scientific reports

Check for updates

OPEN

Nomogram for predicting secondary surgery in patients with concomitant exotropia

Haihua Liu[1,2✉], Rongjun Liu[1,2✉], Ruiying Li[1] & Kaixiu Li[1]

Concomitant exotropia (CX), a common form of strabismus, often requires surgical correction, yet up to 60% of patients undergo secondary surgery (SS) due to recurrence or residual deviation. Current risk prediction tools remain limited by inconsistent variables and short-term follow-up. This study aimed to develop and validate a nomogram integrating long-term follow-up data to predict individualized SS risk in CX patients. This is a retrospective cohort study of patients with CX who underwent surgery at the Peking University First Hospital between January 1, 2008, and December 31, 2010. Of the 355 CX cases included, 70% were randomly assigned to the training set (n = 248) and 30% to the validation set (n = 107). Demographic and clinical variables were ascertained at hospital admission and discharge and screened using multivariate Cox proportional hazards regression analysis to construct predictive models and generate a one-, three-, and six-year alignment rate nomogram. This nomogram provided an estimate of the risk of SS in patients with surgically treated CX. Internal validation was conducted using the concordance index (C-index) and calibration curve for the training and validation sets, respectively. Four independent prognostic factors were identified: age of onset, refraction, types, and deviation angles one week after surgery entered into the nomogram. The proposed nomogram showed favorable discrimination and accuracy in the training and validation sets. The C-indexes of the training and validation sets were 0.83(95%CI:0.71 ~ 0.95) and 0.80(95%CI: 0.70 ~ 0.89), respectively. The proposed nomogram may serve as a predictive tool for prognostic evaluation of CX surgery.

Keywords Exotropia, Surgery, Risk, Prediction

Concomitant exotropia (CX) is defined as an outward ocular misalignment which is the most common types of strabismus and estimated to be present in 1–2% of the pediatric population. CX could greatly influence binocular vision function and appearance of patients. Surgical correction is the main treatment of CX. However, the stability of ocular position after surgery has always been the concern of decision making for surgery and postoperative satisfaction. Many cases of recurrent or residual CX or consecutive esotropia require secondary surgery (SS) to resolve, the reoperation rate varies from 15.7 to 60%[1]. To avoid SS for CX, Previous studies have identified isolated predictors of SS, such as younger age at onset, myopia, and early postoperative overcorrection[2]. However, existing models exhibit certain limitations: (1) most rely on single-center cohorts with small sample sizes ($n < 200$), limiting statistical power; (2) variables are often analyzed in isolation rather than through multifactorial integration; (3) follow-up durations are typically < 2 years, insufficient to capture late SS events. For instance, Ekdawi et al.[3] reported a 76% recurrence rate at 10 years, yet their predictive model lacked key variables like refraction and postoperative deviation angles. Similarly, Wang et al.[4] developed a risk score but excluded patients with congenital CX, reducing generalizability. This study addresses these gaps by analyzing a large, heterogeneous cohort ($n = 355$) with 6-year follow-up, incorporating both preoperative and postoperative factors into a validated nomogram[2–6].

In this study, we sought to combine these suspected clinical factors to build models designed to accurately predict the risk of SS for CX and to help clinicians select appropriate interventions at the time of surgical decision making and early after surgery. We collected demographic and clinical variables at admission and discharge for 355 CX patients and screened them using multivariate Cox proportional hazard regression. We constructed nomograms of four predictors: age of onset, refraction, type, and deviation angles one week after surgery. This nomogram provides an estimate of the risk of SS in surgically treated CX patients. At the same time, the model was validated using multiple indicators, including the concordance index (C-index), the area under the receiver operating curve, and the calibration curve.

[1]Department of Ophthalmology Center, Peking University First Hospital, No.8 Xishiku Road, Xicheng District, Beijing 100034, China. [2]Haihua Liu and Rongjun Liu: contributed equally to this work. ✉email: haihua.liu@pku.edu.cn; drliurongjun@163.com

Materials and methods

Study population

We collected the inpatient and outpatient medical records of 437 patients with CX who underwent surgery at Peking University First Hospital between January 1, 2008, and December 31, 2010. Clinical data were extracted and cross-checked by two experienced doctors who determine whether patients were included or excluded. This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of Peking University First Hospital.Written informed consent was waived due to the use of retrospective data by the Ethics Committee of Peking University First Hospital.

Patients who were diagnosised as CX and underwent uneventful CX surgery were included in this study. The exclusion criteria were as follows: central nervous system defects, paralytic exotropia, anatomical abnormalities of the eye, previous ophthalmic surgery, and a follow-up period < six months.

Among the 437 surgically treated inpatients with CX, 22 (5.1%) were not follow-up in outpatient and 60 (13.7%) were followed up for less than six months. Therefore, 355 (81.2%) patients that were surgically treated for CX were included in this study. The cohort of this study was randomly divided using a computer according to the proportion of 70% and 30%; 248 cases were included in the training set, and 107 cases were included in the validation set (Fig. 1).

Data collection and follow-up

Patients included in this study had a complete medical history and were regularly followed up. A detailed ophthalmological examination was performed before surgery to exclude other ocular diseases, and further strabismus specialist examinations included visual acuity, cycloplegic refraction, deviation angles of distance and near primary, upward and downward 25-degree gaze, extraocular muscle function, and binocular sensory function.The patients visited the outpatient clinic to establish a follow-up medical record. Follow-ups were scheduled one day (1 d), one week (1w), one month (1 m), three months (3 m), and six months (6 m) after surgery. Patients with unstable visual acuity and ocular alignment were followed up three to four times in 12 m, and patients with stable visual acuity and ocular alignment were followed up one to two times in 12 m. During the follow-up period, when the deviation angles of both distance and near were > + 15PD (prism diopters)or < −15PD, we concluded that the patients needed SS. This was a positive endpoint in this study. The follow-up termination time for this study was December 31, 2016. For patients who required SS, the follow-up termination time was the date of SS confirmation. For cases that did not require SS and the last follow-up date was before December 31, 2016, censored data were confirmed by the date of the last follow-up. The cohort was followed up for 6 to 72 months. The six-year follow-up endpoint was chosen based on two considerations: (1) the data set provided complete follow-up coverage from 2008 to 2016, ensuring sufficient time to capture late SS; (2) This time frame aligns with prior studies investigating long-term outcomes in strabismus surgery[3,4].

Potential predictive variables included the patient demographics, medical history, and clinical variables at admission and after discharge. The demographic variables included sex, age at diagnosis, age of onset, and age at surgery. The medical history included family history, history of premature birth, history of febrile convulsion, duration of strabismus. Clinical variables included best-corrected visual acuity, refraction, amblyopia, deviation angles, stereopsis, type of CX, dominant eye or non-dominant eye surgery, surgical muscle, surgical method, surgical amount, accompanying signs (A-V sign, dissociative vertical deviation (DVD), inferior oblique muscle overaction (IOOA), and nystagmus). The distance and near deviation angles and distance and near stereopsis were obtained at one day (1 d), one week (1w), one month (1 m), three months (3 m), and six months (6 m) and the last postoperative follow-up.

Surgical method

Considering to the different deviation angles and types of CX, the surgical method was designed personalized. Mainly three methods were used: unilateral lateral rectus (LR) recession, bilateral LR recession, and bilateral LR recession combined with unilateral medial rectus (MR) resection. The range of surgical amount was as follows: LR recession, 5–10 mm; MR resection, 3–6 mm. The surgical amount calculated in this study was the sum of muscle recession and resection. For basic CX and divergence excess CX choose LR recession, for convergence insufficiency CX choose LR recession combined with MR resection.
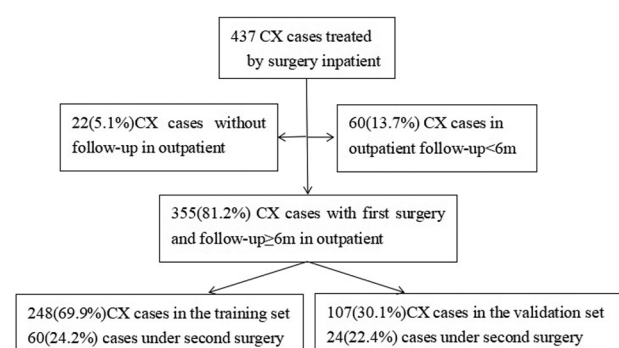
Fig. 1. Flow chart of the grouping of patients with concomitant exotropia(CX).

Missing data

In the study cohort, no diagnostic and surgical data were omitted at admission. Children who were unable to understand and cooperate in the vision test and stereo vision test were judged to have a clear tendency to use one eye as amblyopia, and stereo vision was recorded as none.A differentiated approach addressed missing data: for variables with low missing rates (e.g., near deviation angles at one day post-operation: 5.4%), median imputation was applied to mitigate outlier effects; whereas stereopsis data, exhibiting non-random missingness (46.3% loss during follow-up due to factors like age-related non-compliance) and deemed non-critical to primary outcomes, were excluded to avoid bias amplification. While this exclusion may omit a potential predictor, the conservative strategy aligns with guidelines for managing substantial non-random missingness in secondary variables, thereby safeguarding analytical validity.Table 1.

Statistical analysis

To construct and validate nomograms, patients were randomly divided into a training set and a validation set. 70% is included in the training set and 30% is included in the validation set. Chi-square test or independent sample t test were used to describe and compare the features of the two groups. Results are expressed as percentage and mean ± SD difference. Using Akaike information criteria in Cox proportional risk model, nomogram variables are selected by stepwise selection method. Independent prognostic outcomes for screened SS were expressed as hazard ratio (HR) and 95% confidence interval (95% CI). Nomograms were constructed to predict 1-year, 3-year, and 6-year alignment rate based on predictive models and identified prognostic factors. The outcome of interest was the time from initial surgery to secondary surgery (SS). Patients who did not undergo SS by the end of the study (December 31, 2016) were censored at their last follow-up date. The Cox proportional hazards model was chosen to account for both the occurrence and timing of SS, as it allows for the analysis of time-to-event data while appropriately handling censored observations.

Variable selection was guided by prior literature, clinical expertise, and data quality[2–5]. Predictors with established relevance to SS risk (e.g., age of onset, refraction) were prioritized. Univariate and multivariate Cox regression, combined with stepwise AIC optimization, ensured a balance between statistical significance and clinical utility.

Training sets and validation sets were used for internal validation. Discrimination and calibration consisted of the nomogram validation. The consistency index (C-index) was calculated to evaluate the discriminant ability of the prediction model. For the clinical application of this model, the total score per patient is calculated by nomogram. Calibrations are evaluated using calibration charts and a 1000 sample bootstrap is used to reduce over fitting. The H-L test for the nomogram's predicted probabilities at 1, 3, and 6 years using the hoslem. test function. The test divides the data into 10 deciles of predicted risk and compares observed vs. expected event rates within each group.SPSS (version 25.0) and R (version 3.3.28) were used for statistical analysis.The significance level was set at 0.05, and all tests were bilateral.

Results

Cohort characteristics

A total of 355 consecutive patients with CX were identified. 248 (69.86%) cases were included in the training set.The mean follow-up time was 28.34 ± 25.45 months in the training set, with 24.19% undergoing SS. Patients without SS were censored at their last follow-up, ensuring alignment with time-to-event analysis principles. The 1-year, 3-year and 6-year alignment rates were 90.32%, 81.85% and 75.58%, respectively. There were 107 cases (30.14%) in the validation set.The mean follow-up time was 29.23 ± 24.73 months, with 22.42% undergoing SS. The 1-year, 3-year and 6-year alignment rates were 94.39%, 88.78 and 77.57%, respectively.The characteristics of the two groups of CX cases are shown in Table 2, and there is no significant difference between the training set and the validation set.

Independent prognostic factors in the training set

The results of the multivariate cox regression analysis are listed in Table 3. Age of onset ($P = 0.004$), refraction ($P < 0.001$), type of CX ($P = 0.023$), and deviation angles at near 1w postoperative ($P = 0.026$) were associated with poor postoperative CX prognosis.

Prognostic nomogram for OS

The prognostic nomogram that integrated all variables selected using multivariate cox regression is shown in Fig. 2. The nomogram illustrated that the age of onset, refraction, type of CX, and deviation angles of strabismus at near 1w after surgery. Each subtype within these variables was assigned a score on a point scale. By adding the total score and locating it on the total point scale, we can easily draw a straight line to determine the one-, three-, and six-yearalignment probabilities.

Variable	Missing data proportion	Handling method
Near deviation angles (1 d)	5.4%	Median imputation
Stereopsis (follow-up)	46.3%	Excluded from analysis
Other clinical variables	< 1%	Retained (no imputation needed)

Table 1. Missing data proportions and handling methods.

Variables	Total ($n = 355$)	Validation set ($n = 107$)	Training set ($n = 248$)	Statistic	P-value
Sex, n(%)				$\chi^2=0.40$	0.525
Male	175 (49.30)	50 (46.73)	125 (50.40)		
Female	180 (50.70)	57 (53.27)	123 (49.60)		
Age at Surgery (years)	6.87 ± 3.39	6.66 ± 2.96	6.96 ± 3.56	t=−0.76	0.445
Age of Onset (years)	3.08 ± 2.27	2.86 ± 2.24	3.17 ± 2.27	t=−1.20	0.233
Refraction, n(%)				$\chi^2=0.06$	0.970
≥0.0D	264 (74.37)	79 (73.83)	185 (74.60)		
−3.0D-0.0D	77 (21.69)	24 (22.43)	53 (21.37)		
<−3.0D	14 (3.94)	4 (3.74)	10 (4.03)		
Amblyopia, n(%)				$\chi^2=0.03$	0.874
No	307 (86.48)	93 (86.92)	214 (86.29)		
Yes	48 (13.52)	14 (13.08)	34 (13.71)		
Etiology, n(%)				$\chi^2=5.01$	0.073
Congenital	79 (22.25)	32 (29.91)	47 (18.95)		
Acquired	276 (77.75)	75 (70.09)	201 (81.05)		
Accompanying Sign				$\chi^2=0.01$	0.928
No	231 (65.07)	70 (65.42)	161 (64.92)		
Yes	124 (34.93)	37 (34.58)	87 (35.08)		
Deviation Angles (PD)	−35.10 ± 17.67	−35.92 ± 17.43	−34.75 ± 17.79	t=−0.57	0.568
Type, n(%)				$\chi^2=2.27$	0.322
Basic	193 (54.37)	64 (59.81)	129 (52.02)		
Convergence insufficiency	138 (38.87)	38 (35.51)	100 (40.32)		
Divergence excess	24 (6.76)	5 (4.67)	19 (7.66)		
Surgery Amount (mm)	12.42 ± 3.70	11.93 ± 3.90	12.63 ± 3.59	t=−1.64	0.102
Deviation Angles 1w				-	0.799
−5 ~ 5PD	309 (87.04)	91 (85.05)	218 (87.90)		
−6~−10PD	32 (9.01)	12 (11.21)	20 (8.06)		
>−10PD	8 (2.25)	3 (2.80)	5 (2.02)		
6 ~ 10PD	1 (0.28)	0 (0.00)	1 (0.40)		
> 10PD	5 (1.41)	1 (0.93)	4 (1.61)		

Table 2. Characteristics of the training and validation sets.

|	B	HR	HR 95%CI	Z	P-value
Age of Onset	0.81	2.25	1.30 ~ 3.88	2.91	0.004
Refraction	1.113	0.286	0.171–0.377	4.744	< 0.001
Type of CX	0.016	1.016	0.900–1.744	2.303	0.023
Deviation Angles 1w	0.834	2.28	1.09 ~ 4.77	2.21	0.026

Table 3. Multivariate Cox regression of the training set. Deviation angles 1w = deviation angles of strabismus at near one week postoperative.

Validation of the nomogram

The ROC curves of the training and validation sets are shown in Fig. 3. In the training set, the C-index was 0.83(95%CI:0.71 ~ 0.95), and in the validation set, the C-index was 0.80(95%CI: 0.70 ~ 0.89). The nomogram's predictive accuracy was validated across 1-, 3-, and 6-year time points. ROC curves and calibration plots for all intervals are provided in Figs. 3, 4 and 5. While discrimination slightly declined over time (AUC: 0.83 at 1 year vs. 0.76 at 6 years in the training set), the model retained clinical utility for long-term risk stratification, with sensitivity and specificity exceeding 70% at all time points (Table 4).

The calibration curves for predicting the one-year, three-year, six-year alignment probability of patients with CX in the training and validation sets are shown in Fig. 6, respectively. Both curves showed good agreement between the nomogram predictions of the one-year OS and the actual observations.

Quantitative calibration was assessed using the Hosmer-Lemeshow test, which demonstrated no significant deviation between predicted and observed secondary surgery rates at 1, 3, and 6 years (all $P > 0.05$; Table 5). These results align with the visual calibration curves (Figs. 3, 4, 5 and 6), confirming the nomogram's reliability in stratifying short- and long-term SS risks.
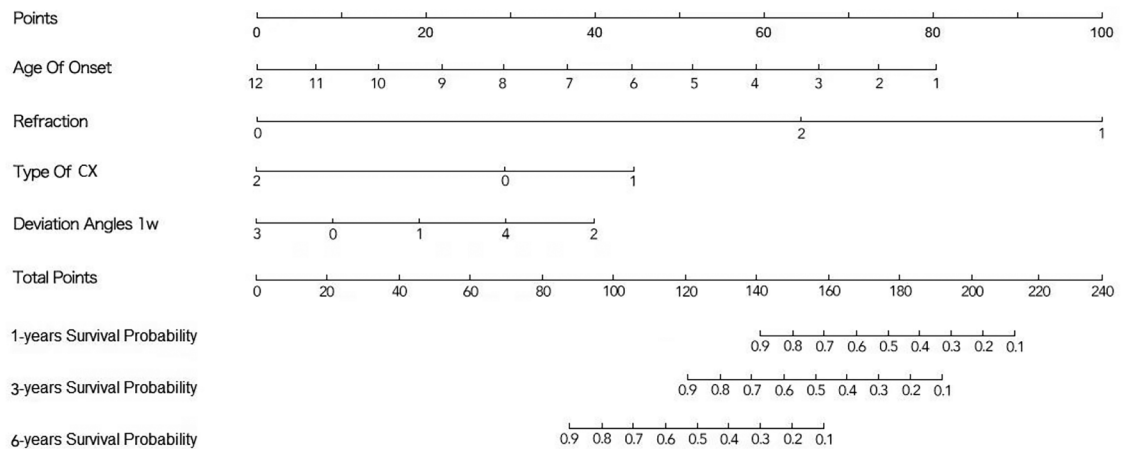
Fig. 2. Nomogram for predicting the one-, three- and six-year alignment probability of patients with CX (To use the nomogram, an individual patient's value is located on each factor axis, and a line is drawn upward to determine the number of points received for each factor value. The sum of these numbers is located on the total points axis, and a line is drawn downward to the alignment a CX to determine one-, three- and six-year alignment probability).
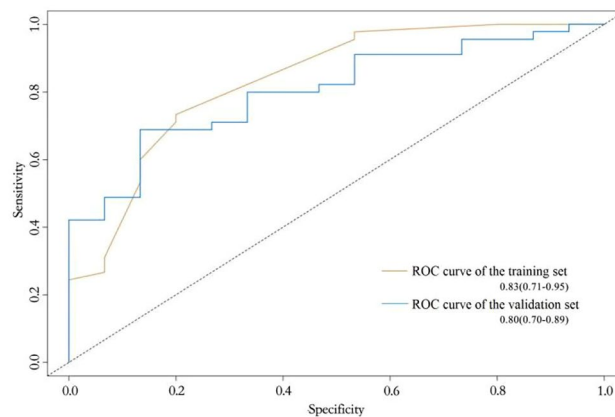
Fig. 3. ROC curve for predicting the one-year alignment probability of patients with CX.
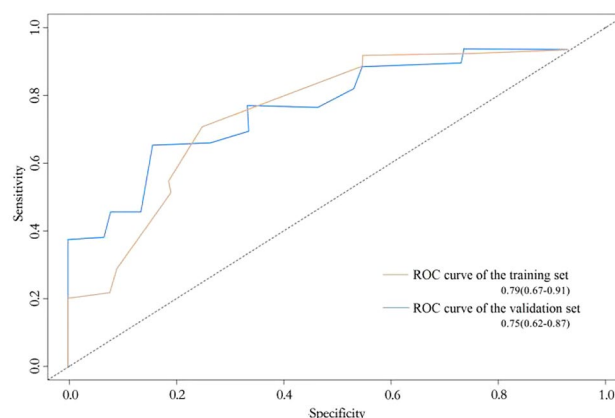
Fig. 4. ROC curve for predicting the three-year alignment probability of patients with CX.
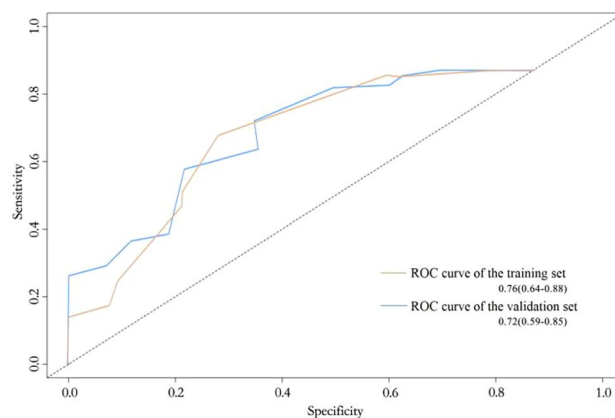
Fig. 5. ROC curve for predicting the six-year alignment probability of patients with CX.

	AUC of training set	AUC of validation set	Sensitivity, %	Specificity, %
1-year	0.83(0.71–0.95)	0.80(0.70–0.89)	86.22	85.74
3-year	0.79(0.67–0.91)	0.75(0.62–0.87)	80.52	82.11
6-year	0.76(0.64–0.88)	0.72(0.59–0.85)	75.84	78.33

Table 4. Accuracy of the nomogram in predicting the alignment probability.

Decision curve analysis demonstrated that the nomogram provides significant net clinical benefit compared to default strategies across threshold probabilities of 5–45% (Fig. 7). At a 25% risk threshold, the model achieves a net benefit of 0.42, meaning it correctly identifies 42% of patients requiring SS while minimizing unnecessary interventions. This underscores its utility in optimizing surgical decision-making and resource allocation.

Discussion

The high recurrence rate has been a troublesome problem for CX surgery. In this study, 84(23.66%) cases underwent SS. Wang[4] reported the SS was 11.90% for 12 months and 27.14% for 24 months. Ekdawi[3] reported the recurrence rate reach 76% by 10 years and 86% by 15 years. There are many clinical reports on the outcome and risk factors of CX surgery, and the results of long-term alignment rates and related risk factors vary widely. Strabismus specialists should make effort to reduce the need for SS due to recurrent or residual CX or consecutive esotropia after CX surgery. It is meaningful to establish a model for predicting the risk of SS for CX so that we could identify high-risk patients early and adjust the treatment plan appropriately at the early stage to reduce the occurrence of SS for CX.

In this study, we aimed to develop and validate a nomogram for predicting secondary surgery (SS) risk in patients with comitant exotropia (CX), incorporating four independent predictors.The proposed nomogram serves as a practical clinical tool. When applying this model, clinicians should first evaluate the four predictors included: age at onset, refractive status, exotropia type, and near deviation angle at 1 week postoperatively. Each predictor corresponds to a specific score on the nomogram's scale (see Fig. 2). By summing these scores, the total points directly translate to the probability of postoperative alignment maintenance at 1, 3, and 6 years, providing quantifiable guidance for individualized risk stratification.

For example: High-risk patients (e.g., total score > 180): These individuals have a > 60% probability of requiring SS within 6 years. For such cases, surgeons might consider more aggressive initial surgical dosing (e.g., larger lateral rectus recession) to minimize residual deviation, schedule closer postoperative follow-ups (e.g., monthly for the first 6 months).Early counseling with patients and families about the likelihood of SS is also critical to manage expectations. Intermediate-risk patients (e.g., total score 120–180): These patients may benefit from standardized surgical protocols but require vigilant monitoring. Clinicians should prioritize adherence to follow-up schedules (e.g., every 3 months for the first year) and emphasize compliance with postoperative visual training or refractive correction (e.g., myopia control in patients with high refractive errors). Low-risk patients (e.g., total score < 120): While these individuals have a lower probability of SS (< 20% at 6 years), routine follow-ups remain necessary to detect late-onset deviations.

Many studies have explored the factors related to recurrence. One cohort study found that a younger age at onset increases the risk of recurrence[5]. Patients with early onset (≤ 2 years) had a significantly higher recurrence rate. Another study indicated that older age of onset was associated with a lower recurrence rate. Additionally, some research has reported that early onset is linked to poorer sensory outcomes[7]. However, several studies demonstrated no significant differences in surgical responses between early-onset and late-onset CX patients[8,9]. In this study, age of onset was identified as a predictor of SS. The older the age of onset, the lower the score; conversely, the younger the age of onset, the higher the score and the greater the risk of SS.

6
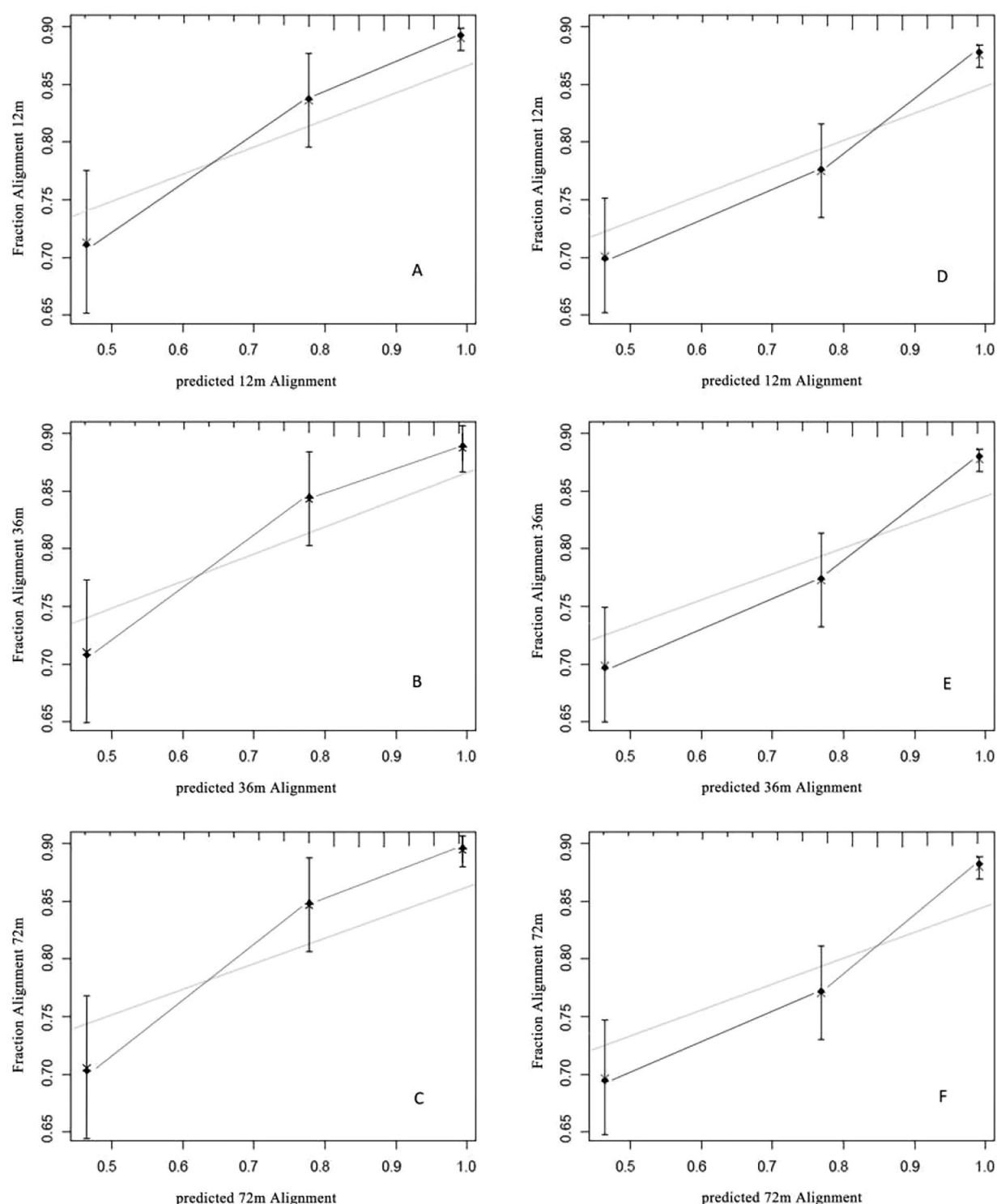
Fig. 6. Calibration curve for predicting the one-year, three-year, six-year alignment probability of patients with CX in the training set(A-C) and validation set(D-F).

In the past, due to a lack of understanding of CX, parents often struggled to provide accurate onset ages for their children, typically reporting ages that were older than the actual age. Recently, with the increased availability of health education, parents are more likely to seek medical advice promptly upon noticing abnormal symptoms in their children, which aids in the early detection of strabismus. As a result, the onset age reported by parents has become more accurate. Congenital CX in infants can be diagnosed at birth, but achieving good

Time point	Cohort	Chi-square statistic	P-value
1-year	Training Set	8.21	0.412
	Validation Set	7.85	0.448
3-year	Training Set	9.34	0.314
	Validation Set	8.97	0.345
6-year	Training Set	10.12	0.258
	Validation Set	11.05	0.198

Table 5. The H-L test results for the training and validation sets.
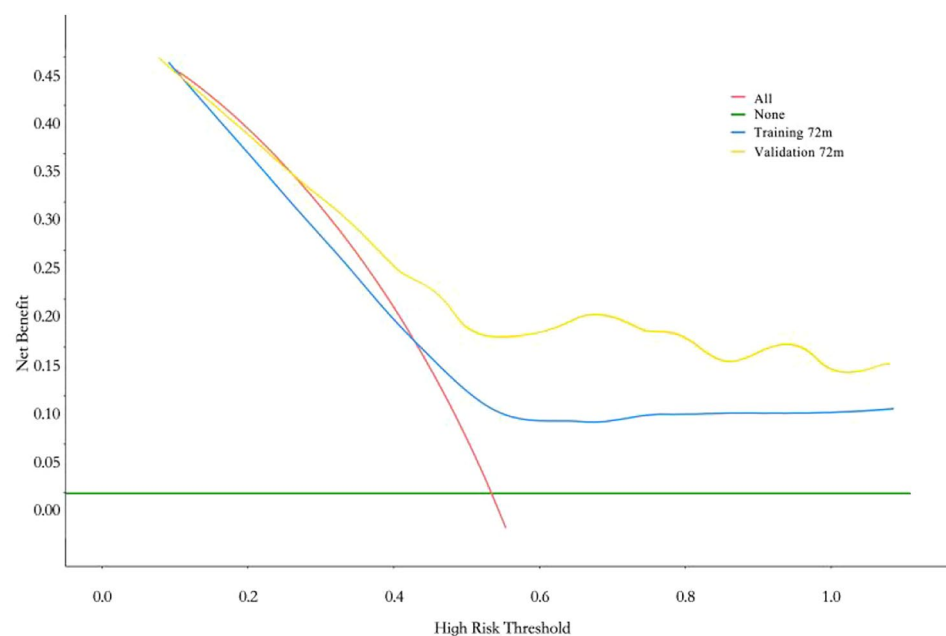
Figs. 7. Decision curve analysis (DCA) for assessing the net benefit of applying Training set 72 m and Validation set 72 m predictive models.

binocular visual function remains difficult, even with optimal postoperative alignment. Nevertheless, the long-term rate of SS is high and warrants special attention.

In this study, refractive error was found to influence the risk of SS, while hyperopia had a lesser impact on SS. Notably, myopia ranging from −3.0D to 0.0D was associated with the highest SS. Some research indicates that refractive error or anisometropia does not significantly affect final surgical outcomes[10]. However, other studies suggest that myopia may lead to less favorable responses to surgery, and anisometropia has also been linked to poorer outcomes[11]. This digital age, the volume of visual information individuals encounter daily has significantly increased, necessitating constant adjustments by the eyes to adapt to varying visual stimuli. This prolonged overload in visual processing has contributed to a notable rise in the incidence of visual fatigue and myopia. The situation has been exacerbated by the COVID-19 pandemic, during which excessive screen time and reduced outdoor activities have led to a marked increase in cases of myopia, particularly among individuals with small-angle exotropia[12]. Intermittent exotropia and myopia are often regarded as co-occurring conditions[13]. In patients with intermittent exotropia, there is competitive inhibition between the two eyes, with the non-dominant eye relying on fusion images to adjust its position. This indicates a scenario where the adjustment response is less robust than the adjustment stimulus, resulting in an adjustment lag. Myopic patients also experience similar lag when focusing on near objects or through negative spherical lenses.

The interplay between accommodative and convergence functions is a physiological mechanism underlying the concurrent progression of these two conditions, and the incidence of myopia among patients with exotropia is on the rise[14]. Additionally, a meta-analysis by Tang[15] identified myopia as a significant risk factor for exotropia and less favorable responses to surgery. In cases of significant myopia (≥−5D), it is advisable to adjust surgical dosages accordingly[16]. Due to the prism effect of negative lenses, the degree of strabismus measured while wearing glasses tends to be greater than when assessed with the naked eye, which can lead to overcorrection during surgery.

Audrey et al.[17] studied the distribution of CX subtypes and found that the proportion of strabismus with divergence excess was the highest 55%, and that of convergence insufficiency was the lowest. The research results of Chen[18] showed that the basic exotropia accounted for the highest proportion of 74.7%, and the divergence

excess exotropia accounted for the lowest proportion. Different from the above-mentioned studies, the results of this study show that the convergence insufficiency type has been the main type of CX. The increased proportion of convergence insufficiency CX may be related to the increased incidence of myopia. The accommodative collective reflex is weakened in myopic patients, especially in moderate and low myopia[19]. Compared with the non strabismus population, CX patients need to mobilize more sets and adjustments to achieve binocular vision. CX patients with myopia not only need to mobilize more sets to produce more regulation, but the existence of regulation lag will cause the imbalance between set and regulation, causing convergence insufficiency and eye abduction, and then the degree of near temporal strabismus increases, which is larger than that of far strabismus, resulting in the increase in the proportion of convergence insufficient CX.

The results of this study showed that the near-eye position 1w postoperatively could be used to predict the long-term eye position. The results of the 1w eye position were divided into five grades, with the highest PD score of deviation angles >−10PD, suggesting that residual exotropia should be avoided as much as possible after CX surgery once there was a high probability that CX will develop. Ekdawi reported that early postoperative overcorrection will decrease with time until it disappears, while early postoperative residual exotropia is very easy to cause the recurrence of exotropia[14]. Therefore, early postoperative overcorrection is widely accepted, and it is emphasized that the surgical design of exotropia should be overcorrected. However, there were also studies that believe that overcorrection in CX surgical design can not effectively predict the trend of postoperative drift back. A small number of exotropia can recover to the normal position in the early postoperative period. However, most of the overcorrection does not disappear and esotropia occurs, which is easy to cause monocular suppression, amblyopia or loss of binocular visual function in children. Therefore, it is proposed that CX should not be overcorrected[20]. Studies have reported that the distal esotropia angle ≥20pd on the first day after surgery is a risk factor for esotropia after exotropia surgery, and 12% of such patients have esotropia[21]. In this study, the proportion of mild overcorrection in the early postoperative period was high, and the strabismus angle was 3.51pd at 1 W after surgery. This study supports that the CX surgical design should be appropriately overcorrected.

This study we use Cox model instead logistic regression. Logistic regression could predict the binary outcome of SS occurrence, it ignores the temporal dimension of the data. In contrast, the Cox model leverages both the occurrence and timing of SS, providing dynamic risk estimates (e.g., 1-, 3-, and 6-year probabilities) critical for clinical decision-making. Furthermore, censoring is inherent in longitudinal studies, as some patients may not experience SS within the observation window. The Cox model inherently accommodates this through partial likelihood estimation, making it more robust for our cohort with variable follow-up durations (6–72 months).

This study aimed to develop a nomogram for predicting the risk of SS following CX surgery, and it has several notable strengths. First, the scores of most predictors aligned with findings from existing literature, suggesting that the scores are both valid and representative. Second, we conducted an internal validation of the nomogram, and its performance in terms of discrimination and precision was superior, indicating that it is less dependent on specific sample characteristics. Furthermore, the model's sensitivity and specificity exceeded 80%, demonstrating its potential clinical utility. While this study analyzed surgeries performed between 2008 and 2010, the surgical techniques (e.g., lateral rectus recession, medial rectus resection) and postoperative management protocols remain standard today. Recent studies continue to validate identical surgical dosages and risk factors, confirming the enduring clinical relevance of our findings. Furthermore, long-term follow-up (up to 6 years) is critical for capturing late SS events, as short-term data often underestimate recurrence rates.Future studies may refine this nomogram with emerging technologies (e.g., AI-assisted alignment measurements), but the core predictors identified here are foundational to SS risk stratification.

Compared to prior models focused on isolated predictors or short-term outcomes, our nomogram uniquely integrates preoperative and postoperative factors to provide a holistic risk assessment. While Ekdawi[3]and Wang[4] advanced our understanding of surgical alignment and preoperative variables, their exclusion of critical factors like refraction limited clinical utility. Our model's superior discrimination (C-index >0.80 vs. 0.68–0.75 in earlier studies) and validation across 6 years address these gaps, offering a tool that balances statistical rigor with practical applicability. Furthermore, decision curve analysis confirms that the nomogram optimizes net clinical benefit, outperforming default 'treat all' or 'treat none' strategies across plausible risk thresholds.

This study has several limitations. First, all data were derived from a single tertiary hospital. Although we allocated two-thirds of patients to nomogram development and one-third for internal validation-a methodologically sound approach—the absence of an external cohort necessitates caution. While the nomogram showed robust performance in discrimination (C-indexes: 0.83 training, 0.80 validation) and calibration, its generalizability requires verification across diverse populations and healthcare settings. The geographic and institutional homogeneity of our cohort may restrict broader applicability. Future validation should prioritize multicenter studies involving varied demographic profiles, surgical techniques, and follow-up protocols, particularly across distinct healthcare systems. Such external validation is critical to confirm the model's predictive accuracy and strengthen its utility in clinical decision-making.

Second, The cohort's follow-up duration ranged from 6 to 72 months, which reflects real-world variability in clinical follow-up. To mitigate bias from censored data, we employed Cox proportional hazards regression, which accounts for varying observation periods by modeling time-to-event outcomes. Sensitivity analyses confirmed that censored cases (e.g., lost to follow-up) did not differ significantly from non-censored cases in baseline characteristics ($p > 0.05$ for all variables). However, the possibility of informative censoring (e.g., patients with poor outcomes being lost to follow-up) cannot be entirely ruled out. Future prospective studies with standardized follow-up schedules are needed to validate these findings.

Third, while median imputation preserved data integrity for minor missingness, it assumes missingness is random, which may not fully hold for stereopsis. Variable follow-up intervals, though addressed statistically, could still influence long-term risk estimates. Finally, the exclusion of stereopsis due to high missingness may

overlook its potential role in surgical outcomes. While our nomogram leverages historical data, its predictive framework is designed to accommodate future refinements. As novel surgical techniques or diagnostic tools emerge, the model can be recalibrated through prospective studies. We strongly advocate for multicenter collaborations to validate and update this tool, ensuring its adaptability to evolving clinical standards.

Future research should prioritize the following directions: Multicenter External Validation: Collaborate with institutions across diverse geographic and healthcare settings to validate the nomogram's performance in contemporary cohorts.Prospective Long-Term Studies: Establish prospective registries to track CX patients for ≥ 10 years, capturing late SS events and late-onset comorbidities (e.g., amblyopia, binocular vision loss). Cost-Effectiveness Analysis: Evaluate the nomogram's impact on healthcare resource allocation by comparing SS rates, patient outcomes, and costs between risk-stratified and standard care cohorts.

In conclusion, we have developed and validated a user-friendly nomogram for predicting the risk of SS following CX surgery. The modest reduction in AUC for 6-year predictions reflects inherent challenges in modeling late SS events. Nevertheless, the nomogram's sustained accuracy (AUC > 0.70) surpasses existing tools focused solely on short-term outcomes, offering clinicians a unique advantage in counseling patients about lifelong risks. This tool has the potential to assist in surgical planning and help strabismus specialists better communicate surgical risks to patients and their families. However, further replication and independent validation in a larger cohort are necessary before it can be routinely applied in clinical practice.

Data availability
Data from this study are available upon reasonable request to the corresponding author.

References
1. Benson, M. D., Wozniak, J. & Macdonald, I. M. Ananlysis of strabismus reoperations in Northern Alberta, Candada from 1995 to 2015[J]. Can J Ophthalmol, **54**(1):94–97. (2019).
2. Knn, W. J. & Kun, M. M. The fast exodrift after the first surgical treatment of Exotropia and its correlation with surgical outcome of second surgery[J].BMC Ophthalmol,**18**(1):67–72. (2018).
3. Ekdawi, N. S. et al. Postoperative outcomes in children with intermittent Exotropia from a population-based cohort[J]. J AAPOS,**13**(1):4–7. (2009).
4. Zijin Wang, T. et al. Preoperative and postoperative clinical factors in predicting the early recurrence risk of intermittent Exotropia after surgery[J]. Am J Ophthalmol, **251**: 115–125. (2023).
5. Kim, D. H., Yang, H. K. & Hwang, J. M. Long term surgical outcomes of unilateral recession-resection versus bilateral lateral rectus recession in basic-type intermittent Exotropia in children[J]. Sci Rep,**11**(1):19383. (2021).
6. Repka, M. X. et al. The relationship of age and other baseline factors to outcome of initial surgery for intermittent exotropia[J]. Am J Ophthalmol, **212**,153–161. (2020).
7. Choi, Y. M. & Kim, S. H. Comparison of clinical features between two different types of Exotropia before 12 months of age based on stereopsis outcome[J]. Ophthalmology, **120**(1):3–7. (2013).
8. Stoller, S. H., Simon, J. W. & Liniger, L. L. Bilateral latetal rectus recession for exotropia: a survival analysis[J]. J Pediatr Ophthalmol Strabismus, **31**(2):836–840. (1994).
9. Suh, S. Y. et al. Outcomes of surgery in children with early-onset exotropia[J]. Eye(Lord), **27**(7):836–840. (2013).
10. Awadein, A., Eltanamly, R. M. & Elshazly, M. Intermittent exotropia: relation between age and surgical outcome: a change-point analysis[J]. Eye **28**:587–593. (2014).
11. Pineles, S. L. et al. Long-term results of the surgical management of intermittent exotropia[J]. J Aapos,**14**:298–304. (2010).
12. Shen Jiang, H. & Chaoqun, L. Hong, et al. Changes in the proportion of inpatient surgery for acute concomitant Esotropia and intermittent Exotropia before and after the outbreak of covid-19[J] Chinese Journal of Eye, **31** (6): 454–457. (2022).
13. Wei, D. & Jing, F. Mechanism of interaction between myopia and intermittent Exotropia and key clinical scientific issues[J]. Chinese Journal of strabismus and pediatric ophthalmology, **28** (01): 36–38. (2020).
14. Ekdawi, N. S. et al. The development of myopia among children with intermittent exotropia[J].**149** (3), 503–507 (2010).
15. Tang, S. M. et al. Refractive errors and concomitant strabismus: A systematic review and meta-analysis[J]. Sci Rep, **12**, 6: 35177. (2016).
16. Kekunnaya, R., Chandrasekharan, A. & Sachdeva, V. Management of strabismus in myopes[J]. Middle East Afr J Ophthalmol,**22**:298–306. (2015).
17. Audrey, C., Lipika, R. & Linley, S. Comitant horizontal strabismus: an Asian perspective[J].The British journal of ophthalmology, **91**(10): 1337–1340 (2007).
18. Chen Danni, Li, R. et al. Prevalence, incidence and risk factors of strabismus in a Chinese population-based cohort of preschool children: the Nanjing eye Study[J]. The British journal of ophthalmology, **105**(9): 1203–1210. (2021).
19. Goss, D. A. & Zhai, H. Clinical and laboratory investigations of the relationship of accommodation and with refractive error. A literature review [J]. Documenta ophthalmologica Advances in ophthalmology, **86**(4): 349–380. (1994).
20. Choi, J., Kim, S. J. & Yu, Y. S. Initial postoperative deviation as a predictor of long term outcome after surgery for intermittent exotropia[J]. J AAPOS, **15**(3): 224–229. (2011).
21. Kim, T. W., Kim, J. H. & Hwang, J. M. Long-term outcome of patients with large over correction following surgery for exotropia[J].**219** (4), 237–242 (2005).

Acknowledgements

Author contributions
Conception and design: Haihua Liu; Data collection: Rongjun Liu, Ruiying Li; Analysis and interpretation: Haihua Liu, Kaixiu Li; writing and editing: Haihua Liu, Rongjun Liu; Overall responsibility: Haihua Liu, Rongjun Liu.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to H.L. or R.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.