Article published online: 2023-09-29

External Validation of a Clinical Nomogram for Predicting Intracranial Hematoma Following Head Computed Tomography in Pediatric Traumatic Brain Injury

Validação externa de um nomograma clínico para previsão de hematoma intracraniano após tomografia computadorizada de cabeça em lesão cerebral traumática pediátrica

Apisorn Jongjit[1] Thara Tunthanathip[1]

[1] Department of Surgery, Division of Neurosurgery, Faculty of Medicine, Prince of Songkla University, Songkhla, Thailand

Arq Bras Neurocir 2023;42(3):e226–e232.

Address for correspondence Thara Tunthanathip, Department of Surgery, Division of Neurosurgery, Faculty of Medicine, Prince of Songkla University, Songkhla, 90110, Thailand (e-mail: tsus4@hotmail.com).

Abstract

Introduction Over-investigation of head computed tomography (CT) has been observed in children with TBI. Long-term effects from a head CT brain scan have been addressed and those should be balanced. A nomogram is a simple prediction tool that has been reported for predicting intracranial injuries following a head CT of the brain in TBI children in literature. This study aims to validate the performance of the nomogram using unseen data. Additionally, the secondary objective aims to estimate the net benefit of the nomogram by decision curve analysis (DCA).

Methods We conducted a retrospective cohort study with 64 children who suffered from traumatic brain injury (TBI) and underwent a CT of the brain. Nomogram's scores were assigned according to various variables in each patient; therefore sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy and F1 score were estimated by the cross-tabulation of the actual results and the predicted results. Additionally, the benefits of a nomogram were compared with "None" and "All" protocols using DCA.

Results There were 64 children with TBI who underwent a head CT in the present study. From the cross-tabulation, the nomogram had a sensitivity of 0.60 (95%CI 0.29–0.90), specificity of 0.96 (0.91–1.0), PPV of 0.75 (0.44–1.0), NPV of 0.92 (0.86–0.99), accuracy of 0.90 (0.83–0.97), and an F1 score of 0.66 (0.59–0.73). Also, the area under the curve was 0.78 which was defined as acceptable performance. For the DCA at 0.1

Keywords

- ► External validation
- ► Nomogram
- ► intracranial hematoma
- ► pediatric traumatic brain injury
- ► head injury

received March 29, 2021
accepted June 16, 2021

DOI https://doi.org/10.1055/s-0043-1775579.
ISSN 0103-5355.

Thieme Revinter Publicações Ltda., Rua do Matoso 170, Rio de Janeiro, RJ, CEP 20270-135, Brazil

high-risk threshold, the net benefit of the nomogram was 0.75, whereas the "All" protocol had the net benefit of 0.40 which was obviously different.

Conclusion A nomogram is a suitable method as an alternative prediction tool in general practice that has advantages over other protocols.

Resumo

Introdução A investigação excessiva da tomografia computadorizada (TC) de crânio tem sido observada em crianças com TCE. Os efeitos a longo prazo de uma tomografia computadorizada de crânio foram abordados e devem ser equilibrados. Um nomograma é uma ferramenta de predição simples que foi relatada na literatura para prever lesões intracranianas após uma tomografia computadorizada de crânio em crianças com TCE. Este estudo tem como objetivo validar o desempenho do nomograma usando dados não vistos. Adicionalmente, o objetivo secundário visa estimar o benefício líquido do nomograma por meio da análise da curva de decisão (DCA).

Métodos Realizamos um estudo de coorte retrospectivo com 64 crianças que sofreram traumatismo cranioencefálico (TCE) e foram submetidas a tomografia computadorizada de crânio. As pontuações do Nomograma foram atribuídas de acordo com diversas variáveis em cada paciente; portanto, sensibilidade, especificidade, valor preditivo positivo (VPP), valor preditivo negativo (VPN), acurácia e escore F1 foram estimados pela tabulação cruzada dos resultados reais e dos resultados previstos. Além disso, os benefícios de um nomograma foram comparados com os protocolos "Nenhum" e "Todos" usando DCA.

Resultados Houve 64 crianças com TCE que foram submetidas a tomografia computadorizada de crânio no presente estudo. A partir da tabulação cruzada, o nomograma apresentou sensibilidade de 0,60 (IC95% 0,29–0,90), especificidade de 0,96 (0,91–1,0), VPP de 0,75 (0,44–1,0), VPN de 0,92 (0,86–0,99), acurácia de 0,90 (0,83–0,97) e uma pontuação F1 de 0,66 (0,59–0,73). Além disso, a área sob a curva foi de 0,78, definida como desempenho aceitável. Para o DCA no limiar de alto risco de 0,1, o benefício líquido do nomograma foi de 0,75, enquanto o protocolo "Todos" teve o benefício líquido de 0,40, o que foi obviamente diferente.

Conclusão Um nomograma é um método adequado como ferramenta alternativa de predição na prática geral que apresenta vantagens sobre outros protocolos.

Palavras-chave
- validação externa
- nomograma
- hematoma intracraniano
- lesão cerebral traumática pediátrica
- ferimento na cabeça

Introduction

Mortality and physical disabilities following traumatic brain injury (TBI) in children have been a concern of major public health problems.[1,2] Head computed tomography (CT) is the gold standard investigation of intracranial injuries. According to Larson et al, the rate of head CT in patients following TBI rose from 13.1% in 1996 to 40.7% in 2007.[3] However, the long-term side effects of CT have been mentioned. From prior studies, children who underwent CTs between 1985–2002, were significantly associated with leukaemia and brain tumors.[4,5] Therefore, the balancing of the unnecessary ionizing radiation exposure in children has been considered in over-investigations.

Investigation criteria has been performed and proposed from previous studies, for example, Children's Head Injury Algorithm for the Prediction of Important Clinical Events,[6] Canadian Assessment of Tomography for Childhood Head Injury,[7] and the Pediatric Emergency Care Applied Research Network[8] that are clinical prediction rules to identify children who need a head CT following mild TBI. Nomogram is one of the clinical prediction tools that has been used for predicting clinical outcomes and prognosis in various neurological conditions such as TBI, neuro-oncology and neurosurgical complications. According to Tunthanathip et al., a clinical nomogram was developed for the prediction of intracranial injuries following TBI from 900 TBI children in 2009–2018. The performance of the prediction tool was reported at an acceptable level as follows: accuracy (0.83), sensitivity (0.42), specificity (1.00), positive predictive value (1.00), and negative predictive value (0.81).[9] Also, this nomogram was further developed as a web-based application for user-friendly application in general practice.[9]

External validation is one of the processes to estimate the nomogram's performance using new and unseen data.[10] Therefore, this study aimed to validate the performance of the nomogram at predicting intracranial injuries following a head CT which was proposed from a prior study. Moreover, the secondary objective aimed to estimate the net benefit of the nomogram by decision curve analysis.

Methods

Study Designs and Study Population

A retrospective cohort study design was performed with patients suffering from TBI registered in the Trauma Registry of the trauma center of southern Thailand between January 2019 and December 2020. Patients were excluded for the following reasons: (1) patients died before arrival or at the emergency department; (2) patients who did not have a head CT. In detail, electronic medical records were reviewed to collect clinical characteristics, treatment, and functional outcomes. The findings from the head CT were evaluated by a neurosurgeon. Severities of TBI were defined according to the Glasgow Coma Scale (GCS) as follows: patients with a GCS score of 13–15 were defined as mild TBI, moderate TBI were patients with a GCS score of 9–12, and severe TBI were patients with a GCS score of 3–8 [4]. The hospital-discharge Glasgow Outcome Scale (GOS) was estimated in the present study. In detail, GOS was divided into 5 scores as follows: Death (1 score), vegetative state (2 scores), severe disability (3 scores), moderate disability (4 scores), good recovery (5 scores).[2,11] Moreover, the GOS was dichotomized into unfavorable outcome (GOS of 1–3) and favorable outcome (GOS of 4–5) for binary proposes.[11] The study was approved by the institutional research ethical committees.

Statistical Analysis

Descriptive statistics were performed for describing the baseline characteristics of the present cohort: mean with standard deviation (SD) or median with interquartile range (IQR) were used for describing continuous variables, while the categorical variables were described in percentages.

According to the primary objective, scoring for the present cohort was performed based on a clinical nomogram of Tunthanathip et al.[9] Nomogram scores were assigned according to various variables in each child; therefore, the cross-tabulation between the actual result and predicted result was done to estimate the nomogram's performance. In detail, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy and F1 score were estimated from the cross-tabulation. Additionally, the receiver operating characteristic (ROC) curve and area under the curve (AUC) were performed. An AUC of $\geq$ 0.7 was defined as acceptable performance, whereas an AUC of $\geq$0.8 and $\geq$0.9 was defined as good and excellent performance, respectively.[12,13]

For the secondary objective, a decision curve analysis (DCA) was conducted to evaluate the benefit of the nomogram compared with other protocols: "None" and "All" protocols. The cost-benefit ratio was used at a threshold of 0.1 according to prior studies.[14–16] The statistical analysis was done by the R version 3.6.2 software (R Foundation, Vienna, Austria).

Results

There were 204 children with TBI in the present study with 140 children being excluded because they did not undertake

Table 1 Demographic data of the present cohort ($N = 64$)

Factor	N (%)
Gender	
Male	39 (60.9)
Female	25 (39.1)
Age -month	
< 60	5 (7.8)
≥ 60	59 (92.2)
Mean of age- month (SD)	92.2 (7.6)
Injured mechanism	
Motorcycle crash	22 (34.4)
Fall at ground level	19 (29.7)
Object hit at the head	7 (10.9)
Pedestrians' injury	6 (9.4)
Bicycle accident	6 (9.4)
Vehicle crash	3 (4.7)
Fall from height	1 (1.6)
Road traffic injury	21 (48.4)
Sign and symptoms	
Scalp wound/hematoma	41 (64.1)
Loss of consciousness	21 (32.8)
Amnesia	19 (29.7)
Vomiting	7 (10.9)
Hypotension	4 (6.3)
Seizure before CT of the brain	2 (3.1)
Bleeding per nose/ear	2 (3.1)
Motor weakness	1 (1.6)
Initial Glasgow Coma Scale score	
13–15	57 (89.1)
9–12	1 (1.6)
3–8	6 (9.4)
Pupillary light reflex	
Normal reactivity both eyes	62 (96.9)
Fixed one eye	2 (3.1)
Fixed both eyes	–
Positive findings on CT of the brain	10 (15.6)
Calvarium skull fracture ($N = 3$)	3 (4.7)
Linear	2 (3.1)
Compound depressed	1 (1.6)
Basilar skull fracture	3 (4.7)
Epidural hematoma	3 (4.7)
Subdural hematoma	4 (6.3)
Contusion	1 (1.6)
Brainstem hemorrhage	1 (1.6)
Subarachnoid hemorrhage	5 (7.8)
Intraventricular hemorrhage	2 (3.1)

Table 2 Nomogram score of Tunthanathip et al.[9]

Variable	Score
Age group	
<= 5 years	0
> 5 years	14
Road traffic injury	
No	0
Yes	9
Loss of consciousness	
No	0
Yes	12
Motor weakness	
No	0
Yes	52
Scalp injury	
No	0
Yes	31
Bleeding per nose/ear	
No	0
Yes	100
Glasgow Coma Scale score	
13–15	0
9–12	37
3–8	72
Pupillary light reflex	
React both eyes	0
Fixed one eye	69
Fixed both eyes	35

*Prediction of positive intracranial injury used the cut off of > 0.5 probability (total score more than 79)

a head CT. Hence, the present study enrolled 64 children whose baseline characteristics are presented in ►**Table 1**. The mean age was 92.2 months (SD 7.6), with a range of 4–168 months, whereas the median age was 384 (IQR 132). For the mechanism of injury, road traffic accidents were found in 48.4% of all cases. A motorcycle crash was the most common cause of injury, whereas a fall at ground level was found in 29.7%. More than two-thirds of children had a scalp injury and post-traumatic seizure was observed in 3.1% of them. According to the severity of TBI, major patients were mild TBI and 11% of the cohort were moderate-severe TBI. Therefore, intracranial injuries were found at 15.6%. Subdural hematoma and subarachnoid hemorrhage were common findings following a CT of the brain.

Therefore, children in the present cohort were individually allocated scores as shown in ►**Table 2**. The prediction of positive intracranial injury was assigned when the total score was more than 79 (probability of positive results more than 0.5). Cross-tabulation between actual results and predicted results are presented in ►**Table 3**. From the cross-tabulation, the nomogram's sensitivity, specificity, PPV, NPV, accuracy, and F1 score using the unseen data was 0.60 (95%CI 0.29–0.90), 0.96 (0.91–1.0), 0.75 (0.44–1.0), 0.92 (0.86–0.99), 0.90 (0.83–0.97), respectively. Additionally, the F1 score was 0.66 (0.59–0.73) and the AUC was 0.78, as shown in ►**Fig. 1**.

For the secondary objective, DCA was performed for evaluating the net benefit of the nomogram compared with other situations, as shown in ►**Fig. 2A**. In detail, the DCA comprises of three lines; None (black line), All (gray line), and Nomogram (red line). "None" means nobody received a head CT brain in the present study; therefore, no net benefit is observed in Y-axis. "All" means a head CT is performed on all children, while "Nomogram" means using the nomogram score in the present cohort for selecting head CT. Vicker et al. used the high-risk threshold of 0.1 (cost: benefit ratio or harm: benefit ratio)[14] That meant that 1 normal person was harmed from treatment/investigation (such as head CT with unnecessary radiation exposure) and 9 actual patients underwent necessary treatment/investigation from a total of 10 children. When we set the harm benefit ratio at 1:9, the net benefit of the nomogram is higher than the head CT all cases protocol (All), as shown in ►**Fig. 2B**.

Discussion

The overall performance of the nomogram was at an acceptable level for predicting intracranial injury in pediatric TBI when we performed temporal external validation. We observed the stability of nomogram's performance in variations of baseline risk (intercept) and covariate effects (regression coefficients) of the prediction model in different time periods.[17,18] The tool had a high specificity and PPV that may be useful for ruling in children who were at high risk of intracranial injury. According to Baeyens et al., SPIN is the acronym for 'Specific test when Positive rules IN the disease' and SPIN relates with the high specificity and high PPV.[19]

Moreover, the DCA was plotted in the present study, which is a novel framework for estimating prediction tools by Vicker et al in 2006.[14] In the field of oncology, Calster et al.

Table 3 Cross-tabulation between actual and predicted results

	Actual results of head CT	
Predicted results of head CT	**Positive finding**	**Negative finding**
Predicted positive finding	6	2
Predicted negative finding	4	52

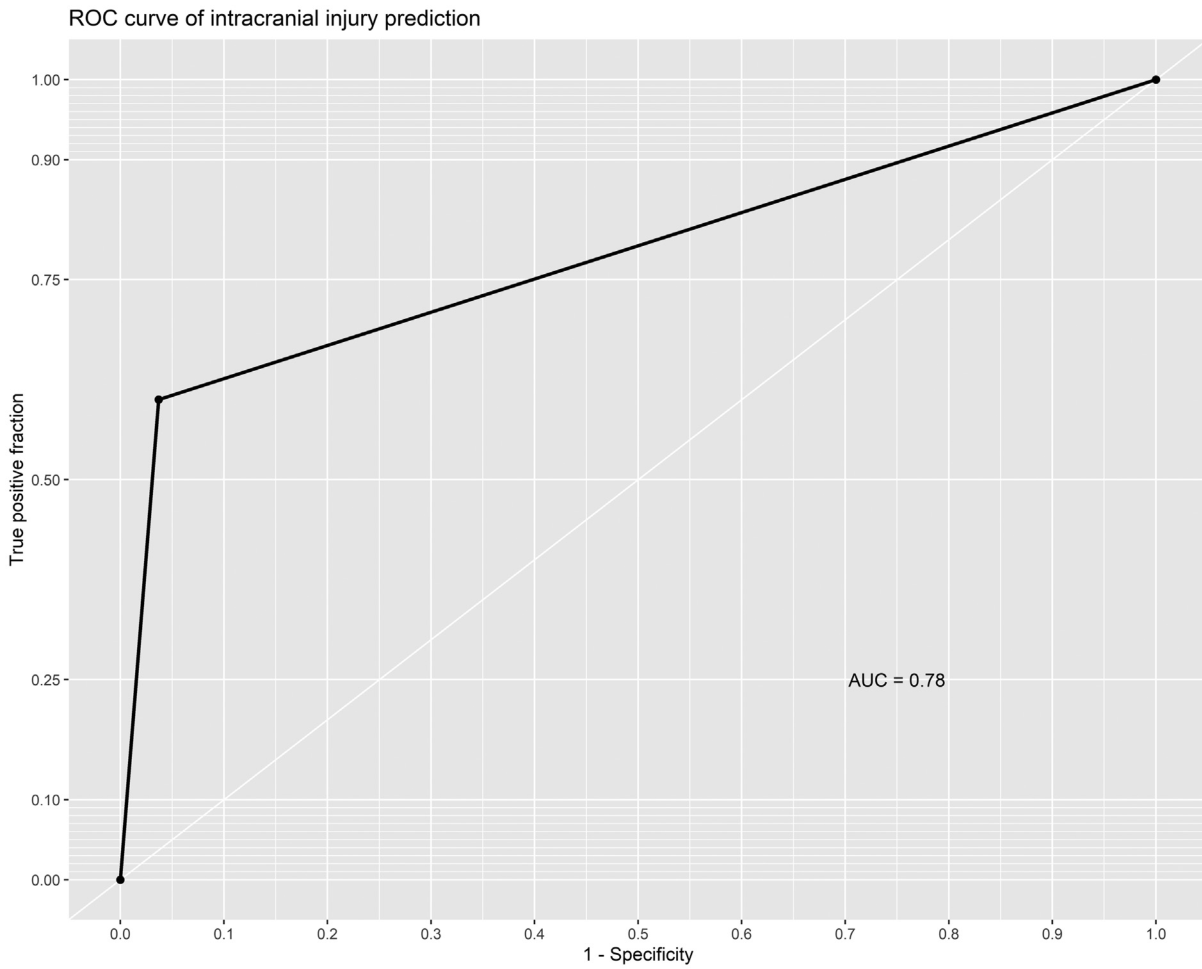

Fig. 1 Receiver operating characteristic curve with area under the curve.

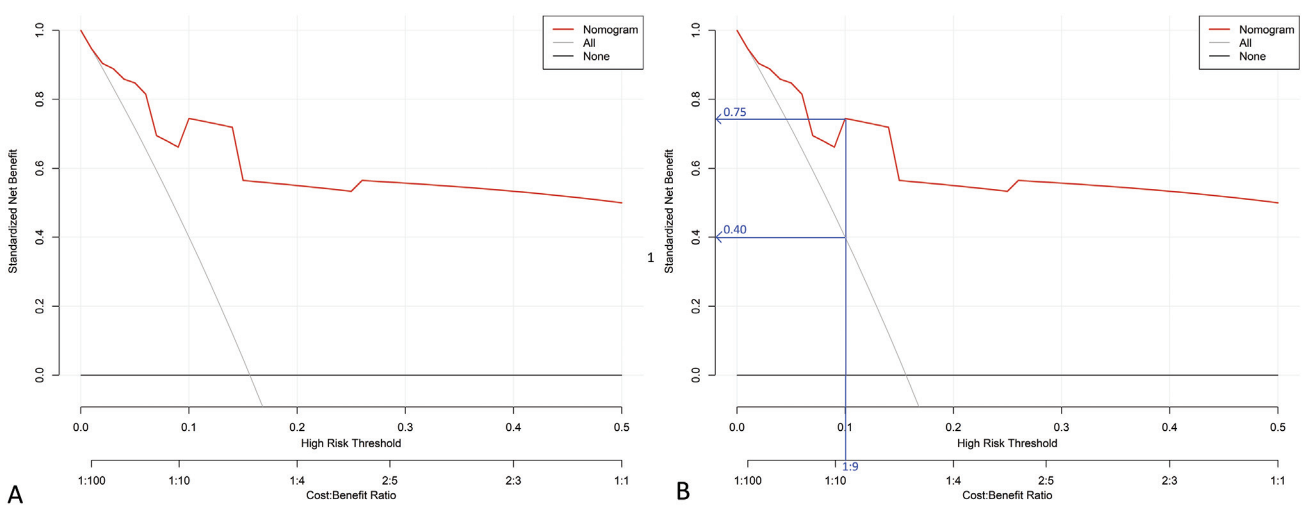

Fig. 2 (A) Decision curve analysis of nomogram. (B) Decision curve analysis with comparison between "All" protocol (gray line) and "Nomogram" protocol (red line). At cost: benefit ratio of 1:9, net benefit of "Nomogram" protocol is higher than "All" protocol (blue line).

used DCA for evaluating the net benefit of the prediction model for high-grade prostate cancer to select who should undertake a biopsy.[16] Therefore, DCA was concluded that it could help the clinicians to make better clinical decisions for treatment or investigation. As a result of the present study, the predictive model of the nomogram has estimated a benefit using DCA and found that it had potential value for implication in general practice.[10,20]

A nomogram is one of the clinical prediction tools that has been used for predicting various outcomes such as neuro-oncology,[10] trauma,[9,20] and various clinical outcomes.[21] Because the scoring system of the nomogram was quite difficult to remember, a web-based application of the nomogram has been developed in the literature review.[9] Moreover, machine learning algorithms have been proposed as alternative approaches for predicting clinical outcomes. Tunthanathip et al. used various algorithms of machine learning and found that the naive Bayes algorithm was highlighted for the prediction of infection following neurosurgical operations.[22] The comparison of predictive performances among various clinical prediction tools should be conducted in the future for selecting the best predictive performance. Hence, the tools will be deployed in general practice.

However, certain limitations should be recognized. First, although high accuracy was observed in the present study, the imbalance of negative and positive findings on head CT may be misleading.[23] Therefore, the F1 score has been suggested for estimating in this situation. The F1 score is calculated from the weighted average of PPV (precision) and sensitivity (recall). This tool may not be appropriate to use for a screening tool, because diminishing vales of recall and F1 score were observed. As mentioned above, the nomogram in the present study may be used for ruling in high-risk patients with an accepted F1 score.[24] Second, the sample size was limited in the present study; therefore, a multicenter study should be conducted in the future to increase the number of TBI children.

For future study, geographic external validation should be performed for estimating the generalizability from differences of both baseline risk and covariate effects of the nomogram's predictive model in different settings and time periods.[17,25] Also, an impact analysis should be conducted to evaluate the diminishing rate of head CT over-investigation in children.[26]

Conclusion

A nomogram is a suitable method for applying an alternative prediction tool in general practice that has advantages over other protocols.

Abbreviations Used in this Paper

AUC: Area under the curve, CT: computed tomography, DCA: decision curve analysis, GCS: Glasgow Coma Scale, GOS: Glasgow Outcome Scale, IQR: interquartile range, NPV: negative predictive value, PPV: positive predictive value, ROC: Receiver operating characteristic, SD: standard deviation, SPIN: Specific test when Positive rules IN the disease, TBI: traumatic brain injury

Declarations

All procedures performed in the study that involved studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee or both and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Author Contributions

AJ and TT conceived the study and designed the method. TT supervised the conduct of the data collection. AJ and TT managed the data, including quality control. AJ and TT provided statistical advice on the study design and analyzed the data and AJ drafted the manuscript, and TT contributed substantially to its revision. TT takes responsibility for the paper.

Funding

None

Conflict of Interest

The authors declare to have no competing interests

Acknowledgment

The authors would like to offer their special thanks to Professor. Nakornchai Phuenpathom and Associate professor. Sakchai Sae-heng for their advice about manuscript preparation.

References

1 The Centers for Disease Control and Prevention CDC newsroom [internet]. 2020 [cited 2020 Jul 5]. Available from: https://www.cdc.gov/media/releases/2018/p0904-tbi-guidelines.html
2 Taweesomboonyat T, Kaewborisutsakul A, Tunthanathip T, Saeheng S, Oearsakul T. Necessity of In-hospital Neurological Observation for Mild Traumatic Brain Injury Patients with Negative Computed Tomography Brain Scans. J Health Sci Med Res 2000; 38:267–274
3 Larson DB, Johnson LW, Schnell BM, Salisbury SR, Forman HP. National trends in CT use in the emergency department: 1995-2007. Radiology 2011;258(01):164–173
4 Sheppard JP, Nguyen T, Alkhalid Y, Beckett JS, Salamon N, Yang I. Risk of brain tumour induction from pediatric head CT procedures: A systematic literature review. Brain Tumor Res Treat 2018; 6(01):1–7
5 Pearce MS, Salotti JA, Little MP, et al. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. Lancet 2012;380 (9840):499–505
6 Dunning J, Daly JP, Lomas JP, Lecky F, Batchelor J, Mackway-Jones KChildren's head injury algorithm for the prediction of important clinical events study group. Derivation of the children's head injury algorithm for the prediction of important clinical events decision rule for head injury in children. Arch Dis Child 2006;91 (11):885–891
7 Osmond MH, Klassen TP, Wells GA, et al; Pediatric Emergency Research Canada (PERC) Head Injury Study Group. CATCH: a clinical decision rule for the use of computed tomography in children with minor head injury. CMAJ 2010;182(04): 341–348
8 Kuppermann N, Holmes JF, Dayan PS, et al; Pediatric Emergency Care Applied Research Network (PECARN) Identification of children at very low risk of clinically-important brain injuries after head trauma: a prospective cohort study. Lancet 2009;374 (9696):1160–1170
9 Tunthanathip T, Duangsuwan J, Wattanakitrungroj N, Tongman S, Phuenpathom N. Clinical nomogram predicting intracranial injury in pediatric traumatic brain injury. J Pediatr Neurosci 2020;15 (04):409–415
10 Tunthanathip T, Ratanalert S, Sae-Heng S, et al. Prognostic Factors and Nomogram Predicting Survival in Diffuse Astrocytoma. J

Neurosci Rural Pract 2020;11(01):135–143. Doi: 10.1055/s-0039-3403446

11 Tunthanathip T, Phuenpathom N. Impact of road traffic injury to pediatric traumatic brain injury in southern Thailand. J Neurosci Rural Pract 2017;8(04):601–608

12 Marques A, Almeida S, Carvalho J, Cruz J, Oliveira A, Jácome C. Reliability, Validity, and Ability to Identify Fall Status of the Balance Evaluation Systems Test, Mini-Balance Evaluation Systems Test, and Brief-Balance Evaluation Systems Test in Older People Living in the Community. Arch Phys Med Rehabil 2016;97(12):2166–2173.e1. Doi: 10.1016/j.apmr.2016.07.011

13 Tunthanathip T, Oearsakul T. Machine Learning Approaches for Prognostication of Newly Diagnosed Glioblastoma. Int J Nutr Pharmacol Neurol Dis 2021;11:57–63

14 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26(06): 565–574. Doi: 10.1177/0272989. Doi: × 06295361

15 Kerr KF, Brown MD, Zhu K, Janes H. Assessing the Clinical Impact of Risk Prediction Models With Decision Curves: Guidance for Correct Interpretation and Appropriate Use. J Clin Oncol 2016;34(21):2534–2540. Doi: 10.1200/JCO.2015.65.5654

16 Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. Eur Urol 2018;74(06):796–804. Doi: 10.1016/j.eururo.2018.08.038

17 Austin PC, van Klaveren D, Vergouwe Y, Nieboer D, Lee DS, Steyerberg EW. Validation of prediction models: examining temporal and geographic stability of baseline risk and estimated covariate effects. Diagn Progn Res 2017;1:12. Doi: 10.1186/s41512-017-0012-3

18 Girardat-Rotar L, Braun J, Puhan MA, Abraham AG, Serra AL. Temporal and geographical external validation study and extension of the Mayo Clinic prediction model to predict eGFR in the younger population of Swiss ADPKD patients. BMC Nephrol 2017;18(01): 241. Doi: 10.1186/s12882-017-0654-y

19 Baeyens JP, Serrien B, Goossens M, Clijsen R. Questioning the "SPIN and SNOUT" rule in clinical testing. Arch Physiother 2019; 9:4. Doi: 10.1186/s40945-019-0056-5

20 Tunthanathip T, Udomwitthayaphiban S. Development and Validation of a Nomogram for Predicting the Mortality after Penetrating Traumatic Brain Injury. Bull Emerg Trauma 2019;7(04): 347–354. Doi: 10.29252/beat-070402

21 Devin CJ, Bydon M, Alvi MA, et al. A predictive model and nomogram for predicting return to work at 3 months after cervical spine surgery: an analysis from the Quality Outcomes Database. Neurosurg Focus 2018;45(05):E9. Doi: 10.3171/2018.8.FOCUS18326

22 Tunthanathip T, Sae-Heng S, Oearsakul T, Sakarunchai I, Kaewborisutsakul A, Taweesomboonyat C. Machine learning applications for the prediction of surgical site infection in neurological operations. Neurosurg Focus 2019;47(02):E7. Doi: 10.3171/2019.5.FOCUS19241

23 Wood T What is the F-score? [Internet]. 2021 [cited 2021 Mar 5]. Available from: https://deepai.org/machine-learning-glossary-and-terms/f-score

24 Shung KP Accuracy, Precision, Recall or F1? [Internet]. 2018 [cited 2021 Mar 5]. Available from: https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9

25 Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. J Clin Epidemiol 2008;61(11):1085–1094. Doi: 10.1016/j.jclinepi.2008.04.008

26 Moons KG, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012;98(09):691–698. Doi: 10.1136/heartjnl-2011-301247

