Odontology (2023) 111:178–191
https://doi.org/10.1007/s10266-022-00716-6

ORIGINAL ARTICLE

External validation of cone-beam computed tomography- and panoramic radiography-featured prediction models for inferior alveolar nerve injury after lower third molar removal: proposal of a risk calculator

Seiko Kubota[1,2] · Tomoaki Imai[1,3] · Ayano Nishimoto[1] · Shigeki Amekawa[2] · Narikazu Uzawa[1]

Received: 26 February 2022 / Accepted: 26 April 2022 / Published online: 23 May 2022

Abstract

We previously developed basic and extended models to predict inferior alveolar nerve injuries (IANI) after lower third molar (LM3) removal based on cone-beam computed tomography (CBCT) images. Although these models comprised predictors, including increased age and inferior alveolar canal-related CBCT factors, external validations were lacking. Therefore, this study externally validated these models and compared them with other related models based on their performance. Original and newly validated samples included patients who underwent LM3 removal following CBCT. Subsequently, 39 and 25 patients with IANI, then 457 and 295 randomly selected patients without IANI were chosen of the observed 1573 and 1052 patients, respectively. CBCT- and panoramic radiograph (PAN)-featured models were validated. Then, models' discrimination and calibration abilities were assessed using C-statistics and calibration plots, respectively. Brier scores were also quantified, after which logistic recalibration was achieved to optimize calibration, and a risk calculator was developed. During the external validation, the extended model exhibited the best C-statistic (0.822) and Brier score (0.064), whereas two CBCT- and two PAN-featured models showed lower performances with C-statistics (0.764, 0.706, 0.584, and 0.627) and Brier scores (0.069, 0.074, 0.075, and 0.072). Besides, all models showed a tendency to overpredict its high-risk range. However, recalibration of the extended model resulted in excellent calibration performance. CBCT-featured models, especially the extended model, conclusively showed a superior predictive performance to PAN models. Therefore, the risk calculator on the extended CBCT model is proposed to be a clinical decision-aid tool that preoperatively predicts IANI risk.

Keywords Cone-beam computed tomography · Tooth extraction · Third molar surgery · Risk calculator · Inferior alveolar canal · Prediction model

Introduction

During the lower third molar (LM3) extraction, inferior alveolar nerve injury (IANI) infrequently occurs with an incidence of 0.4–8.4% [1–3]. The resulting sensory disturbance of the lower lip, in addition to the mental area, frequently compromises patients' quality of life during a protracted period [4, 5]. Hence, to predict the possibility of IANI, oral surgeons are requested to preoperatively evaluate imaging findings, especially the proximity between LM3s' roots and the inferior alveolar canal (IAC) [3, 6–8]. Rood and Shehab [6] have presented seven signs for assessing the IANI risk on panoramic radiography (PAN) and concluded these three signs as the risk indicators, including diversion of IAC, darkening of the root, and an interruption of the white line.

Seiko Kubota and Tomoaki Imai contributed equally as the first author.

✉ Tomoaki Imai
hsc12@hotmail.com

[1] Department of Oral and Maxillofacial Surgery II, Osaka University Graduate School of Dentistry, Suita, Osaka, Japan

[2] Department of Oral and Maxillofacial Surgery, Ikeda City Hospital, Ikeda, Osaka, Japan

[3] Department of Oral and Maxillofacial Surgery, Toyonaka Municipal Hospital, 4-14-1 Shibaharatyo, Toyonaka, Osaka 560-8565, Japan

However, the overlap between LM3 and IAC on PAN does not necessarily indicate the direct contact of LM3 to the inferior alveolar nerve through the IAC cortical wall defect. Therefore, cone-beam computed tomography (CBCT) scan should provide higher-dimensional information on the anatomical relationship than PAN findings. Yet, the application of CBCT is not recommended for routine examinations [9].

Recently, Matzen et al. [10] focused on predictive PAN and CBCT value findings associated with IANI risk in the same group of patients. However, no studies have directly compared the predictive performance of IANI-risk models based on the imaging features of those modalities. We previously reported three predictors of IANI following LM3 surgery: the increase in age above 30 years and two CBCT-related findings, which included the lingual or inter-radicular IAC position to LM3 and the multiple root contacts with the perforated wall of IAC [11]. Additionally, our subsequent study [12] developed an extended CBCT model with five predictors, including two additional CBCT findings: the length of the perforated IAC (> 3.4 mm) and the coronal position of IAC on the enlarged root (Adj-En type). Nevertheless, although the extended model exhibited good discrimination and calibration [12], it remains unsuitable as a predicting tool in clinical settings due to the lack of external validation parameters in independent populations [13–15]. Hence, this study externally validated IANI-risk prediction models containing CBCT-related factors and compared their performances with models possessing PAN features. We also developed an IANI-risk calculator that quickly estimated its preoperative probability.

Materials and methods

Study design and sample

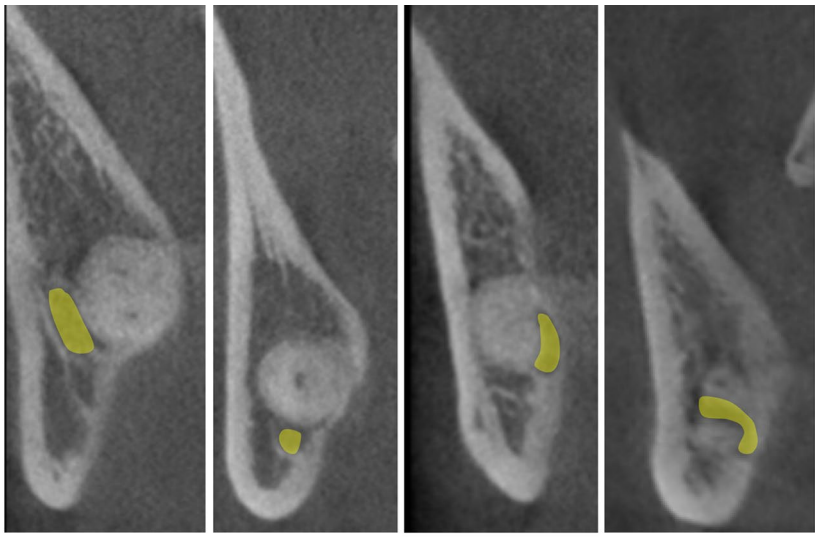
The internal review boards of our institution approved this project, conducted following the principles of the Helsinki Declaration. This retrospective case–control study enrolled patients who underwent LM3 removal after preoperative CBCT scanning from January 2015 to March 2021 at Ikeda City Hospital for external validation, with the original group of our previous study, conducted at Osaka University [12]. Candidates for preoperative CBCT scanning after PAN were those whose radiographs showed a close relationship between IAC and LM3 [6]. Subsequently, CBCT and PAN images were obtained from the validation samples using AUGE SOLIO Z CM (Asahi Roentgen Ind. Co., Ltd, Kyoto, Japan). While exposure parameters for PAN were 72 kV, 12 mA, and a 12-s exposure time with a pixel size of 120 mm, CBCT slice thickness was 0.31 mm (scanning time, 17 s; tube voltage, 85 kV; tube current, 6–8 mA; voxel size, 0.1–0.315 mm; the field of view size, 51–161 mm × 55–100 mm). Regarding the original samples, CBCT scanning was conducted using Alphard-3030 (Alphard VEGA®; Asahi Roentgen Ind. Co., Ltd, Kyoto, Japan), with a slice thickness of 0.65 mm (scanning time, 17 s; tube voltage, 80 kV; tube current, 2–15 mA; voxel size, 0.2 mm; and a field of view size, 102 × 102 mm) [11]. PAN images were obtained using Hyper-X unit (Asahi Roentgen Ind. Co., Ltd, Kyoto, Japan) with exposure parameters of 64 kV, 8 mA, and a 12-s exposure time and a pixel size of 96 mm.

The case group (patients with IANI) corresponded to patients with a documented post-operative neurosensory disturbance of the lower lip and chin area. The control group was randomly selected from a sample of patients with no occurrence of IANI [16, 17]. Exclusion criteria were patients < 16 years; LM3s involving adjacent lesions, such as cysts or tumors; and unavailability of post-operative follow-up at one week data. Subsequently, either side was selected randomly during bilateral surgeries with the same outcome (i.e., bilaterally IANI positive or negative). In contrast, the affected side of IANI was sampled for those with dual results [11, 12].

Study variables

Data on patients' characteristics, surgical situations (surgeon experience and extraction side), and five factors accounting for IAC-related CBCT findings were collected [11, 12]. The buccolingual localization of the IAC to LM3 on the coronal sections was then categorized into the buccal/inferior and lingual/inter-radicular positions [18] (Fig. 1). Afterward, coronally sectioned IAC profiles closest to LM3 were classified into dumbbell and non-dumbbell (round/oval/teardrop) shapes [19], after which the number of roots close to the IAC with cortical perforation was counted [11]. Subsequently, the most extended length of IAC's cortical bone defective on a single coronal CBCT slice in contact with LM3 roots was measured thrice and averaged [20]. Unlike a previous study [12], length as a continuous variable was not converted to a dichotomous variable. The relationship between the perforated IAC position and the adjacent root's shape was classified into the adjacent to the enlarged part (Adj-En) of the root and the non-Adj-En part [12, 21]. Next, the relationship of the Adj-En type, corresponding to the narrowest part of the IAC with any cortical defect located at or coronal to the enlarged portion of the root on the sagittal plane, was evaluated. We also assessed cases when the IAC with cortical defects was located at or coronal to the enlarged portion of the root on the horizontal plane. The cut-off value for age was 30 years in line with previous studies [11, 19, 22].

Additional variables in this study were features of PAN using the Pell–Gregory classification and the seven Rood signs. Class I/II/III indicated the horizontal position of LM3



IAC position	Buccal	Inferior	Lingual	Inter-radicular
IAC shape	Round/oval	Round/oval	Teardrop	Dumbbell
Root contacting perforated IAC	None	None	Single	Multiple

Fig. 1 Parameters related to cone-beam computed tomography (CBCT) imaging on the coronal section. The buccolingual position of the inferior alveolar canal (IAC) to the root of the lower third molar is classified into the buccal, inferior, lingual, and inter-radicular positions. The IAC shape is classified into round/oval, teardrop, and dumbbell in the closest proximity to the root. The number of roots contacting the perforated IAC is counted as none, single, and multiple. Indirect contact between the inferior alveolar nerve and the root is categorized as "none," as shown in the left and the second panels

relative to the anterior border of the ramus. Position A/B/C showed the vertical relationship to the mandibular occlusal plane. Rood's signs included were as follows [6]: darkening of the root (loss of root density impinged over by the canal), interruption of the white line (discontinuity of the radio-opaque line of the IAC), diversion of the canal (direction shift of the IAC at the root's crossing), deflection of the root (abrupt deviation of roots near the IAC), narrowing of the root where the IAC crosses, narrowing of the canal (a sharp decrease in the width of the canal while it crosses the root), and the presence of a dark/bifid root apex (loss of root density that is intruded upon by the IAC with bifid root apex). Then, to simplify these signs, three former PAN findings were defined as the significant Rood's signs [6, 17, 23]. Afterward, we focused on the crucial signs, at least one. Two oral surgeons independently assessed images. Any conflicts in the evaluations were decided by consensus [12].

The outcome was defined as the presence (case group) or absence (control group) of IANI at one week postoperatively. Sensory perception was also evaluated via subjective monitoring of the feeling at the lower lip and chin area. In cases with suggestive neurosensory disturbance, the pin-prick, light touch, or Semmes–Weinstein test was additionally conducted [11]. In the IANI-positive group, the outcome was postoperatively assessed by follow-up at 1, 3, and 6 months. Neurosensory disturbance detected at 6 months postoperatively was considered permanent [11].

Statistical analyses

We reported descriptive analyses as frequencies (%) for categorical variables, median (quartile) for continuous variables, and analyzed the demographic characteristics of patients in the original and validation groups using the χ^2 test. Additionally, univariate logistic regression analysis was conducted to estimate relationships between the IANI and variables, after which we calculated their odds ratios (ORs) and 95% confidence intervals (CIs).

Our extended and basic CBCT models included five predictors (the lingual/inter-radicular IAC position, multiple roots contacted with the perforated IAC, higher age (> 30 years), root shape of the Adj-En type, and the length of the perforated IAC) and three latter variables, respectively [12]. Then, to compare our models with another CBCT-featured model, we focused on the combined high-risk factor for IANI reported by Tachinami et al. [24]. Next, we independently employed these components: the lingual position of the IAC to LM3, the dumbbell shape of the IAC at the

coronal section, and the presence of IAC's cortical bone defect on the root, after which these factors were simultaneously applied to the multivariate analysis referred to as the Tachinami's triads' model.

For PAN-featured models, variables other than CBCT-related factors were simultaneously used as multivariate models to select predictors using a backward stepwise algorithm based on the P-value < 0.05. However, for the simplified panoramic model, all Rood's signs [6] were distilled to reflect major Rood's signals, then multivariate analysis was conducted as aforementioned. Subsequently, these models based on the original sample ($n = 496$) were transported to external validation samples ($n = 320$), with the coefficients calculated in the original model. Then, the validities of various models in both samples were evaluated using metrics of discrimination and calibration [14, 25]. The discrimination was assessed using concordance (C) statistics. This value is equivalent to the area under the receiver operating characteristic (ROC) curve for binary outcomes. Besides, while 0.5 or lower was considered a non-discriminatory model, 0.8 or higher implied good discrimination. Finally, the calibration was assessed graphically using a calibration plot to identify differences between the predicted and observed probability of IANI. A well-calibrated model displays the plotted curve on or near the ideal reference line [25, 26].

During the external validation, the model calibration was also evaluated using a calibration-in-the-large and a calibration slope, with the proposal for a better clinical prediction model [14, 15]. The calibration-in-the-large indicated whether predicted probabilities were entirely low (positive value) or high (negative value) by comparing the observed and expected outcomes of the model in a logistic regression model. The calibration slope was the regression slope of the linear predictor, indicating whether the model was overfitted (negative value) or under fitted (positive value) [14]. Perfect calibration was yielded when the calibration-in-the-large was 0 and the calibration slope was 1. However, in developing prediction models using standard logistic regression, the calibration intercept and slope were, by definition, 0 and 1 on the original cohort [25]. Therefore, the model's overall performance was tested using Brier scores, measuring the mean squared difference between the predicted probabilities for IANI and the observed prevalence of IANI. The lower this score, the lesser the difference and the better the predictions of IANI are calibrated.

Furthermore, the validated model exhibiting the best performance was recalibrated to improve the consistency of the predicted and observed probabilities. The model was updated through logistic recalibration [27], correcting the linear predictor's intercept and regression coefficients using a single adjustment factor. This method can be employed when the coefficient of the original model is proposed to have been overfitted [27]. The performance of the recalibrated model was then evaluated using those metrics as described above.

In logistic regression models, the probability of IANI in an individual patient (p_i) can be calculated with the formula of $p_i = 1/(1 + e^{-\text{linear predictor}})$. Therefore, the formula of the prediction model is as follows:

$$\begin{aligned}\text{Linear predictor}_{\text{original}} &= \log\left(p_i/1 - p_i\right) = \beta_0 + \beta_1 * \text{predictor}_1 \\ &\quad + \ldots + \beta_n * \text{predictor}_n \\ &\quad \left(\beta_0, \text{intercept}; \beta_1, \ldots \beta_n, \text{regression coefficients}\right).\end{aligned}$$

Additionally, logistic recalibration updated the original model by multiplying with the coefficient ($\beta_{\text{recalibration}}$) and adding the estimated intercept as follows [27]:

$$\begin{aligned}\text{Linear predictor}_{\text{recalibration}} &= \text{updated intercept} + \\ &\quad \left(\beta_{\text{recalibration}} * \text{linear predictor}_{\text{original}}\right).\end{aligned}$$

Hence, since p_i indicates the probability based on the IANI prevalence (π) in this case–control study, with all cases and the randomly sampled control, p_i was adjusted using the overall prevalence (π_0) against the whole sample. Subsequently, this probability defined as p'_i was calculated using the formula below [28]

$$p'_i = p_i(1-\pi)\pi_0 / \left[\left(1-p_i\right)\pi\left(1-\pi_0\right) + p_i(1-\pi)\pi_0\right].$$

Statistical significance was considered at $p < 0.05$. Statistical analyses were also conducted using the R 4.0.2 software program (R Foundation for Statistical Computing, Vienna, Austria). Then, the IANI-risk calculator was constructed using Excel (Microsoft).

Results

Table 1 summarizes the demographic, clinical, and imaging characteristics of the original and validation samples. Of the 1573 and 1052 patients subjected to the LM3 surgery following CBCT scanning, 39 (overall prevalence of 2.5%) and 25 (2.4%) patients lapsed into neurosensory disturbance of the lower lip and mental area one week postoperatively. The control group included 457 and 295 patients, randomly selected from the original and validation groups. Therefore, the variable distribution of the operator, the IAC position and CBCT root shape findings, the position of the Pell–Gregory classification, and the narrowing canal of PAN findings significantly differed between original and validation samples.

Table 1 Descriptive preoperative characteristics of the original and external validation samples

Variable	Original sample[a] Osaka Univ $n=496$	Validation sample Ikeda Hosp. $n=320$	p value
IANI (−): control	457 (92.1)	295 (92.2)	
IANI (+): case	39 (7.9)	25 (7.8)	
Patient background			
Age (years)			
≤30	222 (44.8)	123 (38.4)	0.087
>30	274 (55.2)	197 (61.6)	
Gender			
Men	217 (43.8)	122 (38.1)	0.129
Women	279 (56.2)	198 (61.9)	
Situations of surgery			
Operator			
Resident	190 (38.3)	32 (10.0)	<0.001
Fellow	306 (61.7)	288 (90.0)	
Extraction side			
Left	240 (48.4)	144 (45.0)	0.382
Right	256 (51.6)	176 (55.0)	
CBCT findings			
Root shape and IAC contact			
Non-Adj-En type	433 (87.3)	255 (79.7)	0.005
Adj-En type	63 (12.7)	65 (20.3)	
IAC position			
Buccal/inferior	419 (84.5)	242 (75.6)	0.002
Ligual/inter-radicular	77 (15.5)	78 (24.4)	
IAC shape			
Non-dumbbell	446 (89.9)	292 (91.2)	0.611
Dumbbell	50 (10.1)	28 (8.8)	
No. of roots with perforated IAC			
Intact	187 (37.7)	97 (30.3)	0.078
Single	220 (64.2)	153(47.8)	
Multiple	89 (17.9)	70 (21.9)	
Length of perforated IAC wall (mm)	2.22 [0.00, 3.60]	2.47 [0.00, 3.31]	0.704
PAN findings			
Pell-Gregory classification			
Class			
I/II	410 (82.7)	263 (82.2)	0.937
III	86 (17.3)	57 (17.8)	
Position			
A/B	449 (90.5)	273 (85.3)	0.030
C	47 (9.5)	47 (14.7)	
Rood's sign			
Dark and bifid apex of root	462 (93.1)	304 (95.0)	0.353
	34 (6.9)	16 (5.0)	
Darkening of root[b]	390 (78.6)	254 (79.4)	0.867
	106 (21.4)	66 (20.6)	
Deflection of root	488 (98.4)	313 (97.8)	0.742
	8 (1.6)	7 (2.2)	
Diversion of canal[b]	450 (90.7)	277 (86.6)	0.081
	46 (9.3)	43 (13.4)	

Table 1 (continued)

Variable	Original sample[a] Osaka Univ $n=496$	Validation sample Ikeda Hosp. $n=320$	p value
Interrunption of white line of canal[b]	265 (53.4)	156 (48.8)	0.217
	231 (46.6)	164 (51.2)	
Narrowing of canal	440 (88.7)	303 (94.7)	0.005
	56 (11.3)	17 (5.3)	
Narrowing of root	487 (98.2)	317 (99.1)	0.473
	9 (1.8)	3 (0.9)	
Presense of any major Rood's signs	159 (32.1)	87 (27.2)	0.161
	337 (67.9)	233 (72.8)	

In each component of Rood's sign, the upper and lower lines indicate the absence and presence of its indicator, respectively

IAC inferior alvolar canal, *IANI* inferior alveolar nerve injury, *PAN* panoramic radiography

[a]Variables other than PAN-related findings in the original sample have been shown in the previous study[12]

[b]Three major Rood's signs associated with IANI [6]

Predictive factors for IANI of CBCT- and PAN models

Table 2 shows univariate and multivariate logistic regression analyses results. As shown during the multivariate analyses, our extended model included IANI predictors, as follows: increased age (OR 2.25; 95% CI 0.89–5.68; $p=0.086$), the IAC position (OR 3.96; 95% CI 1.60–8.48; $p=0.002$), multiple roots close to the perforated IAC (OR 2.65; 95% CI 1.07–6.54; $p=0.034$), length of the perforated IAC (OR 1.31; 95% CI 1.05–1.64; $p=0.016$), and the Adj-En type (OR 3.37; 95% CI 1.47–7.72; $p=0.004$). Additionally, variables for the basic model were restricted to those reported previously [11]: increased age (OR 3.52; 95% CI 1.49–8.30; $p=0.004$), the IAC position (OR 5.26; 95% CI 2.34–11.8; $p<0.001$), and multiple roots (OR 4.30; 95% CI 1.90–9.71; $p=0.001$). Tachinami's triads model consisted of the IAC position (OR 6.18; 95% CI 2.92–13.1; $p<0.001$), the dumbbell shaped IAC (OR 2.80; 95% CI 1.24–6.33; $p=0.014$), and the IAC's cortical bone defect (OR 3.31; 95% CI 0.93–11.8; $p=0.064$).

Therefore, to construct PAN-featured models, variables other than CBCT-featured ones were included into stepwise backward logistic models. Results showed that the following variables were predictors of IANI in the Rood's model: increased age (OR 3.62; 95% CI 1.51–8.66; $p=0.039$), darkening of the root (OR 3.83; 95% CI 1.40–10.5; $p=0.009$), interruption of the white line (OR 5.77; 95% CI 2.27–14.6; $p<0.001$), diversion of the canal (OR 2.57; 95% CI 1.04–6.37; $p=0.042$), and deflection of the root (OR 10.7; 95% CI 1.57–72.7; $p=0.015$). Meanwhile, simplified PAN model contained three predictors: increased age (OR 3.03; 95% CI 1.33–6.90; $p=0.008$), Position C of the Pell–Gregory classification (OR 2.42; 95% CI 1.04–5.66; $p=0.041$), and the presence of major Rood's signs (OR 2.45; 95% CI 1.05–5.75; $p=0.039$).

Performance of predictive models

The predictive performance of models constructed from the original and validation groups are shown in Table 3. As indicated, the C-statistics showing discrimination ability in all the CBCT-featured models represented superior performances to the PAN-featured models (Fig. 2), with the best values of both C-statistics and Brier score being in the extended model. Furthermore, calibration plots in all models except for the basic model were shown close to the ideal calibration line (Fig. 3).

Results also showed that fitting validation data to the models comprised coefficients and intercepts from the original data. However, all models showed inferior performances to the original with less discrimination (lower C-statistics), whereas shifted calibration plots from the ideal line tended to overestimate (Fig. 4). Of those, the extended CBCT model represented good discrimination (> 0.8) of 0.822, with the best calibration performance and closest values to the zero ideal in the calibration-in-the-large and one in the calibration slope. After extended model recalibration by correcting the intercept and coefficients, the calibration plot fell near the ideal 45-degree line. Additionally, an improved Brier score of 0.062 from 0.064, a calibration-in-the-large score of 0 from − 0.510, and calibration slope of 1 from 0.833 were observed.

Table 2 Logistic regression models to measure the association between inferior alveolar injury and study variables developed in the original sample

Variable	Univariate		Multivariate									
			CBCT-featured model[a]						PAN-featured model[b]			
			Extended model [12]		Basic model [12]		Tachinami's triads [24]		Rood's model		Simplified mode	
	OR (95% CI)	*p* value	OR (95% CI)	*p* value	OR (95% CI)	*p* value	OR (95% CI)	*p* value	OR (95% CI)	*p* value	OR (95% CI)	*p* value
Patients' demographics												
Age												
> 30 years (vs ≤ 30 years)	3.41 (1.54–7.58)	0.003	2.25 (0.89–5.68)	0.086	3.52 (1.49–8.30)	0.004			3.62 (1.51–8.66)	0.039	3.03 (1.33–6.90)	0.008
Gender												
Female (vs male)	1.13 (0.58–2.19)	0.721										
Situations of surgery												
Operator												
Fellow (vs resident)	1.44 (0.71–2.91)	0.315										
Extraction side												
Right (vs left)	0.56 (0.29–1.10)	0.090										
CBCT finginds												
Root shape and IAC contact												
Adj-En type (vs the others)	8.92 (4.43–18.0)	< 0.001	3.37 (1.47–7.72)	0.004					Excluded		Excluded	
IAC position												
Ligual/inter-radicular (vs buccal/inferior)	10.70 (5.34–21.6)	< 0.001	3.69 (1.60–8.48)	0.002	5.26 (2.34–11.8)	< 0.001	6.18 (2.92–13.1)	< 0.001	Excluded		Excluded	
IAC shape												
Dumbbell (vs non-dumbbell)	6.55 (3.13–13.7)	< 0.001					2.80 (1.24–6.33)	0.014	Excluded		Excluded	
Contact of perforarated IAC												
Presence (vs absence)	8.09 (2.45–26.6)	< 0.001					3.31 (0.93–11.8)	0.064	Excluded		Excluded	
No of roots with perforated IAC												
Multiple (vs none/single)	8.52 (4.27–17.0)	< 0.001	2.65 (1.07–6.54)	0.034	4.30 (1.90–9.71)	0.001			Excluded		Excluded	
Length of perforated IAC wall	1.74 (1.46–2.08)	< 0.001	1.31 (1.05–1.64)	0.016					Excluded		Excluded	
PAN fingings												
Pell&Gregory classification												
Class III (vs I/II)	2.30 (1.12–4.75)	0.024										
Position C (vs A/B)	3.31 (1.46–7.48)	0.004									2.42 (1.04–5.66)	0.041

Table 2 (continued)

Variable	Univariate		Multivariate									
			CBCT-featured model[a]						PAN-featured model[b]			
			Extended model [12]		Basic model [12]		Tachinami's triads [24]		Rood's model		Simplified mode	
	OR (95% CI)	*p* value	OR (95% CI)	*p* value	OR (95% CI)	*p* value	OR (95% CI)	*p* value	OR (95% CI)	*p* value	OR (95% CI)	*p* value
Rood's signs												
Dark and bifid apex of root	0.71 (0.17–3.11)	0.658									Excluded	
Darkening of root[c]	1.30 (0.61–2.75)	0.499							3.83 (1.40–10.5)	0.009	Excluded	
Deflection of root	4.06 (0.79–20.8)	0.093							10.7 (1.57–72.7)	0.015	Excluded	
Diversion of canal[c]	2.85 (1.22–6.63)	0.015							2.57 (1.04–6.37)	0.042	Excluded	
Interrunption of white line of canal[c]	3.18 (1.55–6.55)	0.002							5.77 (2.27–14.6)	< 0.001	Excluded	
Narrowing of canal	0.89 (0.30–2.61)	0.832									Excluded	
Narrowing of root	N.A.	N.A.										
Presense of any major Rood's signs[c]	2.28 (0.98–5.28)	0.055							Excluded		2.45 (1.05–5.75)	0.039

CI confidence interval, *IAC* inferior alveolar canal, *OR* odds ratio, *PAN* panoramic radiography

[a]Variables are restricted to three or five predictors reported in previous studies

[b]Variables other than CBCT-related fidings are simultaneously subjected and selected by the backward stepwise analysis

[c]Three major Rood's signs [6] are integrated for simplification

Table 3 Models' performance in the original and validation samples

Models	C-statistic	Calibration-in-the-large	Calibration slope	Brier score
Original sample				
Extended model (5 factors)	0.878	NA[a]	NA[a]	0.055
Basic model (3 factors)	0.850	NA[a]	NA[a]	0.062
Tachinami's triads	0.803	NA[a]	NA[a]	0.061
PAN Rood's model	0.766	NA[a]	NA[a]	0.067
PAN simplified model	0.692	NA[a]	NA[a]	0.069
Validation sample				
Extended model (5 factors)	0.822	− 0.510	0.833	0.064
Basic model (3 factors)	0.764	− 0.60 3	0.802	0.069
Tachinami's triads	0.706	− 1.047	0.590	0.074
PAN Rood's model	0.584	− 1.624	0.326	0.075
PAN simplified model	0.627	− 1.192	0.535	0.072
Recalibrated extended model	0.821	0	1	0.062

A value of C-statistic > 0.8 is considered to be good performance of discrimination

Calibration-in-the-large and calibration slope should be as close to 0 and 1 as possible, respectively

A lower Brier score indicates better performance

[a]In prediction model developed using standard logistic regression, the calibration-in-the-large and calibration slope are, by definition, 0 and 1 on the original sample, respevtively

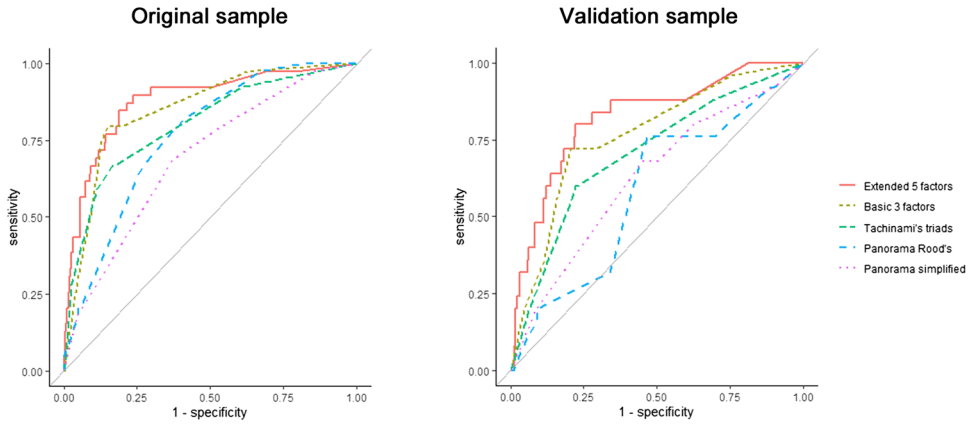


Fig. 2 Comparison of receiver operating characteristic (ROC) curves. Left and right panels, showing ROC curves in the original and validation samples, respectively. The gray 45° line is the line of no discrimination

User-friendly calculator of IANI risk after recalibration

The probability of IANI (p_i) in an individual patient was calculated using the formulas below

$$p_i = 1/\left(1 + e^{-\text{linear predictor}}\right), \text{ and linear predicter}_{\text{original}}$$
$$= -4.9634 + 0.8109 * X_1 + 1.3054 * X_2 + 0.9745 * X_3 + 1.2153 * X_4 + 0.2730 * X_5,$$

where X_1 represents age (1 if age > 30, 0 if otherwise); X_2, the IAC position (1 if the IAC was positioned at the lingual or inter-radicular site, 0 if otherwise); X_3, multiple roots' contact (1, if multiple roots were in contact with the perforated IAC, 0 if otherwise); X_4, the relationship between root shape and contact position of IAC (1, if it was the Adj-En type, 0 if otherwise); and X_5, the contacted length (mm) of the perforated IAC.

Based on the recalibrated model, the regression equation was shown below

$$\text{Linear predictor}_{\text{recalibration}} = -4.6430 + 0.6752 * X_1 + 1.0870 * X_2 + 0.8114 * X_3 + 1.0120 * X_4 + 0.2273 * X_5.$$

Subsequently, we constructed a user-friendly calculator using Microsoft Excel worksheets (Fig. 5 and Supplementary

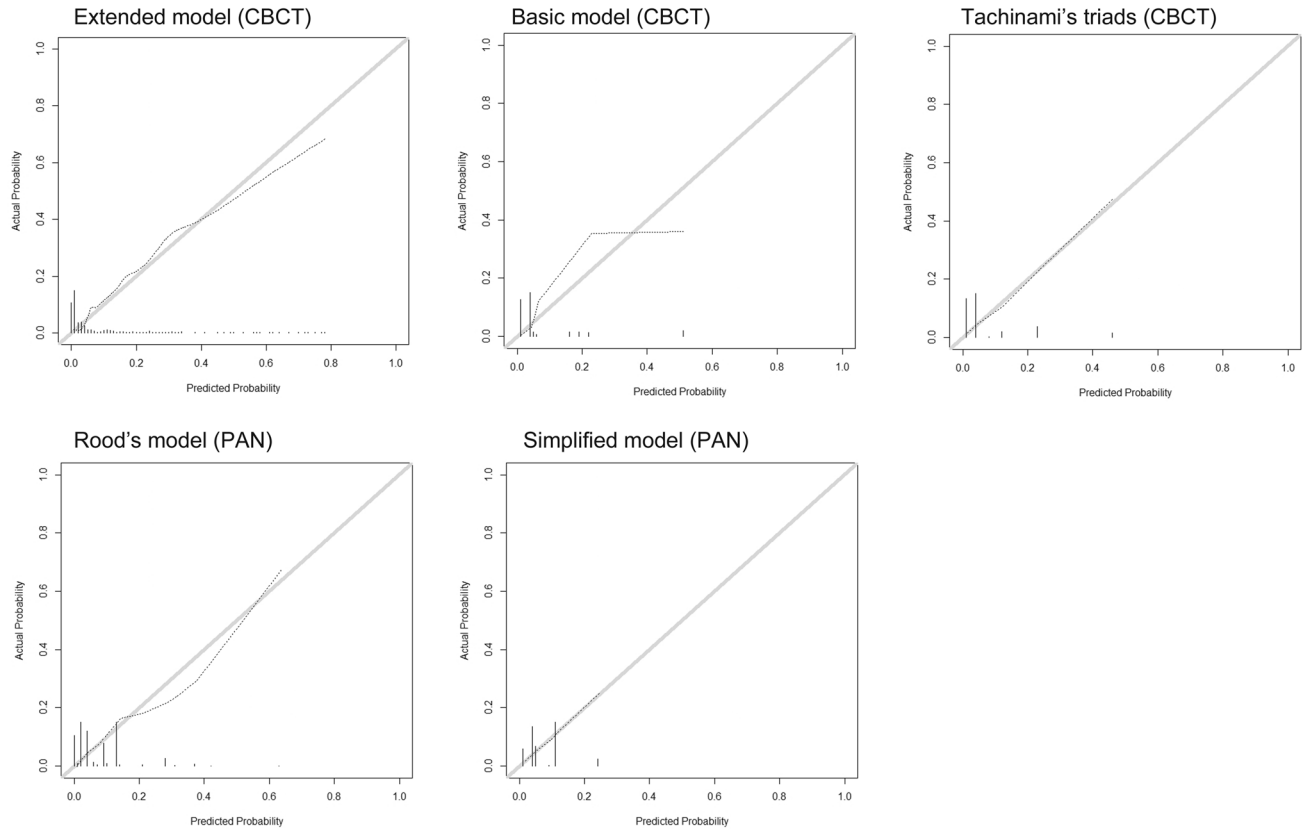


Fig. 3 Calibration plots of the models in the original sample. Calibration plots were generated for CBCT- and PAN-featured models to visualize the deviation from the model's predicted probabilities based on the observed outcome. Straight 45° lines indicate that the predicted and measured rates are similar

Excel file). The interface allowed the user to input the five factors. Then, the output displayed an adjusted IANI (p'_i) probability, calculated using the recalibrated model.

$$p'_i = p_i * (1-0.078) * 0.0244/[(1-p_i) * 0.078 * (1-0.0244) + p_i * (1-0.078) * 0.0244].$$

Note: The IANI prevalence (π) combined original and validation samples accounted for 7.8% (39 + 25/496 + 320) in this case–control study, and its overall prevalence (π_0) was 2.44% (39 + 25/1573 + 1052).

The clinical course of cases with IANI

Of 64 patients with IANI in the original and validation samples, 20, 21, and 4, respectively, recovered within one, three, and six months postoperatively (Supplementary Table). Seven had persistent IANI at six months postoperatively, which was considered permanent. The remaining 12 had IANI at the last visit within six months postoperatively and were lost to follow-up.

Discussion

Prediction models for decision tools are becoming an integral component of individualized treatment and care [25]. However, insufficient documentation that presents external validation, using model discrimination and calibration, has been reported in the field of oral surgery [29]. We further assessed the external validity of the risk-calculating model to the independent institution in this study as geographical validation [14]. CBCT-featured models showed superior performances to PAN-featured models, with the best in the extended model.

Of few studies that have reported IANI predictors on PAN-related findings using multivariate analysis, Szalma et al. [17] showed three significant predictors of seven Rood's signs. However, we identified four factors: deflection of the root and other previously highlighted factors. Furthermore, our Rood's signs model also exhibited unstable metric values with a substantial decrease in the validation sample. This instability was attributed to the difficulty of accurately recognizing Rood's signs on two-dimensional PAN images.

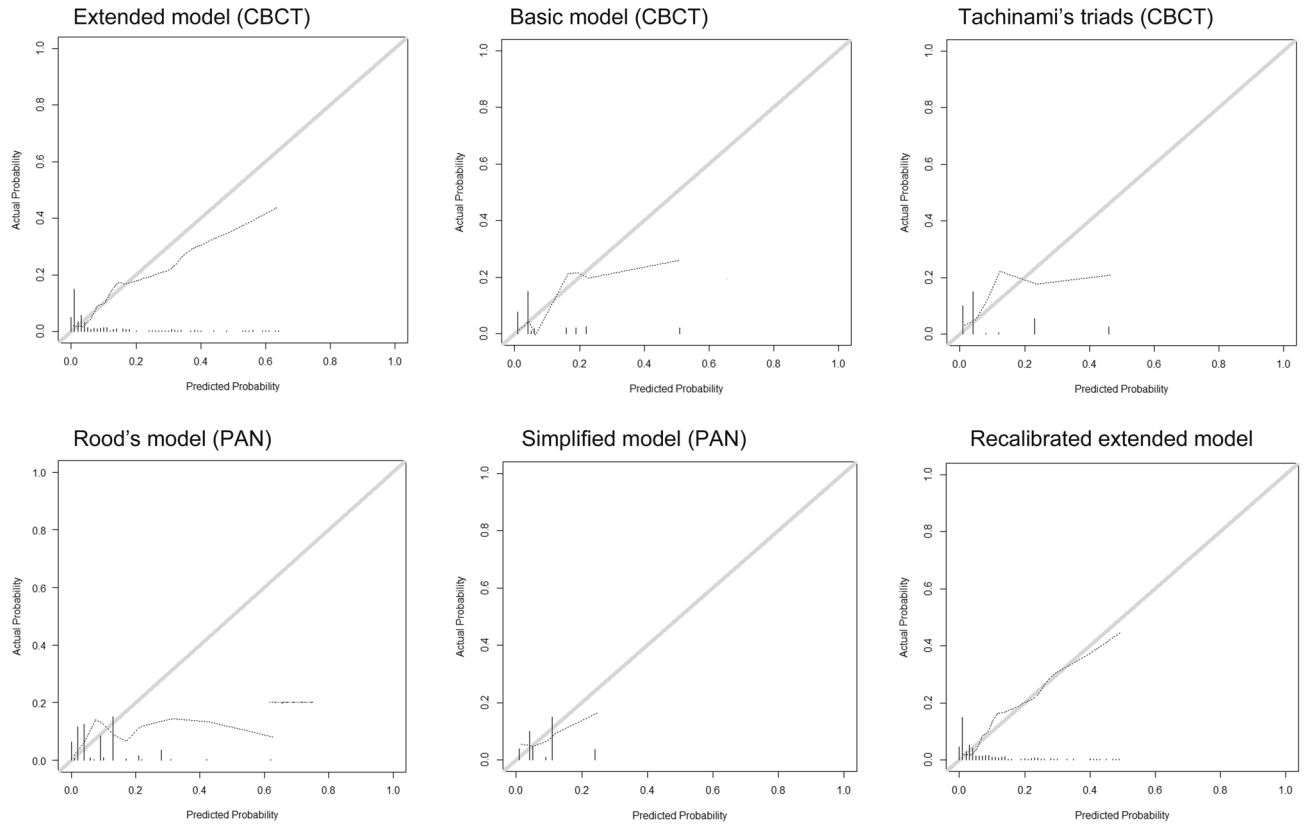


Fig. 4 Calibration plots of the models in the validation sample. Five former calibration plots showing that higher predicted risks tended to overestimate observed risks. Of those, the extended model with acceptable performances was recalibrated, resulting in an improved curve

Three major Rood's signs were integrated into an indicator included in the simplified PAN model, resulting in an improved validation performance. Nevertheless, these PAN-featured models representing inadequate performance are unsuitable as an individualized IANI-risk estimator.

The inferiority in the performance of the prediction models in the validation studies to that in original settings can be attributed to the overfitting of the model and the differences in the patient characteristics [14, 30]. However, developing a new model from the validation is an unnecessarily appropriate choice. Hence, a better alternative is to update existing prediction models according to the local circumstance or validation sample [27]. In this study, original and validation samples similarly shared approximately 1:12 case/control ratios against roughly the exact overall IANI prevalence (2.5% and 2.4%). As observed, the flattened slope of the calibration plot of the models from the validation sample (Fig. 4) implied overfitting of the model. Then, the coefficient and intercept of the extended model formula were adjusted to correct the overfitting using the logistic recalibration method [26, 27], resulting in good discrimination (C-statistic > 0.8) and ideal calibration (calibration-in-the-large, 0; and calibration slope, 1). Although the validation sample, compared with the original, included patients with a slightly higher IANI-risk status (Table 1), two datasets were constructed by sampling from each database of two institutions with similar overall IANI prevalence (2.5% and 2.4%). Therefore, the IANI-risk value calculated in the recalibrated formula was slightly shrunk compared with the original formula. Nevertheless, a validated, followed by an updated model with recalibration can carefully be applied in new patients similar to the actual and validation samples [27].

No consensus has been reached regarding CBCT being more effective than PAN in decreasing IANI risk [7, 9]. This does not, however, deny the utility of CBCT in preoperative assessment. Indeed, Korkmaz et al. [3] reported a higher feasibility of CT in determining the direction of LM3 removal and developing temporary IANI compared with PAN. Mendonca et al. [31] showed that CBCT influences intraoperative procedures, such as osteotomy or tooth sectioning, so as to not increase the chance of IANI. Therefore, oral surgeons considering using CBCT should have strong and clear clinical reservations based on the appropriate evaluation of PAN findings.

This study has some limitations, mainly associated with the retrospectively case-controlled approach, potentially leading to a selection bias. Oral surgeons between independent institutions can also bias the clinical outcomes

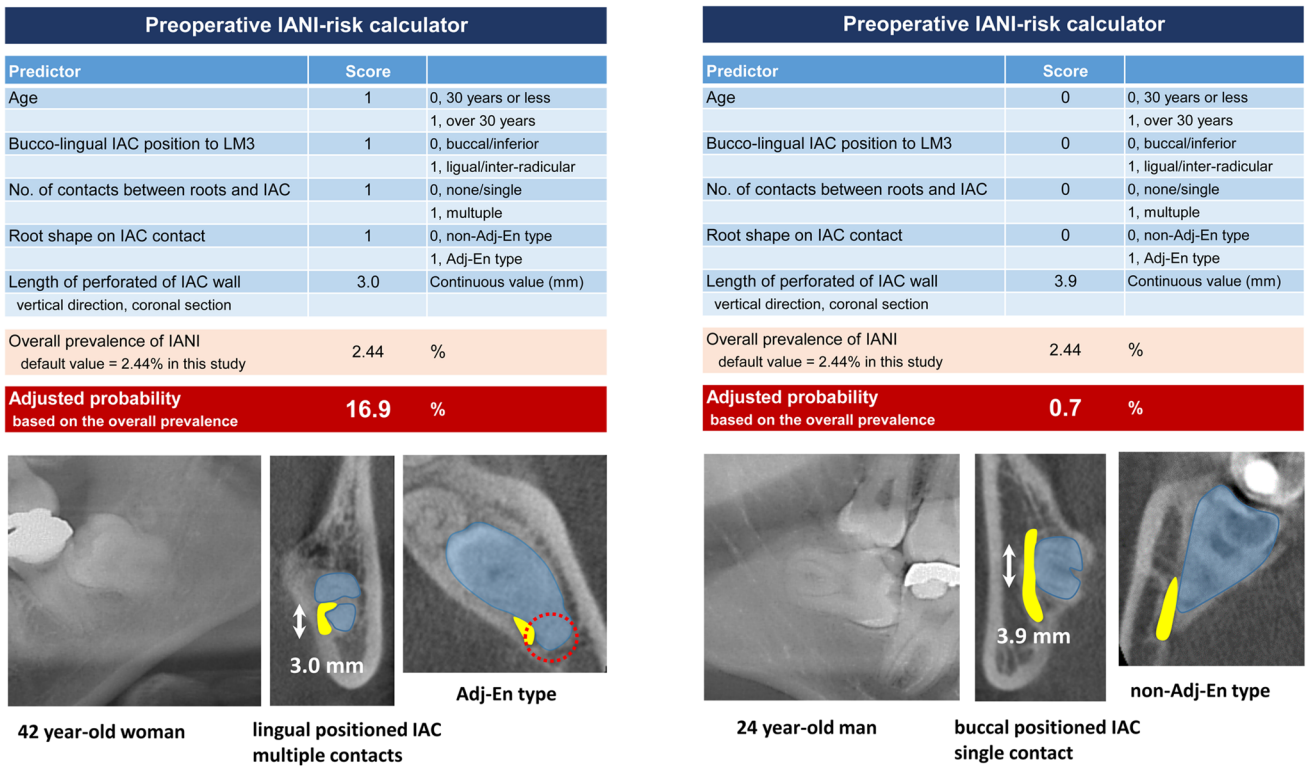


Fig. 5 A preoperative inferior alveolar injury (IANI)-risk calculator. The left panel shows an example of a 42-year-old woman. CBCT images include the lingually positioned inferior alveolar canal (IAC) in contact with multiple roots, Adj-En (adjacent to the enlarged part of the root) type of relationship between the root shape and IAC contact, then the perforated IAC wall of 3.0 mm in length. The IANI probability adjusted by an overall prevalence of 2.44% in this study accounts for 16.9%. The right panel shows an example of a 24-year-old man. CBCT images include the buccally positioned IAC in contact with the single root, the non-Adj-En type, and the perforated IAC wall of 3.9 mm in length. The adjusted probability accounts for 0.7%. Note: the relationship of Adj-En type corresponds to the IAC with a cortical defect situation located at or coronal to the enlarged part of the root (dotted red circle) on the horizontal plane as in the left case. It also corresponds to the proposition that the narrowest (i.e., the most compressed) IAC part is located at or coronal to the enlarged part of the root on the sagittal plane [12, 21]. The overall prevalence affects the calculated probability of IANI. The application of the prevalence in a user's institution is recommended if possible

after LM3 surgery. Nevertheless, the participating institution that served the validation data has been affiliated with Osaka University, and the surgeons attained common skills for managing third molar surgery during postgraduate training. External validation studies are recommended in large observational samples that accurately represent patient characteristics in clinical settings [32]. In this study, Ikeda City Hospital, among our several affiliated institutions, was the only hospital that uses a CBCT imaging system for evaluating dentoalveolar surgery. Thus, further study is warranted for a multicenter external validation study based on a standardized method of evaluating the severity of IANI, which is desirable for an increased number of patients and strengthened generalizability of the results, leading to improved efficiency [33]. Since this study focused on preoperative predictors for IANI, intraoperative factors such as surgical procedures were not investigated. Furthermore, we did not construct a web-based calculator system due to incomplete generalizability. Therefore, these issues should be evaluated by additional validation of the recalibrated model in future considerations.

In conclusion, this study is the first to externally validate the IANI-risk model for patients who will undergo LM3 removal. The extended model with increased age and four CBCT components exhibited the best performance among the PAN- and CBCT-featured models. Additionally, the risk calculator on the recalibrated model can assist clinicians when considering the impact of potential risks for developing IANI, allowing them to adjust surgical procedure strategies, which contributes to obtaining informed consent from patients according to their accurate recognition.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10266-022-00716-6.

Acknowledgements The authors thank Dr. Hiroyuki Kuragami in the Data Coordinating Center, Translational Research Center for Medical Innovation, Osaka University Hospital, for statistical consultation.

Author contributions All authors contributed to the study's conception and design. Material preparation and data collection were performed by SK, AN, and TI. Statistical analysis was conducted by TI and SK. The first draft of the manuscript was written by TI and SK. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding There is no funding.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This study was approved by the internal review boards (H30-E27-1) and was conducted in accordance with the Declaration of Helsinki.

Informed consent The requirement to obtain informed consent was waived by the internal review boards because of the retrospective design.

References

1. Lopes V, Mumenya R, Feinmann C, Harris M. Third molar surgery: an audit of the indications for surgery, post-operative complaints and patient satisfaction. Br J Oral Maxillofac Surg. 1995;33:33–5.
2. Cheung LK, Leung YY, Chow LK, et al. Incidence of neurosensory deficits and recovery after lower third molar surgery: a prospective clinical study of 4338 cases. Int J Oral Maxillofac Surg. 2010;39:320–6.
3. Korkmaz YT, Kayipmaz S, Senel FC, Atasoy KT, Gumrukcu Z. Does additional cone beam computed tomography decrease the risk of inferior alveolar nerve injury in high-risk cases undergoing third molar surgery? Does CBCT decrease the risk of IAN injury? Int J Oral Maxillofac Surg. 2017;46:628–35.
4. Leung YY, McGrath C, Cheung LK. Trigeminal neurosensory deficit and patient reported outcome measures: the effect on quality of life. PLoS One. 2013;8: e77391.
5. Kang F, Sah MK, Fei G. Determining the risk relationship associated with inferior alveolar nerve injury following removal of mandibular third molar teeth: a systematic review. J Stomatol Oral Maxillofac Surg. 2020;121:63–9.
6. Rood JP, Shehab BA. The radiological prediction of inferior alveolar nerve injury during third molar surgery. Br J Oral Maxillofac Surg. 1990;28:20–5.
7. Ghaeminia H, Gerlach NL, Hoppenreijs TJ, et al. Clinical relevance of cone beam computed tomography in mandibular third molar removal: a multicentre, randomised, controlled trial. J Craniomaxillofac Surg. 2015;43:2158–67.
8. Wang D, Lin T, Wang Y, et al. Radiographic features of anatomic relationship between impacted third molar and inferior alveolar canal on coronal CBCT images: risk factors for nerve injury after tooth extraction. Arch Med Sci. 2018;14:532–40.
9. Matzen LH, Berkhout E. Cone beam CT imaging of the mandibular third molar: a position paper prepared by the European Academy of DentoMaxilloFacial Radiology (EADMFR). Dentomaxillofac Radiol. 2019;48:20190039.
10. Matzen LH, Petersen LB, Schropp L, Wenzel A. Mandibular canal-related parameters interpreted in panoramic images and CBCT of mandibular third molars as risk factors to predict sensory disturbances of the inferior alveolar nerve. Int J Oral Maxillofac Surg. 2019;48:1094–101.
11. Kubota S, Imai T, Nakazawa M, Uzawa N. Risk stratification against inferior alveolar nerve injury after lower third molar extraction by scoring on cone-beam computed tomography image. Odontology. 2020;108:124–32.
12. Imai T, Nishimoto A, Kubota S, Nakazawa M, Uzawa N. Predictive scoring model for inferior alveolar nerve injury after lower third molar removal based on features of cone-beam computed tomography image. J Stomatol Oral Maxillofac Surg. 2022;123:136–41.
13. Diamand R, Ploussard G, Roumiguie M, et al. External validation of a multiparametric magnetic resonance imaging-based nomogram for the prediction of extracapsular extension and seminal vesicle invasion in prostate cancer patients undergoing radical prostatectomy. Eur Urol. 2021;79:180–5.
14. Moons KG, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012;98:691–8.
15. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162:W1-73.
16. Kim JW, Cha IH, Kim SJ, Kim MR. Which risk factors are associated with neurosensory deficits of inferior alveolar nerve after mandibular third molar extraction? J Oral Maxillofac Surg. 2012;70:2508–14.
17. Szalma J, Lempel E, Jeges S, Szabo G, Olasz L. The prognostic value of panoramic radiography of inferior alveolar nerve damage after mandibular third molar removal: retrospective study of 400 cases. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2010;109:294–302.
18. Hasegawa T, Ri S, Shigeta T, et al. Risk factors associated with inferior alveolar nerve injury after extraction of the mandibular third molar–a comparative study of preoperative images by panoramic radiography and computed tomography. Int J Oral Maxillofac Surg. 2013;42:843–51.
19. Ueda M, Nakamori K, Shiratori K, et al. Clinical significance of computed tomographic assessment and anatomic features of the inferior alveolar canal as risk factors for injury of the inferior alveolar nerve at third molar surgery. J Oral Maxillofac Surg. 2012;70:514–20.
20. Selvi F, Dodson TB, Nattestad A, Robertson K, Tolstunov L. Factors that are associated with injury to the inferior alveolar nerve in high-risk patients after removal of third molars. Br J Oral Maxillofac Surg. 2013;51:868–73.
21. Qi W, Lei J, Liu YN, et al. Evaluating the risk of post-extraction inferior alveolar nerve injury through the relative position of the lower third molar root and inferior alveolar canal. Int J Oral Maxillofac Surg. 2019;48:1577–83.
22. Shiratori K, Nakamori K, Ueda M, Sonoda T, Dehari H. Assessment of the shape of the inferior alveolar canal as a marker for increased risk of injury to the inferior alveolar nerve at third molar surgery: a prospective study. J Oral Maxillofac Surg. 2013;71:2012–9.
23. Blaeser BF, August MA, Donoff RB, Kaban LB, Dodson TB. Panoramic radiographic risk factors for inferior alveolar nerve injury after third molar extraction. J Oral Maxillofac Surg. 2003;61:417–21.
24. Tachinami H, Tomihara K, Fujiwara K, Nakamori K, Noguchi M. Combined preoperative measurement of three inferior alveolar canal factors using computed tomography predicts the risk of inferior alveolar nerve injury during lower third molar extraction. Int J Oral Maxillofac Surg. 2017;46:1479–83.

25. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J. 2014;35:1925–31.
26. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. BMC Med. 2019;17:230.
27. Janssen KJ, Moons KG, Kalkman CJ, Grobbee DE, Vergouwe Y. Updating methods improved the performance of a clinical prediction model in new patients. J Clin Epidemiol. 2008;61:76–86.
28. Rousson V, Zumbrunn T. Decision curve analysis revisited: overall net benefit, relationships to ROC curve analysis, and application to case-control studies. BMC Med Inform Decis Mak. 2011;11:45.
29. Russo D, Mariani P, Caponio VCA, et al. Development and validation of prognostic models for oral squamous cell carcinoma: a systematic review and appraisal of the literature. Cancers (Basel). 2021;13(22):5755.
30. Debray TP, Vergouwe Y, Koffijberg H, et al. A new framework to enhance the interpretation of external validation studies of clinical prediction models. J Clin Epidemiol. 2015;68:279–89.
31. Mendonca LM, Gaeta-Araujo H, Cruvinel PB, et al. Can diagnostic changes caused by cone beam computed tomography alter the clinical decision in impacted lower third molar treatment plan? Dentomaxillofac Radiol. 2021;50:20200412.
32. Ramspek CL, Jager KJ, Dekker FW, Zoccali C, van Diepen M. External validation of prognostic models: what, why, how, when and where? Clin Kidney J. 2021;14:49–58.
33. Sprague S, Matta JM, Bhandari M, et al. Multicenter collaboration in observational research: improving generalizability and efficiency. J Bone Joint Surg Am. 2009;91(Suppl 3):80–6.