR (95% CI)	P value
Sex		
Male vs female	1.37 (0.98-1.91)	0.058
Age, y		
≤40	1.26 (0.73-2.18)	0.394
40-60	1.72 (1.00-2.97)	0.049
>60	1.00 (Ref)	N/A
Nodule size, mm		
Per 1-mm increase	1.01 (1.00-1.03)	0.047
No. of nodules		
Multiple vs single	1.69 (1.20-2.38)	0.002
Abnormal cervical lymph nodes		
Yes vs no	3.74 (2.36-5.95)	<0.01
Original C-TIRADS category		
3	0.12 (0.04-0.39)	<0.01
4A	5.54 (2.68-11.47)	<0.01
4B	4.52 (2.34-8.70)	<0.01
4C	2.62 (1.47-4.66)	<0.01
5	1.00 (Ref)	N/A

Abbreviations: 95% CI = 95% confidence interval; C-TIRADS = Chinese Thyroid Imaging Reporting and Data System; N/A = not applicable; OR = odds ratio; Ref = reference

* Model performance: area under the receiver operating characteristic curve=0.73 (95% CI=0.69-0.76); P<0.01

SUPPLEMENTARY TABLE 4. Diagnostic performance measures at various risk thresholds of the final prediction model (derivation population)

Risk threshold for C-TIRADS optimisation	Sensitivity	Specificity	Accuracy	Predictive value (95% CI)		Likelihood ratio (95% CI)	
	(95% CI)	(95% CI)		Positive	Negative	Positive	Negative
≥30%	0.96 (0.94-0.98)	0.17 (0.13-0.21)	0.654	0.64 (0.63-0.65)	0.77 (0.67-0.85)	1.17 (1.11-1.23)	0.18 (0.11-0.31)
≥50%	0.88 (0.85-0.91)	0.38 (0.33-0.43)	0.688	0.68 (0.66-0.70)	0.68 (0.62-0.74)	1.43 (1.31-1.56)	0.29 (0.22-0.38)
≥60%	0.67 (0.63-0.71)	0.66 (0.61-0.71)	0.671	0.75 (0.72-0.78)	0.57 (0.53-0.60)	2.01 (1.72-2.35)	0.48 (0.42-0.56)
≥64%	0.75 (0.70-0.79)	0.63 (0.58-0.67)	0.677	0.79 (0.76-0.82)	0.57 (0.54-0.59)	2.51 (2.08-3.04)	0.49 (0.43-0.55)
≥80%	0.24 (0.21-0.28)	0.95 (0.92-0.97)	0.525	0.89 (0.83-0.93)	0.45 (0.43-0.46)	5.52 (3.34-9.11)	0.78 (0.74-0.83)

Abbreviations: 95% CI = 95% confidence interval; C-TIRADS = Chinese Thyroid Imaging Reporting and Data System

SUPPLEMENTARY TABLE 5. Diagnostic performance of the final prediction model at various risk thresholds in the external validation population

Risk threshold for C-TIRADS optimisation	Sensitivity	Specificity	Accuracy	Predictive value (95% CI)		Likelihood ratio (95% CI)	
	(95% CI)	(95% CI)		Positive	Negative	Positive	Negative
	≥30%	0.91 (0.88-0.93)	0.54 (0.47-0.60)	0.782	0.79 (0.76-0.81)	0.75 (0.69-0.81)	1.97 (1.72-2.25)
≥50%	0.87 (0.84-0.90)	0.64 (0.57-0.69)	0.793	0.82 (0.79-0.84)	0.72 (0.67-0.77)	2.40 (2.04-2.83)	0.19 (0.15-0.25)
≥60%	0.85 (0.81-0.88)	0.69 (0.62-0.74)	0.796	0.83 (0.81-0.86)	0.71 (0.66-0.75)	2.71 (2.26-3.26)	0.21 (0.17-0.26)
≥74%	0.71 (0.66-0.75)	0.87 (0.82-0.91)	0.766	0.91 (0.88-0.94)	0.61 (0.57-0.64)	5.56 (4.02-7.68)	0.33 (0.28-0.38)
≥80%	0.66 (0.61-0.70)	0.89 (0.85-0.93)	0.742	0.92 (0.89-0.94)	0.58 (0.54-0.61)	6.55 (4.52-9.48)	0.37 (0.33-0.43)

Abbreviations: 95% CI = 95% confidence interval; C-TIRADS = Chinese Thyroid Imaging Reporting and Data System

SUPPLEMENTARY FIG. (a-c) Construction and evaluation of the predictive model for optimising Chinese Thyroid Imaging Reporting and Data System (C-TIRADS) classification. (a) Receiver operating characteristic curve illustrating diagnostic performance in the derivation cohort. (b) Calibration curve evaluating agreement between predicted and observed outcomes in the derivation cohort. (c) Receiver operating characteristic curve assessing discriminative ability of the predictive model for radiologist-optimised C-TIRADS classification in the independent validation cohort. (d) Calibration curve demonstrating predictive accuracy and goodness-of-fit of the model in the validation cohort

