



External validation of the IHXGboost-P model to predict incisional hernia after midline laparotomy

Edgard Efren Lozada Hernandez^{1,2} · Tania A. Ramirez-delreal^{3,4} · Dagoberto Armenta-Medina^{1,3} · Sebastián Salazar-Colores⁵

Received: 12 November 2025 / Accepted: 22 December 2025

© The Author(s), under exclusive licence to Springer-Verlag France SAS, part of Springer Nature 2026

Abstract

Background Incisional hernia (IH) is a significant complication that occurs after midline laparotomy and is associated with high morbidity and economic impacts. A fundamental goal of preventing IH is to determine which patients are considered low- or high-risk, as modifications in prevention techniques have been justified in high-risk patients.

AIM of this study was to externally validate the IHXGBoost-P model to assess its accuracy, generalizability, and clinical applicability in an independent cohort.

Methods: A prospective cohort study was conducted in a tertiary hospital in Mexico (March 2021–December 2022) to externally validate the IHXGBoost-P model. Patients older than 18 years who underwent midline laparotomy and have a minimum follow-up of 24 months were included. The performance of the model was evaluated via area under the receiver operating characteristic curve (AUROC), accuracy, sensitivity, precision, specificity and calibration metrics.

Results Of the 438 patients analyzed, 62 (14.1%) developed IH. The model demonstrated good discriminative capacity (Accuracy: 0.94 ± 0.015) and calibration (Brier score: 0.051). Key predictors included the risk of surgical site infection (odds ratio (OR): 3.01, 95% CI: 2.32–3.91), previous surgery, and body mass index (BMI). The specificity (0.97 ± 0.013) was determined to be high and useful for identifying lowrisk patients.

Conclusions The IHXGBoost-P model is a reliable tool for predicting the risk of IH, with robust performance being observed in external validation. Its integration into clinical practice through a web application could optimize surgical decision-making to prevent IH.

Keywords Incisional Hernia · XGBoost · Artificial Intelligence · External Validation.

Introduction

Incisional hernia (IH) is a serious complication due to its high incidence (12.8% overall and 30% in high-risk patients) [1] and high recurrence (ranging from 40% to 70% depending on whether mesh was used during the initial repair) [2].

This incidence is influenced by several factors, such as the type of investigated population (e.g., bariatric surgery vs. emergency patients), the location of the incision (midline vs. outside of the midline), the duration of follow-up, and the method utilized (clinical or imaging methods) for the diagnosis [3].

IH has an enormous economic impact on health institutions [4]; moreover, it is an important cause of morbidity (e.g., pain) and has a negative effect on both patients' quality of life and their body image [3]–[4].

✉ Edgard Efren Lozada Hernandez
edgardlozada2@gmail.com

¹ INFOTEC (Centro De Investigacion E Innovacion En Tecnologías De La Informacion Y Comunicación), Aguascalientes, Mexico

² Hospital Regional de Alta Especialidad-IMSS Bienestar, Leon, Guanajuato, México

³ SECIHTI (Secretaría de Ciencia, Humanidades, Tecnología e Innovación), CDMX, Mexico

⁴ CentroGeo (Centro de Investigación en Ciencias de Información Geoespacial), Aguascalientes, México

⁵ CIO (Centro de investigaciones en Óptica), León Guanajuato, México

For all of the above-mentioned factors, the prevention of IH is a priority in international consensus. A fundamental goal of preventing IH is to determine which patients are considered low- or high-risk, as modifications in prevention techniques have been justified in high-risk patients. Furthermore, techniques involving closure or placement of a prophylactic mesh have been utilized to reduce complications [5]–[6].

Although several predictive scales have been developed to identify key risk factors (including emergency surgery, chronic obstructive pulmonary disease (COPD), BMI > 25, and surgical site infection (SSI)) [7–14], their clinical adoption has been limited. This limitation is due to significant methodological heterogeneity, various biases (such as overfitting), and a lack of robust external validation that confirms the usefulness of these scales in environments other than those in which the scales were developed (Table 1) [15].

Unlike traditional statistical models, machine learning models (including algorithms such as XGBoost) can capture

complex and nonlinear relationships between variables, thereby increasing the predictive accuracy [16]. Our group previously developed the IHXGBoost-P model to predict IH after midline laparotomy. This model demonstrated excellent performance in its developmental phase (area under the curve (AUC): 0.93) [17].

The objective of this study was to perform external validation of the IHXGBoost-P model to evaluate its accuracy, generalizability and clinical applicability in an independent cohort, thus facilitating its future implementation in medical practice.

Methods

Study design

A longitudinal and comparative prospective cohort study was conducted. The performance of the study and the reporting of its results were based on the TRIPOD-IA guidelines [18] Supplementary Information, Appendix A. This study was registered with the research ethics committee of the Hospital Regional de Alta Especialidad del Bajío IMSS-Bienestar, Guanajuato, Mexico (registration number: DPEI/CI/HRAEB/002/2021). Development and Internal validation were registered with Clinical Trials: NCT05718999.

Selection criteria

The following inclusion criteria were utilized: patients who were older than 18 years, who underwent midline laparotomy in either urgent or scheduled scenarios (regardless of background diagnoses) from March 2021 to December 2022 and who completed the minimum follow-up time period of 24 months to objectively determine the presence of IH. The exclusion criteria included patients who participated in other research protocols associated with laparotomy and for whom no information was available regarding the variables of the model.

Variables of IHXGBoost-P

The IHXGBoost-P model includes 13 variables selected through a structured, multi-phase process based on eight previously published predictive scales for IH. This approach integrated prior evidence, eliminated redundant or non-significant predictors, and retained only those consistently associated with IH across studies [17]. A detailed description of the variable selection procedure is provided in Supplementary Information, Appendix B.

This model consists of 13 variables. These values were obtained from the electronic file and included age (years),

Table 1 Predictive scales for IH

Author	Sample size	Characteristics	C-Statistic
Veljkovic et al., 2010 [7]	646	Open surgery; follow-up > 6 months; general abdominal surgery; development phase only; prospective study.	0.92(0.88–0.96)
Hernia project Goodenough et al., 2015 [13]	872	Open and laparoscopic surgery; 41-month follow-up; general abdominal surgery; development and external validation phases; prospective study.	0.77(0.74–0.80)
Cherla et al., 2017[14]			
Basta M et al., 2016 [8]	2161	Open and laparoscopic surgery; 28-month follow-up; bariatric surgery; development phase; retrospective study.	0.85(0.78–0.92)
Fisher et al., 2016 [9]	12,373	Open surgery; 32-month follow-up; general abdominal surgery; development phase; retrospective study.	0.77(0.74–0.88)
Lanni et al., 2016[10]	30,865	Open and laparoscopic surgery; 30-month follow-up; colon surgery; development phase; retrospective study.	0.70(0.69–0.71)
Tecce et al., 2017 [11]	2145	Open surgery; follow-up not specified; hysterectomy; development phase; retrospective study.	0.82(76–0.88.88)
Basta M et al., 2019 Pen Hernia Risk Calculator [12]	29,729	Open and laparoscopic surgery; 32-month follow-up; general abdominal surgery; development and internal validation phases; retrospective study.	0.83(0.76–0.89)

C-Statistic: reported as mean and 95% confidence interval

sex (female/male), colon surgery (yes/no), COPD (yes/no), ascites (yes/no), jaundice (yes/no), anemia (yes/no), emergency surgery (yes/no), previous surgery (yes/no), BMI (which was calculated by dividing weight by height squared and reported as kg/m²), ostomy (yes/no), oncological patient (yes/no) and the risk of infection marked in the web app as RISKWINF (risk of infection), which was obtained from the National Surgical Quality Improvement Program (NSQIP) Surgical Risk Calculator of the American College of Surgeons (ACS). This tool has multiple published validations and ensures a good estimation of risk, which is reported as a percentage from 0% to 100% [19]–[20]. The risk factors for each patient were entered into the risk calculator (<https://hi-eelh-app-1035249958862.us-central1.run.app>) by two researchers; additionally, this calculator is an open access tool [21].

The primary outcome was the presence of IH, detected either clinically or with an imaging study, as defined by the European Hernia Society (EHS). Specifically, the following EHS definition of IH was used: “any site of the abdominal wall with or without protuberance in the area of a postoperative scar perceptible or palpable by clinical examination or imaging” [3, 22]. The patient was assessed at 1 month, 3 months, 6 months, and subsequently every 6 months until completing at least two years of follow-up. If a hernia was detected clinically, no additional imaging study was performed. During all these assessments, an abdominal wall ultrasound was conducted, and in cases of diagnostic uncertainty, the patient underwent a computed tomography scan. Our hospital is a public institution; therefore, performing these imaging studies did not incur any additional cost for the patient.

Sample size

The sample size was calculated by using the R programming language in the R studio environment with the use of the library: pmsampsize [23]. A sample size of 395 patients was determined to be necessary for the external validation of the model on the basis of obtaining 56 events and an event ratio of 4.25 for each parameter used in the validation of the tool. Supplementary Information, Appendix C.

Structure of the model for external validation

For external validation, the final risk prediction model, based on the Extreme Gradient Boosting (XGBoost) algorithm, was applied to the independent cohort. The entire validation process was executed using the Python programming language within the Google Colab environment. The model was optimized prior to being locked to prioritize sensitivity, thereby ensuring a high capacity to correctly identify

patients at the greatest risk of IH. To ensure the integrity of this locked model, it was saved in.pkl format using the joblib library, thereby preserving its complete architecture and optimized parameters. Performance was evaluated using clinically relevant metrics, including the area under the receiver operating characteristic curve (AUC-ROC), sensitivity, specificity, and positive predictive value. Complete technical details regarding the algorithm parameters, optimization process, and code are available in the Supplementary Material, Appendix D.

Statistical analysis

Continuous baseline variables are expressed as the mean and standard deviation because these variables were normally distributed. Categorical variables are presented as numbers with percentages. The cohort was divided into two groups according to whether IH was present. Quantitative variables were compared via the Student's t-test, and qualitative variables were compared with the χ^2 test.

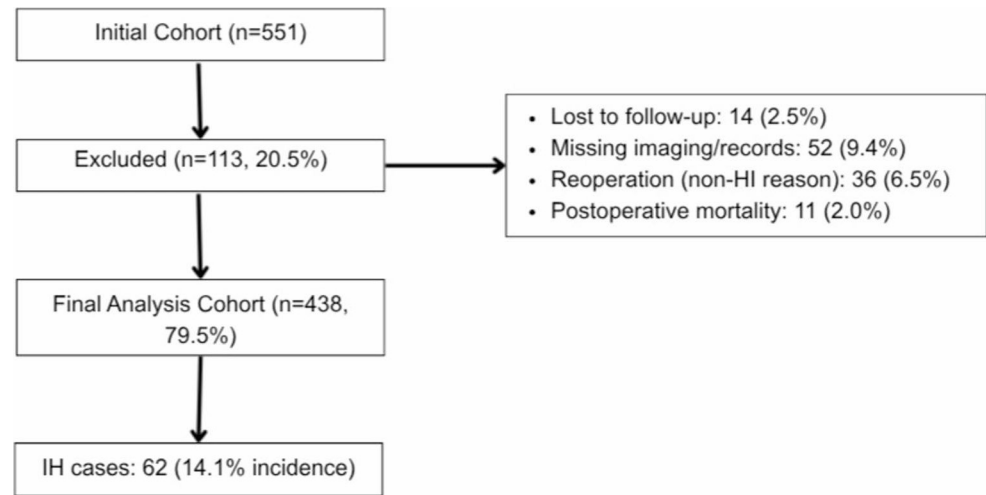
To evaluate the model, its predictive capacity was analyzed by using the ROC-AUC and the Youden index, as well as classification metrics such as sensitivity, specificity, predictive values (positive and negative), overall accuracy (Accuracy) and the balance between precision and accuracy, which involved the recall metric (F1 score). A calibration analysis was performed to evaluate the agreement between the predicted and observed probabilities by using goodness-of-fit tests and calibration graphs.

A Bayesian analysis was performed with the results of the confusion matrix, and the Fagan nomogram was constructed to determine the prevalence after the test. Both analyses were implemented from the post-test probabilities derived from the observed sensitivity, specificity and prevalence values by using the following libraries: caret for the confusion matrix and mada for the nomogram. Any value of $p < 0.05$ was considered to indicate statistical significance.

Results

Study population

A total of 551 patients were included in this study; 113 patients (20.5%) were excluded for various reasons, including a lack of completion of follow-up, unconfirmed presence of IH in their files, reoperations for causes other than IH, or death during the immediate postoperative period. Ultimately, 438 patients were included in the analysis, of whom 62 developed IH (representing 14.1% of the sample) (Fig. 1, STROBE format flowchart).

Fig. 1 STROBE format flowchart

Baseline characteristics

Significant predictors of IH included age (OR: 1.02, 95% CI: 1.00–1.38), colon surgery (OR: 2.51, 95% CI: 1.47–4.29), RISKWINF (OR: 3.01, 95% CI: 2.32–3.91), previous surgery (OR: 2.55, 95% CI: 1.24–5.27) and BMI (OR: 1.06, 95% CI: 1.00–1.12). Table 2.

Model performance

The IHXGBoost-P model achieved a ROC-AUC: 0.88 ± 0.011 , Accuracy: 0.94 ± 0.01 , Precision (PPV, positive predicted value): 0.84 ± 0.01 , Recall (sensitivity): 0.79 ± 0.05 , NPV (negative predicted value): 0.96 ± 0.01 , F1 score: 0.81 ± 0.02 and Specificity: 0.96 ± 0.03 . Compared with the accuracy of the internal validation assessment (0.93), the accuracy of the model increased by 0.01, which was likely due to population variability and exclusion.

The SHAP analysis corroborated the findings of the multivariate logistic regression, highlighting RISKWINF, previous surgery, age, colon surgery, and BMI as the most influential predictors of IH. This analysis added value by quantifying each variable's contribution to the model output and revealing non-linear interactions that traditional regression may overlook. Thus, SHAP not only confirmed the robustness of the adjusted model but also provided individualized interpretability, reinforcing the biological and clinical plausibility of the identified risk factors. Figure 2.

Calibration and discrimination

The discrimination plot (Fig. 3A) shows an area under the ROC curve (AUC=0.88), confirming that the model effectively distinguishes between patients who developed an IH and those who did not, thus supporting its predictive

Table 2 Bivariate analysis between predictive variables and the presence of IH in the external validation

Variable	IH (n=62)	No IH (n=376)	p	OR (CI95%)
Age (year)	54.7(15.8)	49.9(14.2)	0.026 *	1.02(1–1.38.38)
Sex	40(64.5)	220(58.5)	0.37	1.29(0.74–2.26)
Female	22(35.5)	156(41.5)		
Male				
Colon surgery	31(50)	123(33)	0.008	2.51(1.47–4.29)
COPD	5(8.1)	14(3.7)	0.12	2.64(0.92–7.58)
Ascites	11(17.5)	51(13.5)	0.39	1.62(0.8–3.31)
Jaundice	1(1.6)	24(6.3)	0.13	0.28(0.04–2.12)
Anemia	17(27.4)	175(46.2)	0.005	0.43(0.24–0.78)
Emergency surgery	26(41.9)	113(30.1)	0.06	1.63(0.94–2.82)
RISKWINF	9.9(3.2)	4.03(1.8)	0.001*	3.01(2.32–3.91)
Previous surgery	12(19)	32(8.4)	0.009	2.55(1.24–5.27)
BMI (Kg/m ²)	26.9(4.7)	25.5(4.6)	0.041*	1.06(1–1.12.12)
Ostomy	9(14.3)	34(9.2)	0.18	1.65(0.75–3.62)
Oncologic patient	49(79.4)	254(67.6)	0.07	1.62(0.89–3.16)

*Mean (SD), Student's *t*-test. Qualitative variables are reported as frequency and percentage, analyzed using the χ^2 test. *COPD* Chronic Obstructive Pulmonary Disease, *BMI* Body Mass Index, *RISKWINF* Risk of infection, score obtained from the ACS Surgical Risk Calculator

accuracy and clinical reliability. The calibration curve (Fig. 3B) compares the predicted probabilities generated by the IHXGBoost-P model with the observed incidence of IH, demonstrating good agreement, particularly across low- and

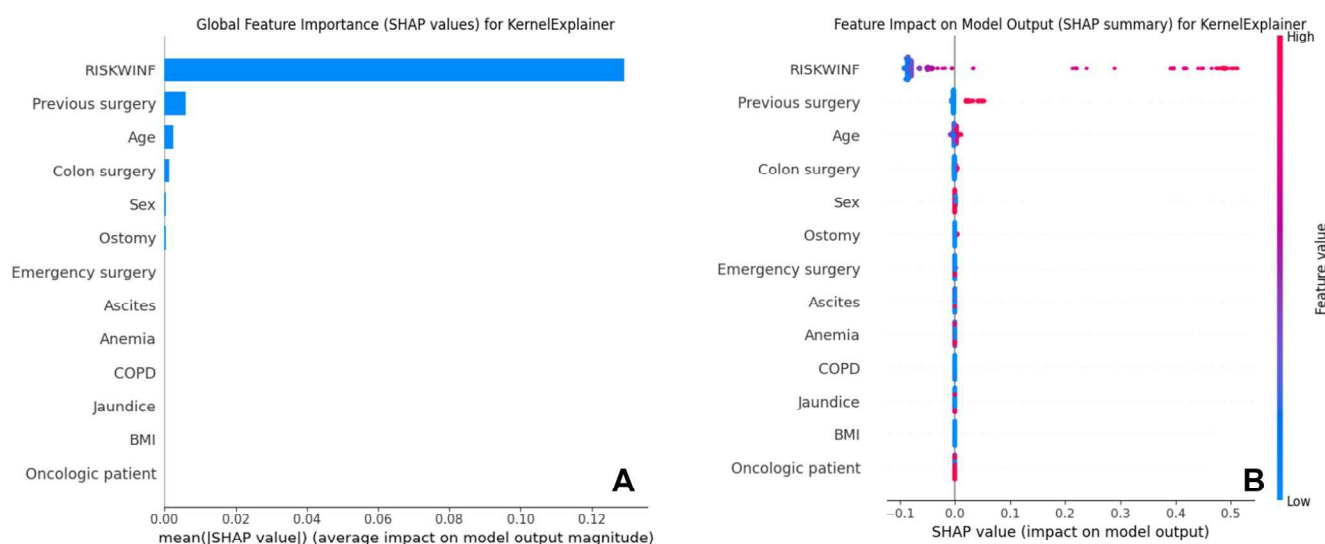


Fig. 2 SHAP values for the IHXGBoost-P model. **(A)** Global feature importance. **(B)** Feature impact on model output

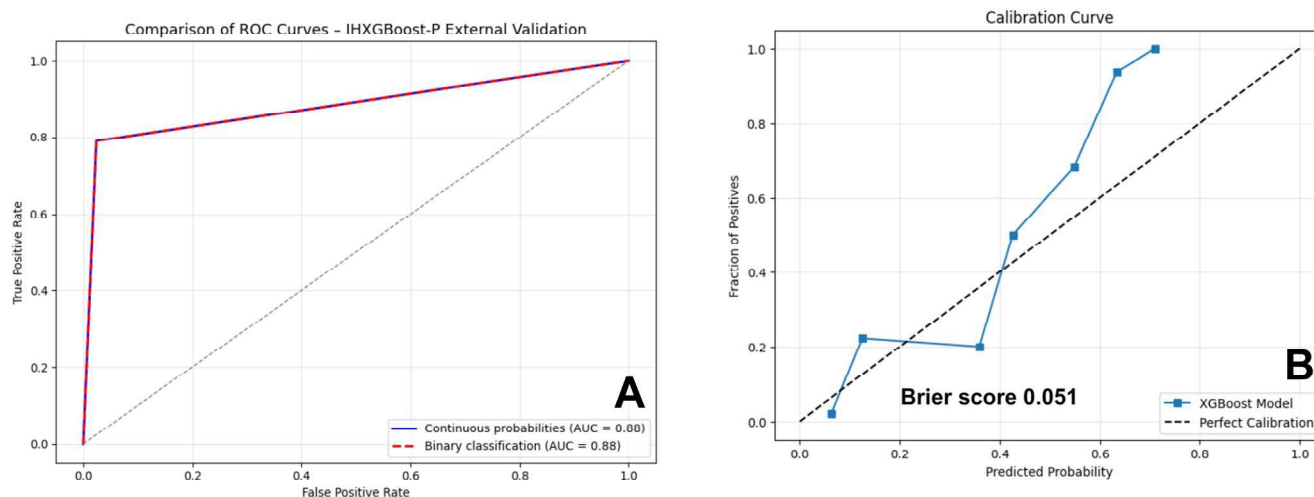


Fig. 3 **(A)** Discrimination (ROC-AUC) **(B)** Calibration (Brier score and Calibration plot)

intermediate-risk groups. The Brier score of 0.051 reflects excellent overall calibration. Figure 3 illustrates the model's performance in terms of calibration and discrimination.

In the Bayesian analysis, the Fagan nomogram, revealed that the IHXGBoost-P model demonstrated good diagnostic performance, significantly improving the post-test risk estimation and both confirming and ruling out IH with observed prevalences of 84% and 3%, respectively, which were outside the classic zone of uncertainty. Figure 4.

In the threshold analysis, the probability distribution of predicted risk (Fig. 5) demonstrated a clear separation between patients with and without IH. A 10% probability cut-off provided the best balance between sensitivity and specificity, yielding sensitivity of 0.89, specificity of 0.93, and an AUC of 0.96. This threshold accurately identified

high-risk patients while maintaining a low false-positive rate, supporting its use for clinical decision-making.

Discussion

The objective of this study was to externally validate the IHXGBoost-P model. The results demonstrated good discriminatory capacity (ROC-AUC=0.88, Accuracy 0.94) and calibration (Brier score of 0.051) for predicting the occurrence of IH in patients after midline laparotomy, similar to the results from its development and internal validation.

The model demonstrated an expected reduction in the ROC-AUC from 0.94 to 0.88 in the external validation assessment, which was most likely associated with

Fig. 4 Fagan nomogram for Bayesian post-test probability of IH

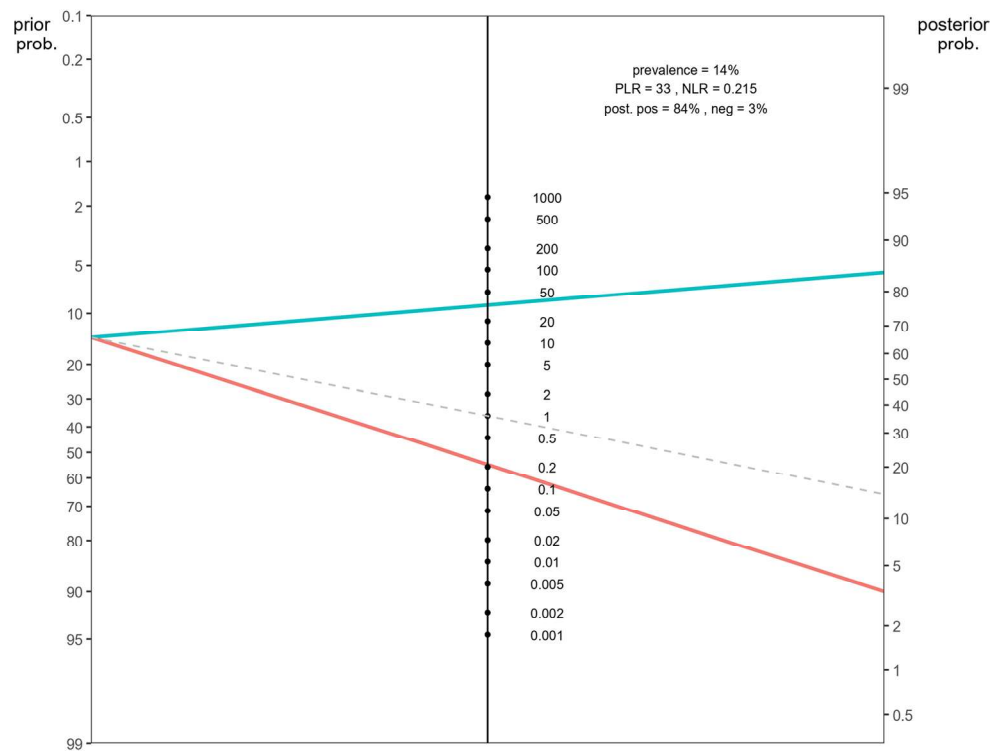
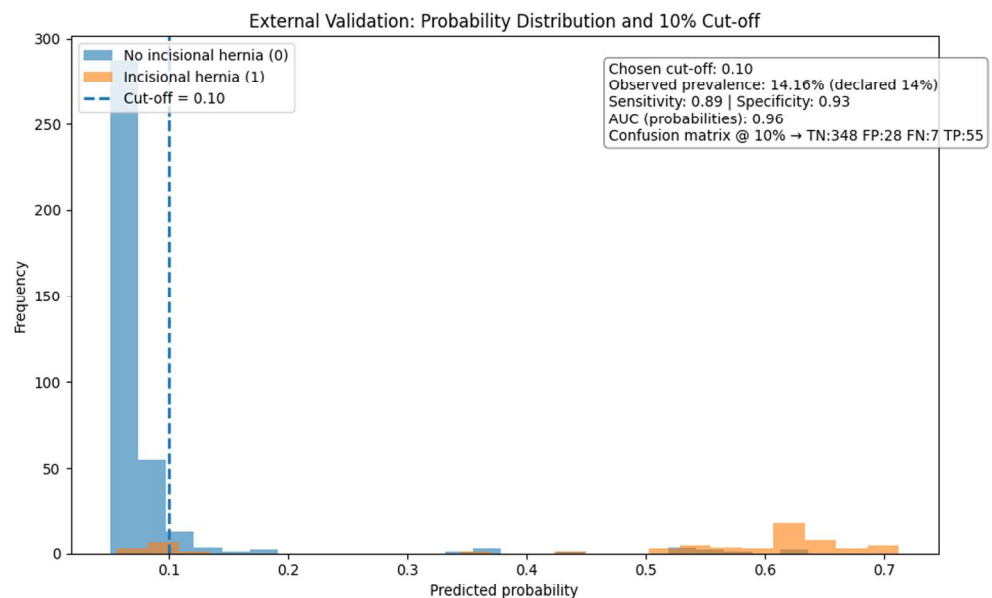


Fig. 5 Determination of the optimal cut-off point for IHXGBoost-P model



population variability. Its sustained specificity (0.96) supports its use to identify low-risk patients. Calibration and discrimination are fundamental factors for determining the validity of a prediction model; however, good calibration does not guarantee good discrimination (and vice versa) [24].

Of the 551 patients who met the inclusion criteria, 113 were excluded from the final sample due to lack of follow-up, unconfirmed IH in the files, reoperations for other causes, or early postoperative death. Moreover, this

exclusion was implemented prior to the analysis, and there were no “losses” of the final sample. Therefore, we do not believe that these losses compromised the statistical power of the model.

Currently, no similar model exists in the literature for predicting IH; Cole et al. [25] published a study in 2021 with similar characteristics but evaluated the prediction of acute dehiscence without assessing IH. Thus, the current research will have a considerable impact on determining which patients can be considered either high risk or low

risk, with the results being utilized to modify the technique accordingly.

IH and SSI are recognized as the main complications that occur after laparotomy and are associated with increased morbidity, hospital stays and costs. In the United States, agencies such as the Centers for Disease Control and Prevention (CDC) and the Joint Commission consider colon surgery and its complications (particularly SSI) to be key indicators of surgical quality. In our model, both variables significantly affected the occurrence of IH [26].

A meta-analysis by Tansawet et al. [15] revealed that only two of seven predictive scales included SSI in their final models, assessing it dichotomously (yes/no) and retrospectively, although they reported a strong association between SSI and IH (OR: 5.54). In our cohort, we confirmed this relationship, demonstrating that SSI increases the risk of IH threefold (OR: 3.00, 95% CI: 2.32–3.91).

A specific strength of our study involves the use of the NSQIP Surgical Risk Calculator of the ACS to prospectively estimate the risk of SSI, which allowed for the incorporation of this variable in advance in the model, thus increasing its clinical utility and predictive capacity [19]. Lozada et al. (2025) conducted the external validation of the ACS NSQIP Surgical Risk Calculator in patients undergoing midline laparotomy, demonstrating excellent diagnostic performance (AUC = 0.88; Brier score = 0.072) and good calibration for predicting postoperative SSI. A cut-off points of 6.6% optimally discriminated against low- and high-risk patients, allowing stratification that improved predictive accuracy and clinical applicability. These findings support the inclusion of the prospectively estimated RISKWINF in the IHXGBoost-P predictive model [27].

The Bayesian analysis and the Fagan nomogram were calculated by using the prevalence of the cohort and the metrics obtained from the confusion matrix. The post-test probability after a positive result was 84%, whereas it decreased to 3% after a negative result, consistent with a good ability to confirm and rule out IH at the utilized cutoff point of 10%, which demonstrated an optimal balance between sensitivity and specificity in the external validation.

Another strength of this study involved the performance of external validation in a different cohort compared with that used for development and internal validation. However, the same clinical characteristics were used; thus, selection bias was reduced. Another strength involved the correct identification of the variables that form the model, in addition to its prospective nature and the utilization of an adequate, timely follow-up, along with the diagnosis of the hernia. An additional strength included the development and use of a web app that allowed for the simple identification of patients and the completion of the variables comprising the model.

The effective clinical translation of a predictive model critically depends on its reproducibility and accessibility across diverse hospital settings. Recognizing that heterogeneity in resources among centers can limit the adoption of tools based on advanced statistical methods, the development of a free, intuitive web application constitutes the fundamental mechanism to overcome this barrier. This web app eliminates the need for local statistical resources, specialized expertise, or costly infrastructure. Any surgeon, in any center, can obtain a standardized risk prediction immediately using only a web browser. Thus, the tool not only validates the model from a statistical standpoint but also ensures its operational reproducibility and democratizes access, enabling universal and equitable implementation that is essential for a real-world impact on IH prevention.

Limitations include its single-center nature and focus on a specific population, without including other types of incisions (such as paramedian or transverse). Additionally, applicability to other ethnicities or countries may require multicenter external validation and local recalibration (including intercept and slope adjustments), as differences in variable incidence (e.g., proportions of patients undergoing surgery, colon surgery, SSI rates, or BMI) can modify baseline risks. A further limitation is that the current model predicts the patient's intrinsic risk without incorporating technical factors related to abdominal wall closure (such as suture technique or mesh use), which also influence the final risk of IH. We also acknowledge other considerations that contextualize our findings. The limited number of outcome events may affect the precision of certain performance estimates and precludes extensive subgroup analyses. Furthermore, this validation did not include a direct comparison with other existing risk prediction scores. Ultimately, the clinical impact and practical utility of this model must be evaluated in prospective implementation studies.

Conclusions

The IHXGBoost-P model accurately predicted the risk of IH (ROC-AUC: 0.88; Brier score: 0.051) in patients who underwent midline laparotomy. Its high specificity and prospective evaluation of the risk of SSI can allow for the use of personalized surgical strategies (such as the use of prophylactic mesh) to reduce the incidence of IH. Moreover, integrating this model into clinical workflows through a web tool could improve patient outcomes and reduce care costs.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10029-025-03579-2>.

Author contributions EELH: Conceptualization, Data curation, Methodology, Writing & Editing, Visualization. TARR: Conceptualization, Methodology, Writing-Review & Editing, Visualization. DAM: Methodology, Data Curation, Writing-Review & Editing, Visualization. SSC: Methodology, Data Curation, Writing-Review & Editing, Visualization.

Funding None.

Data availability The authors confirm that the data supporting the findings of this study are available upon reasonable request.

Declarations

Informed consent For this observational study with anonymized data, formal consent is not required.

Declarations of generative AI and AI-assisted technologies in the writing process During the preparation of this work the authors did not use any tool or service to modify/adapt/revise its content.

Conflict of interest All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

1. Bosanquet DC, Ansell J, Abdelrahman T, Cornish J et al (2015) Systematic review and meta-regression of factors affecting midline incisional hernia rates: analysis of 14,618 patients. *PLoS One* 21(9):e0138745. <https://doi.org/10.1371/journal.pone.0138745>
2. Bhardwaj P, Huayllani MT, Olson MA, Janis JE (2024) Year-over-year ventral hernia recurrence rates and risk factors. *JAMA Surg* 1(6):651–658. <https://doi.org/10.1001/jamasurg.2024.0233>
3. Deerenberg EB, Henriksen NA, Antoniou GA, Antoniou SA et al (2022) Updated guideline for closure of abdominal wall incisions from the European and American Hernia Societies. *Br J Surg* znac302. <https://doi.org/10.1093/bjs/znac302>
4. Yheulon C, Davis S Jr (2019) Adopting the STITCH trial: crossing the chasm from publication to practice. *JAMA Surg* 1(12):1087–1088. <https://doi.org/10.1001/jamasurg.2019.3358>
5. Jairam AP, Timmermans L, Eker HH, Pierik REGJM et al (2017) Prevention of incisional hernia with prophylactic onlay and sublay mesh reinforcement versus primary suture only in midline laparotomies (PRIMA): 2-year follow-up of a multicentre, double-blind, randomized controlled trial. *Lancet* 5(10094):567–576. [https://doi.org/10.1016/S0140-6736\(17\)31332-6](https://doi.org/10.1016/S0140-6736(17)31332-6)
6. Lozada Hernández EE, Hernández Bonilla JP, Hinojosa Ugarte D et al (2023) Abdominal wound dehiscence and incisional hernia prevention in midline laparotomy: a systematic review and network meta-analysis. *Langenbecks Arch Surg* 7(1):268. <https://doi.org/10.1007/s00423-023-02954-w>
7. Veljkovic R, Protic M, Gluhovic A et al (2010) Prospective clinical trial of factors predicting the early development of incisional hernia after midline laparotomy. *J Am Coll Surg* 210(2):210–219
8. Basta MN, Mirzabeigi MN, Shubinets V et al (2016) Predicting incisional hernia after bariatric surgery: a risk stratification model based upon 2161 operations. *Surg Obes Relat Dis* 12(8):1466–147324
9. Fischer JP, Basta MN, Mirzabeigi MN et al (2016) A risk model and cost analysis of incisional hernia after elective abdominal surgery based on 12,373 cases. The case for targeted prophylactic intervention. *Ann Surg* 263(5):1010–100725
10. Lanni MA, Tecce MG, Hsu Y et al (2016) Development of a clinically actionable, longitudinal incisional hernia risk model after colectomy surgery using all-payer claims data. *J Am Coll Surg* 223(4):S29–S3026
11. Tecce MG, Basta MN, Shubinets V et al (2017) A risk model and cost analysis of post-operative incisional hernia following 2,145 open hysterectomies defining indications and opportunities for risk reduction. *Am J Surg* 213(6):1083–109027
12. Basta MN, Kozak GM, Broach RB et al (2019) Can we predict incisional hernia? Development of a surgery-specific decision support interface. *Ann Surg* 270(3):544–553
13. Goodenough CJ, Ko TC, Kao LS et al (2015) Development and validation of a risk stratification score for ventral incisional hernia after abdominal surgery: hernia expectation rates in intraabdominal surgery (the HERNIA project). *J Am Coll Surg* 220(4):405–413
14. Cherla DV, Moses ML, Mueck KM, Hannon C et al (2017) External validation of the HERNIA score: an observational study. *J Am Coll Surg* 225(3):428–434. <https://doi.org/10.1016/j.jamcollsurg.2017.05.010>
15. Tansawet A, Numthavaj P, Techapongsatorn S, McKay G et al (2021) Risk-benefit assessment of onlay and retrorectus mesh augmentation for incisional hernia prophylaxis: a secondary analysis from network meta-analysis. *Int J Surg* 92:106053. <https://doi.org/10.1016/j.ijssu.2021.106053>
16. Smeele NVR, Chorus CG, Schermer MHN, de Bekker-Grob EW (2023) Towards machine learning for moral choice analysis in health economics: a literature review and research agenda. *Soc Sci Med* 326:115910. <https://doi.org/10.1016/j.socscimed.2023.115910>
17. Lozada-Hernández EE, Ramírez-del Real TA, Salazar-Colores S, Armenta-Medina D (2025) Incisional hernia prediction using machine learning models. *BMC Med Inform Decis Mak*. Manuscript under review
18. Collins GS, Moons KGM, Dhiman P, Riley RD et al (2024) TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* 16:e078378. <https://doi.org/10.1136/bmj-2023-078378>
19. Bilimoria KY, Liu Y, Paruch JL, Zhou L et al (2013) Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. *J Am Coll Surg*. Nov;217(5):833–42.e1-3. doi: 10.1016/j.jamcollsurg.2013.07.385. Epub 2013 Sep 18. PMID: 24055383; PMCID: PMC3805776
20. ACS Risk Calculator - Home Page. American College of Surgeons (2018) Available from: <https://riskcalculator.facs.org> [accessed April 30, 2025]
21. IHXGBoost-P predictor (2025) Available from: <https://hi-eel-h-app-1035249958862.us-central1.run.app> [accessed July 30, 2025]
22. Muysoms FE, Antoniou SA, Bury K, Campanelli G et al (2015) European hernia society. European hernia society guidelines on the closure of abdominal wall incisions. *Hernia* Feb 19(1):1–24.

- <https://doi.org/10.1007/s10029-014-1342-5Epub> 2015 Jan 25. PMID: 25618025
23. Riley RD, Snell KI, Ensor J, Burke DL et al (2019) Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat Med* 38(7):1276–1296. <https://doi.org/10.1002/sim.7992>
 24. van der Hulst HC, Dekker JWT, Bastiaannet E, van der Bol JM et al (2022) Validation of the ACS NSQIP surgical risk calculator in older patients with colorectal cancer undergoing elective surgery. *J Geriatr Oncol* 13(6):788–795. <https://doi.org/10.1016/j.jgo.2022.04.004>
 25. Cole J, Hughey S, Metzger A, Geiger P et al (2021) Machine learning to predict fascial dehiscence after exploratory laparotomy surgery. *J Surg Res* 268:514–520. <https://doi.org/10.1016/j.jss.2021.06.068>
 26. Ju MH, Ko CY, Hall BL, Bosk CL et al (2015) A comparison of 2 surgical site infection monitoring systems. *JAMA Surg* 150(1):51–7. <https://doi.org/10.1001/jamasurg.2014.2891>
 27. Lozada-Hernández EE, Martín-Del-Campo LA, Fernández-Vázquez-Mellado LA, Valenzuela Alpuche HA et al (2025) External validation of the ACS NSQIP surgical risk calculator for the prediction of surgical site infection (SSI) and its association with the postoperative occurrence of incisional hernia (IH) in midline laparotomy patients. *Langenbecks Arch Surg* 410(1):293. <https://doi.org/10.1007/s00423-025-03807-4>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.