Scholars Journal of Physics, Mathematics and Statistics

OPEN ACCESS

Abbreviated Key Title: Sch J Phys Math Stat
ISSN 2393-8056 (Print) | ISSN 2393-8064 (Online)
Journal homepage: https://saspublishers.com

Development of a Multivariable Binary Logistic Regression Model to Predict Transudative Pleural Effusions Using Total Adenosine Deaminase Enzyme Activity and Demographic Variables in a Pleurology Unit at the Federal Fluminense University, Brazil

Cyro Teixeira da Silva Junior[1]*, Bernardo Henrique Ferraz Maranhão[2], Jorge Luiz Barillo[3], Patricia Siqueira Silva[4], Roberto Stirbulov[5], Evaldo Marchi[6]

[1]School of Medicine, Full Professor, Department of Clinics and Pleurology Teaching and Research Laboratory from Professor Mazzini Bueno, Fluminense Federal University, Niterói, State of Rio de Janeiro, Brazil. Certificate in Biostatistics from the Faculty of Medical Sciences of Santa Casa de São Paulo
[2]Federal University of the State of Rio de Janeiro, Department of Specialized Medicine, Brazil
[3]General Hospital Santa Teresa, Petrópolis, State of Rio de Janeiro, Brazil
[4]Professor Mazzini Bueno Tuberculosis Research and Assistance Center, Fluminense Federal University, Niterói, State of Rio de Janeiro, Brazil
[5]Faculty of Medical Sciences, Department of Clinics from Santa Casa de Misericórdia de São Paulo, State of São Paulo, Brazil
[6]Jundiaí School of Medicine, Department of Surgery, State of São Paulo, Brazil

DOI: https://doi.org/10.36347/sjpms.2026.v13i09.001 | **Received:** 21.07.2026 | **Accepted:** 07.09.2026 | **Published:** 10.09.2026

***Corresponding author:** Cyro Teixeira da Silva Junior
School of Medicine, Full Professor, Department of Clinics and Pleurology Teaching and Research Laboratory from Professor Mazzini Bueno, Fluminense Federal University, Niterói, State of Rio de Janeiro, Brazil. Certificate in Biostatistics from the Faculty of Medical Sciences of Santa Casa de São Paulo

Abstract **Original Research Article**

Total adenosine deaminase activity in pleural fluid (P-ADA) has been proposed as a biomarker for differentiating the causes of pleural effusion syndrome (PES). This retrospective study included 157 patients with PES, classified as transudative (n = 33; 21%) or exudative (n = 124; 79%). A model was created to predict transudative pleural effusions (TPE) using age, sex, and P-ADA activity as predictors. According to the Akaike information criterion, P-ADA was included as a continuous variable (U/L). A univariable logistic regression model was employed for all independent variables, but only those with a P-value of ≤0.25 were used for the Enter multivariable binary logistic regression. The variance inflation factor (VIF = 1) indicated no multicollinearity. The model was internally validated using 1,000 bootstrap resamples. In the final multivariable model, P-ADA activity (U/L) was the only independent predictor of TPE, whereas age and sex were not retained in the model. The Youden index on the ROC curve established a P-ADA < 9.0 U/L for pleural transudates. The aOR for P-ADA was 0.89 (95% CI: 0.82–0.96; P = 0.001). The regression equation was logit (p)= −1.106 −0.12183 × P-ADA (U/L), that is, increasing P-ADA (U/L) reduced the estimated probability of TPE. The discriminative performance was good, with an AUC of 0.82 (95% CI, 0.76–0.87; P < 0.05). Calibration was satisfactory according to the Hosmer–Lemeshow test. Bootstrap validation demonstrated stable model performance, with an optimism of approximately 2%. The overall classification accuracy was 82%. The model demonstrated substantial explanatory power or predictive performance, with a Nagelkerke pseudo-R^2 of 0.33. The clinical utility was considered high because the AUC exceeded 70%. The final model was not excessively overfitted, despite the relatively small number of TPE. In conclusion, P-ADA is a useful independent biomarker that may enhance diagnostic decision-making for pleural transudates.

Keywords: Adenosine deaminase; transudate; prediction model; logistic regression; calibration, discrimination; ROC curve.

Citation: Teixeira da Silva Junior, C., Ferraz Maranhão, B. H., Barillo, J. L., Silva, P. S., Stirbulov, R., & Marchi, E. (2026). Development of a Multivariable Binary Logistic Regression Model to Predict Transudative Pleural Effusions Using Total Adenosine Deaminase Enzyme Activity and Demographic Variables in A Pleurology Unit at the Federal Fluminense University, Brazil. *Sch J Phys Math Stat, 13*(9), 277-291. https://doi.org/10.36347/sjpms.2026.v13i09.001

1. INTRODUCTION

Pleural effusion syndrome (PES) is a frequent clinical and/or surgical condition characterized by abnormal fluid accumulation in the pleural space. Accurate cause identification remains essential in routine practice. The cornerstone of the diagnostic evaluation is the classification of pleural fluid as exudative or transudative. This distinction directly guides subsequent diagnostic strategies and therapeutic management [1–3].

Transudative pleural effusions typically result from systemic or mechanical disturbances that increase hydrostatic pressure, decrease plasma oncotic pressure, lower intrapleural pressure, promote the transfer of fluid from another body compartment, and cause fluid filtration to exceed lymphatic clearance, through the diaphragm. Common underlying conditions include heart failure, cirrhosis, nephrotic syndrome, and other systemic disorders. In these circumstances, extensive investigation of pleural fluid is generally unnecessary, and management should primarily focus on treating the underlying cause. Although transudative effusions are generally reversible, prognosis largely depends on the severity of the associated systemic disease. In contrast, exudative pleural effusions are commonly associated with local and/or systemic pathological processes, including infection, malignancy, pulmonary embolism, drug reactions, connective tissue diseases, and other conditions. These conditions often require targeted diagnostic and therapeutic interventions and are associated with substantial morbidity and mortality [1].

Total pleural adenosine deaminase activity (P-ADA, U/L; EC 3.5.4.4) is a well-established biomarker for differentiating exudative from transudative pleural effusions, particularly in the context of pleural tuberculosis [4, 5]. However, its role as a predictor of transudative pleural effusion has not been systematically evaluated, and no validated predictive models incorporating P-ADA are currently available. Therefore, this study aimed to develop and validate a logistic regression–based prediction model to estimate the probability of transudative pleural effusion using demographic variables and pleural ADA activity.

2. METHODS

2.1. Design

This retrospective cohort study was conducted between March 2015 and December 2019 at two hospitals, Antonio Pedro University Hospital and Santa Teresa Hospital, in the state of Rio de Janeiro, Brazil, in accordance with the TRIPOD and STROBE guidelines [6, 7]. The study was approved by the local ethics committee in accordance with the guidelines of the Helsinki Declaration (No. 48946121.9.0000.5243), and written informed consent was obtained from all participants.

2.2. Statistical Approach

Descriptive and inferential statistical analyses were performed using GraphPad software package (GraphPad Software, Inc., version 6.0, La Jolla, CA, USA) and MedCalc package for Windows (version 20.027; MedCalc Software, Ostend, Belgium). Data normality was assessed using the Shapiro–Wilk test (W), along with evaluation of variance homogeneity. Variables with non-normal distributions were summarized using the median and interquartile range (IQR), and group comparisons were performed using the Mann–Whitney U test. Categorical variables were expressed as absolute frequencies and proportions, and comparisons between proportions were conducted using the Chi-squared test. A two-tailed P-value of ≤ 0.05 was considered statistically significant and interpreted as evidence against the null hypothesis or H_0 [8].

2.3. Eligibility and Exclusion Criteria

The causes of pleural effusion syndrome were determined through clinical evaluation, imaging findings, serum and/or pleural fluid biomarkers, and, if necessary, surgical procedures. Thoracentesis with or without video-assisted thoracoscopic surgery was performed when indicated [1–3]. Pleural effusions were classified as transudates or exudates according to the Maranhão and Silva Junior criterion, which is based only on pleural fluid total protein and total lactate dehydrogenase levels and has been validated against Light's criteria [2, 3]. The applied cutoff values were 3.4 g/dL for total protein and 328.0 U/L for total pleural LDH [3]. Patients with contraindications to invasive procedures, immunosuppressive therapy, hemolyzed samples, renal failure, HIV infection, or pleural effusion of unknown cause were excluded [1–3].

2.4. Assessment of total adenosine deaminase (ADA), total lactate dehydrogenase (LDH), and total protein levels in pleural fluids

Total pleural adenosine deaminase (ADA) activity was measured in pleural fluid samples using a validated kinetic method with a commercial assay. Total pleural lactate dehydrogenase (LDH) activity was determined using an optimized kinetic method. Total protein concentrations were measured using the biuret colorimetric method [3, 9].

2.5. Framework with key assumptions and main strategies for multivariable binary logistic regression modeling in this study

These strategies are recommended in major methodological references and the TRIPOD Statement guidance for developing and reporting prediction models [6, 10, 11].

2.5.1. Adequate sample size with sufficient events for reliable coefficient estimation

In this study, there were three explanatory variables (age, sex, and P-ADA). Logistic regression sample size rules do not specify a fixed ratio of controls

to cases in the study. The sample size in this study consisted of 157 pleural fluid samples from 157 patients with proven transudative or cases (n = 33) or exudative pleural effusion or non-cases (n = 124). The ratio of cases to non-cases was 1:3.75. However, a balanced design (50% cases, 50% non-cases), usually provides efficient estimation of predictor effects. According to the Riley framework, it is often advantageous compared to highly imbalanced datasets and is commonly seen in diagnostic and prognostic research [10].

2.5.2. Definition of the research question: Identification of dichotomous dependent variable, predictor variables, Akaike information criterion, missing values, and potential outliers

The dependent variable was the diagnosis of transudative pleural effusion, coded as 1, whereas exudative pleural effusion served as the control outcome and was coded as 0. The candidate predictor variables selected a priori for model development included demographic characteristics of the study population (age and female sex) and total pleural adenosine deaminase (P-ADA) activity measured in all patients with PES. The recommendations of Altman and Royston are to retain in its original scale the continuous predictor variables in their original scale and not categorize them using ROC curve–based cutoffs to preserve information and statistical power [12]. However, the Akaike Information Criterion (AIC) was incorporated into the regression analysis to compare two alternative model specifications for P-ADA: modeling P-ADA as a continuous variable (U/L) and dichotomizing P-ADA to indicate pleural transudates based on a previously established cutoff [13]. The AIC provides an information-theoretic framework for comparing models fitted to the same dataset by jointly considering model fit and complexity. Lower AIC values indicate a more favorable balance between goodness of fit and parsimony, reflecting reduced information loss and a lower risk of overfitting [6]. In our that prior study, a pleural transudate was defined by a P-ADA value < 9.0 U/L, corresponding to an area under the ROC curve (AUC) of 81% [13]. For continuous candidate predictor variables, missing values were addressed by substituting the median or arithmetic mean according to the normality distribution [8]. To identify potential outliers, the interquartile range (IQR) method, also known as Tukey's method or the boxplot rule, was used as a nonparametric approach for continuous variables. [8]

2.5.3. Univariable logistic regression for candidate predictor selection based on the Hosmer–Lemeshow strategy

Predictor variables in the univariable analysis were initially evaluated for statistical significance using the Wald test. Consistent with the recommendations of Hosmer and Lemeshow, variables with a P-value < 0.25 in the univariable logistic regression analysis were considered for inclusion in the initial multivariable model [14]. The use of a more liberal screening threshold ($P < 0.25$), rather than the conventional $P < 0.05$, is justified because univariable analyses often have limited power to detect variables that may act as confounders. A confounder is defined as a variable that distorts the estimated relationship between an independent variable of interest and the outcome variable. The independent variables may not appear statistically significant in unadjusted analyses but can become important after multivariable adjustment. Therefore, adopting this relaxed criterion helps minimize false-negative exclusions and ensures that potentially relevant predictors are retained for subsequent multivariable modeling. A univariable logistic regression model was first fitted between each candidate predictor and transudative pleural effusion (outcome). Subsequently, a multivariable binary logistic regression model was developed using the Enter (forced-entry) method, in which all prespecified candidate predictors (age, sex, and P-ADA activity) that were significant in the univariable analysis were entered simultaneously in the multivariable model to estimate their independent associations with the outcome [14].

2.5.4. Multivariable modeling: Multicollinearity assessed using the variance inflation factor (VIF) to prevent highly correlated continuous predictors

Multicollinearity among the predictors in the multivariable model was evaluated using variance inflation factors (VIFs). Higher VIF values indicate an inflation of the variance of regression coefficients, which leads to reduced precision and reliability of the estimated effects. A VIF value of 1 indicates the absence of multicollinearity. In general, VIF values between 1 and 5 suggest moderate correlation that is typically acceptable, whereas values exceeding 5 may indicate multicollinearity. A VIF greater than 10 is commonly interpreted as evidence of severe multicollinearity requiring corrective measures. Multicollinearity reflects a high degree of correlation among potential predictor variables, typically indicated by strong linear relationships (e.g., correlation coefficients ≥ 0.9), which may compromise the stability and interpretability of the regression estimates. Singularity or perfect multicollinearity occurs when one predictor variable is an exact linear combination of other predictors included in the model [11].

2.5.5. Fit the multivariable logistic regression model

The multivariable model was fitted using the Enter method only after prespecifying the predictors. The multivariable logistic regression model was used to estimate adjusted odds ratios (AORs) and their corresponding 95% confidence intervals (CIs), allowing assessment of the independent association between each predictor variable and the outcome while controlling for potential confounders. In addition, a logistic regression equation was derived to model the probability of a transudative pleural effusion based on the significant independent predictor variables. Only variables with a P-

value < 0.05 were retained in the final prediction model [6, 14].

2.5.6. Traditional performance measures for a good model fit of the multivariable logistic regression model

The following approaches represent traditional performance measures used to evaluate the final logistic regression model, as recommended by several authors [6, 8, 11, 14].

1. Overall model fit and predictive performance: Nagelkerke pseudo-R^2

Overall model fit was assessed using the Nagelkerke pseudo-R^2. In the context of logistic regression, the coefficient of determination (pseudo-R^2) measures provide an estimate of the proportion of variation in the binary outcome that is explained by the model. Because logistic regression involves a dichotomous dependent variable and is based on maximum likelihood estimation rather than minimization of squared residuals, the conventional R^2 used in linear regression is not directly applicable. Therefore, the Nagelkerke pseudo-R^2 was employed to quantify the explanatory power of the model for transudative pleural effusion. In addition to evaluating overall model fit, the Nagelkerke pseudo-R^2 was also employed as a scalar indicator of predictive performance. This measure is both robust and interpretable, provided that the final model is appropriately specified [15].

2. Discriminative ability

Model discrimination was assessed using the concordance statistic (C-statistic), expressed as the area under the receiver operating characteristic (ROC) curve (AUC), along with corresponding 95% confidence intervals (CIs). The AUC quantifies the model's ability to correctly distinguish between transudative and exudative pleural effusions across all possible classification thresholds. Discriminative performance was interpreted according to the classification proposed by Hosmer and Lemeshow, whereby an AUC of 0.90–1.00 indicates excellent discrimination, 0.80–0.90 very good discrimination, 0.70–0.80 good discrimination, 0.60–0.70 sufficient discrimination, 0.50–0.60 poor discrimination, and values below 0.50 indicate a non-informative or clinically unhelpful biomarker [14].

3. Calibration plot with graphical illustration of the Hosmer-Lemeshow goodness-of-fit test

Calibration refers to the agreement between predicted probabilities generated by the model and the observed outcomes. Model calibration was assessed using the Hosmer–Lemeshow goodness-of-fit test (HL test) and visually examined through a calibration plot. In addition, a classification table was constructed to evaluate the predictive capacity of the model by cross-classifying observed and predicted outcome probabilities. HL test is based C statistic with ten groups. A large Chi-squared value associated with a P-value < 0.05 indicates poor model fit, whereas smaller Chi-squared values with P-values > 0.05 indicate good model fit and adequate calibration. Graphical assessment of calibration for transudative pleural effusions was performed by plotting observed outcome frequencies against predicted probabilities within each decile of risk [14].

4. Clinical utility

The clinical utility was assessed after evaluating model discrimination using the AUC of the ROC curve [16].

5. Validation of predictive model: Internal validation technique (bootstrap) and external validation

Internal validation evaluates how well a model developed using a specific dataset is expected to perform when applied to new, unseen data drawn from the same source population, thereby assessing potential overfitting and optimism. Common resampling-based approaches for internal validation include cross-validation and bootstrap methods; however, only the bootstrap approach was applied in the present study. The primary performance parameter evaluated during bootstrapping was the area under the receiver operating characteristic curve (AUC). Bootstrap resampling was used to estimate and correct for optimism, defined as the degree to which model performance is overestimated when assessed on the development dataset. The optimism-corrected AUC provides a more realistic estimate of discriminative performance after accounting for overfitting and was reported as the internally validated AUC. In the present study, external validation was not performed [11, 17, 18].

3. RESULTS

3.1. Case distribution, demographic characteristics, and laborartory analysis in pleural exudates and transudates

As shown in Table 1, exudative pleural effusions were more prevalent than transudative effusions, accounting for 124 patients (79%) compared with 33 patients (21%), respectively (Chi-squared test = 39.288, P<0.0001). Transudative pleural effusions were primarily attributed to congestive heart failure (n = 26), followed by liver cirrhosis with ascites (n = 3), chronic renal failure (n= 3), and isolated low total protein levels (n = 1). Exudative pleural effusions were mainly caused by tuberculosis (n = 44) and adenocarcinoma (n = 37), followed by uncomplicated parapneumonic effusion (n =15), complicated parapneumonic effusions or empyema (n = 8), lymphoma (n = 7), squamous cell carcinoma (n = 7), and other causes (n = 6).

Table 1 also summarizes the demographic characteristics and laboratory findings of the 157 patients with PES, comprising exudative and transudative effusions. Median P-ADA levels differed significantly between exudates and transudates (Mann–Whitney U = 735.5; P < 0.0001), with higher values observed in exudates (18.4 U/L; IQR 9.85–41.4) compared with transudates (6.85 U/L; IQR 2.67–11.26). No significant

differences were observed in the distribution of sex between groups, as assessed by the Chi-squared test, for either males (P = 0.9438) or females (P = 0.9416). In contrast, age differed significantly between groups, with higher median age among patients with transudative effusions (76.0 years; IQR 63.0–86.25) compared with those with exudative effusions (58.0 years; IQR 41.5–73.5) (U = 1068.0; P < 0.0001). Total pleural proteins and total pleural lactate dehydrogenase (P-LDH) levels in both groups were within the reference ranges established for biochemical diagnostic criteria [3].

Table 1: Demographic characteristics and laboratory findings in transudative and exudative samples from 157 patients

Parameters	Transudate	Exudate	Test (P-value)
Sample size, n	33	124	-
Prevalence (%)	21.0	79.0	$X^2 = 39.288$ (P<0.0001)
Age, years, median, interquartile range ($75^{th} – 25^{th}$ %)	76.0 (63.0 – 86.25)	58.0 (41.5 – 73.5)	U=1068.0 (P<0.0001)
Male, n (%)	16 (48.0)	58 (47.0)	X^2=0.005 (P=0.9438)
Female, n (%)	17 (52.0)	66 (53.0)	X^2=0.005 (P=0.9416)
Total P-ADA, median, interquartile range ($75^{th} – 25^{th}$ %), g/L	6.85 (2.67 – 11.26)	18.4 (9.25 – 41.4)	U=735.5 (P<0.0001)
Total P-ADA, cutoff by Youden index at the ROC curve (U/L)	<9.0	≥9.0	$X^2 = 29.5138$ (P<0.00001)
Total pleural proteins, mean ± standard deviation (g/L)	2.64 ±1.52	5.05 (4.47 – 5.60)*	t=2.498 (P=0.0186)
Total P-LDH, median, interquartile range ($75^{th} – 25^{th}$ %), U/L	190.5 (100.5 – 278.8)	568.5 (400.3 – 822.5)	U=306.0 (P<0.0001)

Abbreviations: P-ADA, total pleural adenosine deaminase; P-LDH, total pleural lactate dehydrogenase. Statistical notes: The Shapiro–Wilk test (W) rejected the normal data from P-ADA, age, and P-LDH in exudates (P < 0.05), but not total pleural proteins in transudates (P = 0.0881); *Logarithmic transformation of data (W = 0.9582; P = 0.2972)

3.2. Variable coding, missing data, outlier detection with replacement, and Akaike information criterion

Table 2 provides an overview of the variable coding, missing data patterns, and outlier detection, including imputation procedures. For continuous candidate predictor variables, missing values were addressed by substituting the median. Unlike missing values, outliers identified by Tukey's method (IQR rule) should not be automatically imputed or replaced**.** The IQR rule is primarily a screening tool for identifying observations that warrant investigation, not an automatic criterion for replacement or removal [**8**]. For our clinical prediction model, the most defensible approach was to not retain biologically plausible outliers. We investigated their influence using diagnostic measures and only imputed values that were confirmed to be correct. There were no values suspected to be erroneous by digitation. Only one outlier value of the total P-ADA (U/L) was substituted with the median to reduce undue influence on model estimates and improve model stability.

Table 2 also presents the Akaike Information Criterion (AIC) values used to compare two competing logistic regression model specifications: modeling pleural adenosine deaminase (P-ADA) as a continuous variable (U/L) and dichotomizing P-ADA using a cutoff value of < 9.0 U/L to indicate transudative pleural effusions, as defined in **Table 1**. The observed pattern of missing data was consistent with a missing-at-random (MAR) mechanism.

Table 2: Variable coding, missing data handling, outlier detection, median imputation, and Akaike information criterion

Predictor variable	Coding	Missing data			Outlier detection (median imputation)		AIC
		E (%)	T (%)	Median replacement (E/T)	E	T	
Female sex	1	0/124 (0.0)	0/33 (0.0)	-	-	-	165.46
P-ADA	U/L	0/124 (0.0)	0/33 (0.0)	-	1121,1 (18.4)	-	127.40
P-ADA	1	0/124 (0.0)	0/33 (0.0)	-	-	-	163.49
Age	Years	4/124 (3.0)	2/33 (6.0)	76.0/58.0	-	-	149.14

Abbreviations: P-ADA, total pleural adenosine deaminase enzyme; E, exudate; T, transudate; Statistical notes: AIC, Akaike information criterion provides an alternative to various forms of P-ADA by evaluating both continuous and categorical variables. Lower AIC values suggest a better-adjusted model fit.

3.3. Univariable analysis

Table 3 presents the results of the univariable analysis of each potential variable for the final multivariable modeling among 157 patients. Only the coefficients for P-ADA (U/L) and age were significant in the Wald test (P < 0.25).

Table 3: Univariable analysis of each potential variable with crude odds ratios (unadjusted) for pleural transudate association and exploration of possible predictors for the final multivariable modeling among 157 patients

Potential predictor variable of pleural transudate	Beta-coefficient	Standard error	Wald test (Z)	uOR	P-value
Female	−0.20723	0.39245	0.2788	0.8128	0.5975
P-ADA (U/L)	0.13382	0.037276	12.8888	0.8747	0.0003
Age (years)	−0.039485	0.012262	10.3691	1.0403	0.0013

Abbreviation: P-ADA, total pleural adenosine deaminase; uOR, unadjusted Odds Ratio; Statistical note: The selection criterion was P < 0.25.

3.4. Multivariable analysis of the logistic model

The results in **Table 4** show that only P-ADA was a significant independent predictor of pleural transudate in the proposed model. The VIF indicated no multicollinearity because all values were below the 5–10 threshold.

Table 4: Final model development using Enter multivariable logistic regression according to univariable regression analysis

Independent predictor variable	Beta-coefficient	Standard error	aOR (95% CI)	P-value	VIF ($1/1-R^2$)
Constant or y-intercept (α)	−1.10600	1.10240	-	-	-
P-ADA - U/L (β_1)	−0.12183	0.036938	0.8853 (0.82 – 0.96)	0.0010	1.03
Age - years (β_2)	0.021452	0.014499	1.0217 (0.99 – 1.05)	0.1390	1.02

Abbreviations: 95% CI, 95% confidence interval; aOR, adjusted odds ratio; P-ADA, total pleural adenosine deaminase; AUC, area under the ROC curve. VIF, Variance Inflation Factor; R^2, pseudo-coefficient of determination. Statistical notes: Nagelkerke pseudo-R2 = 0.3332. Discrimination with AUC = 0.820 (95% CI, 0.0871 – 0.757; P < 0.05; SE = 0.0392). The logistic regression equation was logit (p)= −1.106 −0.12183 × P-ADA. The selection criterion was P < 0.05. VIF > 5-10 indicates multicollinearity.

3.5. Clinical utility

The model demonstrated good discriminative performance, with an area under the ROC curve (AUC), not shown graphically, greater than 0.75 (AUC=82%). This value supports its potential clinical applicability [**16**].

3.6. Correlation between transudate classification and pleural adenosine deaminase levels

Figure 1 illustrates the inverse relationship between P-ADA levels and the probability of transudative pleural effusion, demonstrating that the likelihood of a transudate decreases progressively as P-ADA values increase. A marked decline in the predicted probability of transudate is observed around P-ADA values of approximately 60 U/L, indicating a sharp reduction in transudate probability beyond this range [**19**].

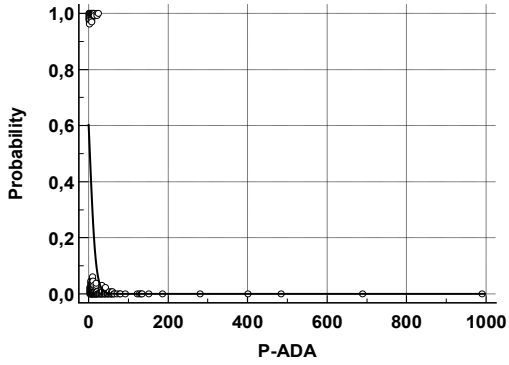

Figure 1: Fitted line plot showing the relationship between transudate classification (response variable) and the significant independent predictor, pleural adenosine deaminase (P-ADA, U/L), in 157 patients with pleural effusion syndrome. The logistic regression equation (logit (p)= −1.106 - 0.12183) × P-ADA expresses the probability of a transudate, where a value of 1 on the y-axis represents a transudate (success). The plot indicates that the probability of a transudate decreases as P-ADA values increase. Around P-ADA levels of 60%, the line declines sharply, suggesting a sudden drop in the probability of a transudate with increasing P-ADA values.

3.7. Calibration and classification table

Calibration, defined as the agreement between predicted and observed risks, is considered a more informative indicator of the clinical utility of a multivariable logistic regression model than discrimination alone. This principle also applies to the evaluation of decision thresholds and net clinical benefit, as demonstrated by decision curve analysis [**11, 14**]. **Figure 2** presents the scatter plot derived from the Hosmer–Lemeshow goodness-of-fit test, based on ten risk deciles, used to assess the calibration of the logistic regression model. The classification table, not shown here, indicated that 82% of pleural transudative effusions were correctly classified [**14**].

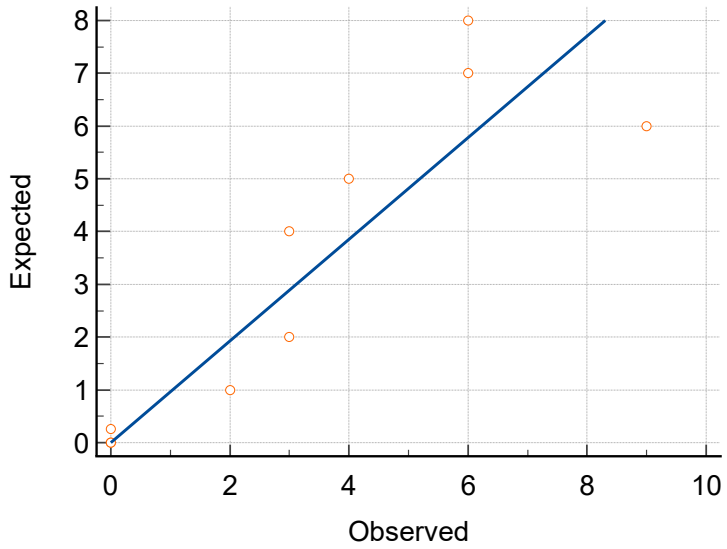

Figure 2: Scatter plot from Hosmer–Lemeshow test for assessing the calibration of the logistic regression model according to the contingency table. Statistical notes: The plot shows the details of the P-ADA with observed and expected number of pleural transudate cases. The data were divided into ten groups. The observed and expected number of cases in each group was calculated using the Chi-squared test, which was equal to 5.5774 (P = 0.6975). A P > 0.05 was considered evidence of good calibration. The graphical presentation shows a significant correlation between the observed and expected probabilities of transudative pleural effusion. The Pearson's correlation coefficient (r) was equal to 0.8916 (P = 0.0005; 95% IC for r = 0.59 – 0.97).

3.8. Bootstrapping for internal validation

Table 5 presents the internal validation performed to assess potential overfitting and reduce bias using bootstrap resampling with 1,000 iterations. This procedure was used to estimate the 95% confidence interval of the area under ROC curve (AUC) for classifying transudative pleural effusions based on total pleural adenosine deaminase activity [**17, 18**].

Table 5: Internal validation using 1,000-resample bootstrapping of the area under the ROC curve to assess potential overfitting in the logistic regression model classifying pleural transudates based on total pleural ADA activity in 157 patients

AUC development		**Internal validation using AUC**	**Optimism correction from AUC to interpret overfitting**
95% CI	Apparent observed	Bootstrap expected	n (%)
Minimum	0.751	0.736	0.015 (+1.5)
Maximum	0.877	0.883	-0.006 (-0.6)

Abbreviations: 95% CI, 95% confidence interval; AUC, area under the ROC curve. Statistical notes: A tolerance level of 2% between observed and expected values of AUC with 1,000 replications was deemed appropriate to evaluate optimism correctly and to interpret overfitting. Rules of thumb: < 0.02: minimal overfitting; 0.02–0.05: moderate; and 0.05: strong overfitting (model likely too flexible).

4. DISCUSSION

In accordance with the TRIPOD guidelines and methodological recommendations from several expert sources, a systematic and structured approach was adopted for the development and interpretation of the multivariable binary logistic regression model used in this study [6, 8, 11, 14].

4.1. Multivariable binary logistic regression: Strategies for logistic regression and keys assumptions

Multivariable binary logistic regression is a statistical method used to predict a yes/no outcome using two or more independent variables, estimate adjusted

odds ratios, and control for confounding factors [6, 8, 11, 14].

Logistic regression is an appropriate analytical approach when the outcome of interest is binary. The term "logistic" originates from the use of the logit function, which models the logarithm of the odds of the outcome as a function of the predictor variables. Diagnostic prediction models play a central role in clinical decision-making by estimating disease probabilities and integrating multiple predictors into a single quantitative framework. These models aim to enhance diagnostic accuracy, standardize clinical evaluation, optimize resource utilization, and support more personalized patient care. By promoting transparency, reproducibility, and methodological rigor, prediction models strengthen evidence-based practice and facilitate earlier and more precise diagnoses. Their reliability, however, depends on robust internal and external validation, and their predictive estimates must be accurate to appropriately inform patient management and therapeutic decision-making [6, 11].

4.2. Practical considerations and contextual relevance

Understanding the clinical context of this study is crucial. Pleural transudates are generally less prevalent than pleural exudates in many hospital-based PES series because of a combination of epidemiologic and clinical factors, such as, many transudative conditions are treated before a pleural effusion develops; transudates are often clinically obvious and may not undergo thoracentesis; pleural exudative diseases are common indications for diagnostic thoracentesis, and there is a referral bias in tertiary centers because specialized pulmonary and pleural clinics receive more patients with complex exudative effusions. This phenomenon increases the observed prevalence of exudates; population characteristics matter because in older populations or cardiology settings, transudates may constitute 20–50% of pleural effusions, and in pulmonary referral centers and studies evaluating biomarkers (such as P-ADA), exudates commonly represent 70–90% of the cases [**1, 3, 13**].

4.3. Clinical relevance of P-ADA activity

The relative contributions of adenosine deaminase isoenzymes ADA1 and ADA2 to total ADA activity vary according to the biological fluid and the underlying pathological condition. Although no fixed proportion applies universally, well-established ranges have been described in the literature. ADA1 typically accounts for approximately 20%–40% of total ADA activity in serum and transudative pleural effusions and is predominantly located intracellularly in lymphocytes, neutrophils, and erythrocytes. In contrast, in tuberculous pleural effusions, ADA2 is the predominant isoenzyme, contributing approximately 70%–90% of total ADA activity, likely reflecting increased monocyte–macrophage activation and turnover [**13**]. The activity of ADA isoenzyme 2 was not specifically evaluated in this study.

4.4. Sample-size justification based on the Riley framework

Calculating the sample size for multivariable models is a complex issue. Our sample size was adequate for the validation of this study according to the framework proposed by Riley *et al.,*for prediction models [10]. The framework proposed by Riley *et al.,*is a modern methodology for determining the minimum sample size required to develop a multivariable clinical prediction model. Instead of relying on simple rules such as "10 events per variable (EPV)," it uses statistical criteria to ensure that the model is likely to perform reliably in new patients. This study included 33 events and 124 non-events for the evaluation of three candidate predictors, providing approximately 11 events per predictor and supporting the model development with acceptable precision and limited overfitting. Thus, the sample size was likely adequate for fitting the model, although the number of events remained relatively modest. In a widely cited paper, Peduzzi *et al.,*[20] concluded that for logistic regression or the logit function, a minimum of ten events (not subjects) is required for each predictor variable (referred to as events per variable, or EPV). This guideline is often recommended as a rule of thumb. A study published by Ogundimu *et al.,*[21] based on simulation studies concluded that for logistic regression, a "higher EPV is needed when low-prevalence predictors are present in a model to eliminate bias in regression coefficients and improve predictive accuracy. In such cases, up to 20 EPVs may be required." Neither the Riley framework nor the TRIPOD guidelines require a minimum ratio of non-cases to cases, such as 1:1, 2:1, or 3:1 [6, 10].

The focus is on having an adequate number of events and non-events. A 1:1 ratio is acceptable if the total sample size satisfies the Riley criteria and provides sufficient events and non-events for stable model estimation. If the modern framework of Riley *et al.,*is used for sample size calculation of prediction models, a 1:1 event-to-non-event ratio is entirely acceptable; the adequacy depends on the resulting numbers of events and non-events rather than on the ratio alone. There is no rationale for the 1 variable per 10 events criterion for binary logistic regression analysis [22].

With 157 pleural effusions, including 33 transudative and 124 exudative effusions, and three candidate predictors, the study achieved approximately 11 events per variable. Assuming a moderate-to-large effect size and a two-sided significance level of 0.05, the available sample size provided an estimated statistical power of approximately 80–90% to detect significant associations in the logistic regression model. However, an effect size of 80% is not a standard parameter for logistic regression power analysis. A post-hoc assessment indicated that the available sample provided

adequate statistical power for the planned logistic regression analysis. Assuming a significance level of 0.05 and an anticipated model performance of Nagelkerke R2=0.33 (AUC > 0.70), the estimated power exceeded 80%, supporting the detection of clinically relevant predictor effects. However, for a prediction-model paper, reviewers from journals such as those affiliated with TRIPOD guidelines often prefer a sample-size justification based on the Riley framework rather than a post hoc power analysis. Many methodological statisticians discourage post hoc power calculations because they add little information beyond the observed effect estimates and confidence intervals [6, 10]. In conclusion, the Riley framework is a stepwise method for calculating the minimum sample size for a new prediction model. It rejects the old "rules of thumb" (such as having 10 events per variable). Instead, it uses specific mathematical formulas based on the outcome type. To build a reliable model, Riley suggested that the sample size should achieve three specific goals: keep model optimism low, limit R^2 shrinkage, and estimate the overall risk precisely. How is the calculation? For binary outcomes (transudates, exudates), the estimated sample size was based on the expected event fraction (the rate at which the "yes" outcome occurs) and the expected Nagelkerke's R^2 [6, 15].

4.5. Univariable logistic regression analysis using the Hosmer–Lemeshow strategy

There is no universally accepted optimal method for constructing a logistic regression model; the appropriate strategy depends on the study objectives, sample size, theoretical framework, and risk of overfitting. In this study, an exploratory binary univariable logistic regression analysis was initially performed to examine crude associations between transudative pleural effusion and candidate predictors. This analysis served as an initial preparatory phase for multivariable modeling, facilitating data exploration, assessment of crude associations, and identification of potentially relevant predictors and confounding effects [6, 14]. Although univariable screening using liberal significance thresholds (e.g., P<0.10, P<0.20, or P<0.25), as historically recommended by Hosmer and Lemeshow, has been widely used to reduce the number of candidate variables and limit overfitting, statistical significance in univariable analysis is now generally considered insufficient as the sole criterion for predictor selection [11]. Predictors may be important confounders, clinically established risk factors, or improve model performance despite having a nonsignificant univariable association; conversely, variables significant in univariable analysis may lose significance after adjustment [6, 11, 14]. Therefore, contemporary modeling strategies generally favor a priori selection based on clinical relevance, biological plausibility, previous evidence, causal reasoning, and subject-matter expertise. Nevertheless, univariable analysis remains useful for describing crude associations, exploring the data, assessing variable coding, and identifying potential nonlinearities and interactions [11]. In the present study, univariable analysis was consequently considered an exploratory and preparatory component of model development rather than the sole basis for determining predictor importance, with clinically relevant candidate variables subsequently considered for multivariable logistic analysis using the Enter method.

Considering the above, we performed a binary univariable analysis in accordance with the Hosmer and Lemeshow strategy [14]. Importantly, our study had only three clinically prespecified candidate predictors: age, sex, and P-ADA (U/L). The strongest justification for this approach is that these predictors were specified a priori based on clinical and biological plausibility, and the study objective. Univariable analysis was used primarily as an initial screening and descriptive step rather than as the sole scientific basis for determining the importance of predictors. The final multivariable model with candidate predictors meeting the prespecified univariable criterion entered simultaneously rather than through an automated variable-selection procedure. Furthermore, with only three prespecified candidate predictors and 33 transudative events, the model was relatively parsimonious, corresponding to approximately 11 events per predictor.

4.6. Multivariable analysis

In multivariable analysis, one predictor can substantially change the apparent importance, effect size, statistical significance, or even the direction of the association of another predictor. It estimates the effect of each predictor, representing its independent association with the outcome after adjusting for the remaining predictors. Consequently, confounding, collinearity, and interactions between predictors may alter the magnitude, direction, and statistical significance of the individual regression coefficients. This is a fundamental reason why multivariable models are used [11, 14].

In this study, a multivariable binary logistic regression model was used to estimate the probability of transudative pleural effusion and to assess the independent association of each predictor with the odds of this outcome while accounting for the other predictors in the model. The final model was developed using P-ADA (U/L) and age (years), which were identified as significant predictors in the univariable analysis (Table 3). It was entered simultaneously into the multivariable model [6,11,14].

4.7. Principal findings and comparison with previous studies

According to Table 1, the distribution of sex and median age in the study population was comparable to that reported in previous investigations by our study group [3, 13, 23]. Several studies have demonstrated that P-ADA activity decreases with age, whereas sex has not consistently been shown to independently influence P-ADA levels. Consequently, age-adjusted interpretation

of ADA values may improve diagnostic accuracy, particularly in older patients, whereas sex-specific reference values are not currently recommended [24]. P-ADA demonstrated high accuracy in distinguishing transudative from exudative pleural effusions, as evidenced by a significant difference between the groups, as shown in Table 1 (Mann–Whitney U = 735.5; $P < 0.0001$). These findings are consistent with prior studies that have also reported the diagnostic utility of P-ADA in pleural effusion classification [3, 13, 23, 25]. In addition, total pleural lactate dehydrogenase (P-LDH) and total pleural protein concentrations remain classical and well-established biomarkers for differentiating transudative from exudative pleural effusions [2, 3, 5]. The statistically significant differences observed in the present study further support the diagnostic relevance of these biomarkers (Table 1; $P < 0.05$).

In the present study, P-ADA was analyzed both as a continuous and a categorical variable (Table 1). Based on the AIC, P-ADA (U/L) was modeled as a continuous variable in the univariable logistic regression analysis in order to preserve information, increase statistical power, estimate the effect of each unit increase, and assess potential dose–response relationships. As shown in Table 2, the lower AIC value indicated a superior model, reflecting an optimal balance between goodness of fit and model parsimony. Accordingly, models with lower AIC values are preferred because they provide a more favorable trade-off between accuracy and complexity. The Akaike Information Criterion (AIC), introduced by the Japanese statistician Hirotugu Akaike in 1973, is a widely used statistical measure for model comparison. It is frequently applied to evaluate alternative sets of predictors, including continuous versus categorical representations, as well as different transformations or interaction terms [26].

As shown in **Table 2**, median imputation was required for missing values in only 3% of exudative cases and 6% of transudative cases. This approach is considered acceptable, given that the proportion of missing data remained low and well within the commonly cited threshold of 5–10%. Missing data should not be disregarded, as inappropriate handling may introduce bias and affect model performance. Previous empirical and simulation studies indicate that when fewer than 10% of observations are missing under missing completely at random (MCAR) or missing at random (MAR) assumptions, the impact on parameter estimates and predictive performance is generally limited [27].

In the present study, an extreme value of pleural adenosine deaminase (P-ADA) activity (1121.1 U/L) was identified using the Tukey's method in an exudative pleural effusion sample from a patient with lymphoma. This value was verified as a true observation, with no evidence of data entry error, and was considered compatible with the underlying causal diagnosis. Nevertheless, to optimize internal calibration of the logistic regression model, this extreme value was replaced with the median P-ADA value (18.4 U/L), as reported in Table 2. This approach did not materially influence the estimated regression coefficients. Outliers were systematically evaluated to minimize potential bias, while arbitrary cutoff points were avoided in order to preserve underlying data trends and statistical precision. In addition, careful selection of covariates was undertaken to address potential confounding, thereby enhancing the robustness and reliability of the observed associations. This rigorous analytical strategy was intended to provide clearer insights into the relationships underlying the data [8, 11]. In statistical analysis, an outlier is defined as an observation that differs markedly from the remaining data points in a dataset, appearing unusually large or small relative to the overall distribution. The Interquartile Range (IQR) method, also known as Tukey's method or the boxplot rule, is a nonparametric approach for identifying potential outliers. If an outlier is detected, removed, or substituted, the test may be repeated iteratively, although repeated testing increases the risk of false positive findings [8].

Table 3 included 33 outcome events and three predictor variables. To enhance model robustness, non-significant predictors were excluded during the univariable analysis (Table 4), resulting in a parsimonious final model. Model performance was subsequently evaluated rigorously, with particular attention to overfitting and calibration (Table 5). Firth's penalized likelihood approach was not applied because this method is primarily indicated to address small-sample bias, which was not evident in the present analysis [28].

According to the Wald test, only the beta coefficients of P-ADA and age were identified as potential predictor variables in the univariable analysis (Table 3). In logistic regression, the beta coefficient represents the change in the logarithm of the odds of the outcome associated with a one-unit increase in the explanatory variable. Understanding the relationship between beta coefficients and odds ratios is essential for correct interpretation of model results. Although closely related, beta coefficients and odds ratios are not equivalent. Beta coefficients are primarily used for model estimation and statistical inference, whereas odds ratios—obtained by exponentiating the beta coefficients—offer a more intuitive and clinically interpretable measure of association and are therefore more commonly reported in the medical and epidemiological literature [8, 11].

As shown in Table 4, P-ADA was the only independent predictor of transudative pleural effusion in the final model, with an adjusted odds ratio (aOR) of 0.8853 (95% CI, 0.82–0.96). However, an odds ratio alone cannot be directly translated into an absolute

probability unless a baseline (reference) probability or odds is specified, a common source of misunderstanding in applied medical statistics. Logistic regression provides an appropriate analytical framework for probability estimation by modeling the log-odds of the outcome as a function of the predictor variables. In the present study, the fitted model was expressed as logit(p) = −1.10600 − 0.12183 × P-ADA (Table 4). For illustrative purposes, a P-ADA value of 20.0 U/L yields a logit(p) of −3.5426, corresponding to an estimated probability of approximately 2.8% after transformation from the logit scale. It should also be noted that the cutoff value for P-ADA used to classify pleural transudates in this study was < 9.0 U/L, as defined by the Youden index at the ROC curve in Table 1 [11, 13].

In addition, Table 4 shows that an aOR of 0.8853 for transudative pleural effusion indicates that pleural adenosine deaminase (P-ADA, U/L) is associated with a reduced likelihood of the outcome. An aOR below 1 denotes a decrease in the odds of classification as a transudate. Specifically, an AOR of 0.8853 implies that each one-unit increase in P-ADA activity is associated with an approximately 11.5% reduction in the odds of a pleural effusion being classified as transudative (1 − 0.8853 = 0.1147). In practical terms, higher P-ADA concentrations are therefore more frequently observed in exudative pleural effusions, reinforcing the inverse association between P-ADA levels and the probability of a transudative diagnosis (Figure 1).

Adjusted odds ratios provide a useful measure of the magnitude of association between predictors and outcomes in a given sample, while 95% confidence intervals (CIs) quantify the precision of these estimates. Wider confidence intervals indicate lower precision, whereas narrower intervals reflect greater precision. In the present study, the 95% CI for P-ADA in Table 4 was relatively narrow (0.82–0.96), indicating a precise estimate of effect. Statistical significance was assessed by examining whether the confidence interval included the null value of 1. Because the CI did not include 1, the association was considered statistically significant. This interpretation was further supported by the corresponding P-value for P-ADA (P = 0.0010), which was well below the conventional significance threshold of 0.05, confirming a statistically significant association [8, 11].

4.8. Diagnostic performance of the model

In our study, the performance of the prediction model was evaluated using multiple methods and metrics appropriate for binary outcomes. Overall model accuracy was assessed using a classification table (confusion matrix), which compares observed outcomes (0 or 1) with predicted classifications based on a predefined probability threshold. Pseudo-R^2 measure was used to provide a general indication of model fit and explanatory power. Discriminative ability was evaluated using the concordance (C) statistic, expressed as the AUC. Model calibration was assessed using the Hosmer–Lemeshow goodness-of-fit test. The Brier score, which jointly summarizes calibration and discrimination without reliance on arbitrary classification thresholds, was not evaluated in the present analysis [8, 11, 14].

With respect to predictive performance (Table 4), the Nagelkerke pseudo-R^2 value of 0.3332 indicated that the model explained approximately 33% of the variability in the diagnosis of transudative pleural effusion. According to Cohen's criteria for effect size interpretation, pseudo-R^2 values of 0.26 are considered substantial, 0.13 moderate, and 0.02 weak [15, 29].

A traditional goodness-of-fit test used to assess the calibration of logistic regression models is the Hosmer–Lemeshow test [14]. In the present study, the Hosmer–Lemeshow Chi-squared statistic was 5.5774 (P = 0.6975), indicating adequate model fit, as P-values greater than 0.05 suggest no evidence of lack of fit. In addition, the graphical assessment shown in Figure 2 demonstrated a strong correlation between predicted and observed probabilities of transudative pleural effusion, with a Pearson correlation coefficient of r = 0.8916 (P = 0.0005; 95% CI, 0.59–0.97). These findings indicate good agreement between predicted probabilities and observed outcomes, supporting satisfactory model calibration [14].

Overall classification accuracy, as assessed by the classification table (confusion matrix), indicated that 82% of pleural transudative effusions were correctly classified by the logistic regression model. The classification table summarizes model performance by comparing observed outcomes with predicted categories (e.g., transudative vs. exudative pleural effusion) based on a predefined probability threshold. Overall accuracy was calculated as the proportion of correctly classified cases divided by the total number of cases, multiplied by 100. Accordingly, an accuracy of 82% indicates that the model correctly classified 82% of all evaluated cases [14].

The model demonstrated very good discriminative performance, with an area under the AUC of 0.820 (95% CI, 0.757–0.887; P < 0.05), according to the classification proposed by Hosmer and Lemeshow [14]. In the present study, several advanced metrics for evaluating discrimination were not assessed, including modified AUC measures for survival data, reclassification tables, net reclassification improvement (NRI), integrated discrimination improvement (IDI), and decision-analytic approaches such as decision curve analysis, which are designed to quantify the net clinical benefit of model-based predictions in decision-making contexts [30].

Clinical utility refers to the practical value and impact of a method, test, or prediction model on clinical decision-making and patient outcomes, beyond mere

statistical performance. It is typically evaluated after assessment of model discrimination. In the present study, clinical utility was assessed based on the concordance statistic (C-statistic or AUC) and its corresponding 95% confidence interval. Although a high AUC, generally exceeding 0.70, is considered a necessary condition for clinical utility [16]. Other approaches for evaluating clinical utility—such as decision curve analysis (DCA), net benefit assessment, clinical impact curves, and comparisons with standard care or existing decision rules—were not performed in this study [6, 11, 30].

Table 5 presents the internal validation of the logistic regression model using bootstrap resampling with 1,000 iterations to assess potential overfitting. In the sample of 157 patients, the observed AUC for identifying transudative pleural effusions ranged from 0.751 (minimum) to 0.877 (maximum). After bootstrap resampling, the expected AUC values ranged from 0.736 (minimum) to 0.883 (maximum).

Table 5 also reports the optimism correction applied to the AUC to facilitate interpretation of potential overfitting. The observed optimism correction from the AUC values (0.015 and −0.006) was minimal, indicating negligible deviation between the apparent and bootstrap-corrected performance estimates. Consequently, the optimism-corrected AUC values were effectively 0.736 (minimum) and 0.883 (maximum), suggesting stable model discrimination with minimal overfitting [17, 18].

In ROC curve analysis, bootstrap resampling is widely used to estimate confidence intervals and to assess the stability of model performance. Optimism correction is applied to adjust the apparent performance of a logistic regression model for the upward bias that arises when model evaluation is conducted on the same dataset used for model development. Because predictive models tend to overfit the development sample, performance metrics such as the AUC, calibration slope, and Brier score may appear overly optimistic. Internal validation—most commonly performed through bootstrap resampling—quantifies this optimism by repeatedly refitting the model in bootstrap samples and comparing its performance in each resample with that observed in the original dataset. The average difference between these performances represents the degree of optimism. Subtracting this value from the apparent performance yields an optimism-corrected estimate, which more accurately reflects the model's expected performance in new data. In this way, optimism correction provides a more realistic assessment of predictive performance by accounting for overfitting inherent to model development [17, 18, 31].

4.9. Strengths, limitations, and future research

Despite the observational design and the interpretation of logit functions, odds ratios, and probabilities derived from logistic regression modeling, this study has limitations. External validation assesses the generalizability and transportability of a prediction model, specifically its ability to predict outcomes in populations or settings different from those used for model development (e.g., different institutions, time periods, or geographic regions). Internal validation using bootstrap resampling does not replace external validation. However, it provides an estimate of how the model is expected to perform when applied to new, unseen data drawn from the same source population. For this reason, internal validation is sometimes loosely associated with the concept of "generalizability." This misconception arises because optimism-corrected performance metrics are intended to approximate how the model would behave in future patients who are similar to those included in the development dataset. Nevertheless, this approach remains fundamentally distinct from true external validation, which evaluates model performance in entirely new settings, populations, or time periods. Conducting studies that include external validation is challenging for many researchers, particularly those without substantial funding or institutional support, as external validation is resource-intensive. It requires access to independent datasets, adequate infrastructure for data sharing and management, interinstitutional collaboration, time to obtain approvals and collect data, specialized software for advanced statistical analyses, and sufficient financial resources to support these activities. Riley *et al.,*proposed a framework for external validation of binary outcome prediction models, emphasizing the number of participants and outcome events needed for precise calibration, discrimination, and clinical utility. For example, to estimate an AUC of 0.80 with a confidence interval width of 0.10, approximately 1,170 participants would be required. Calibration is often the main factor in determining sample size requirements for validation studies, although external validation is crucial before clinical implementation [**32].**

In his book *Problem Solving: A Statistician's Guide (1988)*, Chris Chatfield observes that students often possess knowledge of the technical aspects of regression but lack understanding of its appropriate application. This highlights the necessity for a more balanced approach in both the literature and statistical education, emphasizing techniques alongside problem-solving strategies. Accordingly, it is important to emphasize that Su *et al.,*(2026) concluded that traditional test-based approaches, including univariable predictor selection, are generally discouraged in clinical prediction modeling. However, their study was a systematic review of practices actually used by researchers rather than a guideline establishing that univariable screening is never permissible [**33**]. Another highly relevant systematic review published in *The Lancet Healthy Longevity* in 2026 identified univariable predictor selection as one of the major sources of risk of bias across 250 mortality prediction models in older adults with cancer [**34**]. However, our study differs from the models included in that review in several important respects. Specifically,

key demographic variables and a diagnostic biomarker were prespecified as predictors, and all predictors that were significant in the univariable analysis were entered simultaneously into the final multivariable logistic regression model using the Enter method.

4.10. Considerations regarding the term Pleurology in the title

With regard to the title of this study, it is important to acknowledge that the term "pleurology" is not a standard or widely accepted designation in contemporary medical literature. Etymologically, the prefix "pleuro" refers to the pleura—the serous membrane surrounding the lungs—while the suffix "logy" denotes the study of a given subject. In a literal sense, "pleurology" would therefore signify the study of the pleura or a discipline dedicated to pleural diseases. Despite this linguistic construction, "pleurology" is not formally recognized as an official medical specialty.

A relevant perspective, however, is offered by Professor Richard W. Light (1942–2021), widely regarded as a pioneer and the "father" of modern pleural medicine. Light, best known for developing the diagnostic framework now referred to as Light's criteria, emphasized in the preface to a seminal work on pleural effusions that pleural diseases do not belong to a single medical specialty. Instead, they are encountered across multiple disciplines—including pulmonology, critical care, infectious diseases, oncology, rheumatology, cardiology, and gastroenterology—resulting in diagnostic and management approaches that are frequently suboptimal [**35**].

In light of Professor Light's observations, the fact that pleural diseases do not belong to a single medical specialty and are often managed in a fragmented manner raises important conceptual considerations. The lack of a clearly defined disciplinary focus may contribute to suboptimal diagnosis and management when pleural conditions are addressed exclusively by non-specialists. These considerations have led some authors to reflect on whether pleural diseases might benefit from a more clearly defined conceptual identity, analogous to the evolution of other organ-based disciplines. Although the term "pleurology" is not formally recognized as an official medical specialty and is not widely adopted in contemporary literature, it has occasionally been proposed as a theoretical designation to emphasize the growing complexity of pleural disorders and the potential need for dedicated expertise in this field.

5. CONCLUSION

Total pleural adenosine deaminase activity demonstrated strong independent discriminative ability, providing valid, clinically meaningful, and relevant predictive performance for the identification of transudative pleural effusions. In contrast, age and female sex did not contribute significantly as independent predictors in the final multivariable model.

Acknowledgments

The remaining authors thank Cyro Teixeira da Silva Junior for his contributions to the planning, design, statistical analyses, interpretation, writing of the statistical section of the study, and publication. We thank and dedicate this paper to Full Professor José Manoel Jansen ("in memoriam") of the Rio de Janeiro State University and Federal Fluminense University for his contributions to Brazilian pulmonology associations, medical education, and research.

Additional Information

The authors composed this manuscript without employing any artificial intelligence tools for text generation, data analysis, or the creation of tables and figures.

Abbreviations

AIC, Akaike information criterion; aOR, adjusted ORs, AUC, area under the ROC curve; CI, confidence interval; DCA, decision curve analysis; EPV, event per variable; HL test, Hosmer–Lemeshow goodness-of-fit test; IQR, interquartile range and interquartile range (IQR) method; MAD, median absolute deviation; MAR, missing at random; MCAR, missing completely at ramdom; P-LDH, pleural total lactate dehydrogenase; uOR, unadjusted Odds Ratio; P-ADA, total adenosine deaminase enzyme in pleural fluids; PES, pleural effusion syndrome; PPV, potential predictor variables; ROC curve, Receiver Operating Characteristic Curve; STROBE, STrengthening the Reporting of OBservational studies in Epidemiology; TRIPOD, Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis; VATS, video-assisted thoracoscopic surgery; VIF, variance inflation factor.

REFERENCES

1. Hooper, C., Lee, Y. C. G., & Maskell, N. (2010). Investigation of a unilateral pleural effusion in adults: British Thoracic Society pleural disease guideline 2010. *Thorax, 65*(Suppl. 2), ii4–ii17.
2. Light, R. W., Macgregor, M. I., Luchsinger, P. C., & Ball, W. C., Jr. (1972). Pleural effusions: The diagnostic separation of transudates and exudates. *Annals of Internal Medicine, 77*(4), 507–513.
3. Maranhão, B. H. F., Silva Junior, C. T., Chibante, A. M. S., & Cardoso, G. P. (2010). Determination of total proteins and lactate dehydrogenase for the diagnosis of pleural transudates and exudates: Redefining the classical criterion with a new statistical approach. *Jornal Brasileiro de Pneumologia, 36*(4), 468–474.
4. Atalay, F., Ernam, D., Hasanoglu, H. C., Karalezli, A., & Kaplan, O. (2005). Pleural adenosine deaminase in the separation of transudative and exudative pleural effusions. *Clinical Biochemistry, 38*(12), 1066–1070.

5. Mehta, M., Marwah, S., Shah, H., & Trivedi, A. (2015). A study of the utility of lactate dehydrogenase, total proteins, and adenosine deaminase in the diagnosis of pleural exudates: A new statistical approach. *International Journal of Medical Science and Public Health, 4*(2), 286–290.
6. Moons, K. G. M., Altman, D. G., Reitsma, J. B., Ioannidis, J. P. A., Macaskill, P., Steyerberg, E. W., Vickers, A. J., Ransohoff, D. F., Collins, G. S., & Reitsma, J. B. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. *Annals of Internal Medicine, 162*(1), W1–W73.
7. von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Bulletin of the World Health Organization, 85*(11), 867–872.
8. Altman, D. G. (1991). *Practical statistics for medical research*. Chapman & Hall.
9. Delacour, H., Sauvanet, C., Ceppa, F., & Burnat, P. (2010). Analytical performances of the Diazyme ADA assay on the Cobas® 6000 system. *Clinical Biochemistry, 43*(18), 1468–1471.
10. Riley, R. D., Snell, K. I. E., Ensor, J., Burke, D. L., Harrell, F. E., Jr., Moons, K. G. M., & Collins, G. S. (2019). Minimum sample size for developing a multivariable prediction model: Part II—Binary and time-to-event outcomes. *Statistics in Medicine, 38*(7), 1276–1296.
11. Steyerberg, E. W. (2019). *Clinical prediction models: A practical approach to development, validation, and updating* (2nd ed.). Springer.
12. Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ, 332*(7549), 1080.
13. Maranhão, B. H. F., Silva Junior, C. T., Barillo, J. L., Souza, J. B. S., Silva, P. S., & Stirbulov, R. (2025). Total adenosine deaminase cases as an inflammatory biomarker of pleural effusion syndrome. *World Journal of Clinical Cases, 13*(19), 101850.
14. Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). John Wiley & Sons.
15. Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika, 78*(3), 691–692.
16. Fan, J., Upadhye, S., & Worster, A. (2006). Understanding ROC curves. *Canadian Journal of Emergency Medicine, 8*(1), 19–20.
17. Bandos, A. I., Rockette, H. E., & Gur, D. (2006). Resampling methods for the area under the ROC curve. In *Proceedings of the ICML 2006 Workshop on ROC Analysis in Machine Learning* (pp. 1–8).
18. Wu, J. C., Martin, A. F., & Kacker, R. N. (2016). Validation of nonparametric two-sample bootstrap in ROC analysis on large datasets. *Communications in Statistics—Simulation and Computation, 45*(5), 1689–1703.
19. Mitchell, M. N., & Chen, X. (2005). Visualizing main effects and interactions for binary logit models. *The Stata Journal, 5*(1), 64–82.
20. Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology, 49*, 1373–1379.
21. Ogundimu, E. O., Altman, D. G., & Collins, G. S. (2016). Adequate sample size for developing prediction models is not simply related to events per variable. *Journal of Clinical Epidemiology, 76*, 175–182.
22. van Smeden, M., de Groot, J. A. H., Moons, K. G. M., Collins, G. S., Altman, D. G., Eijkemans, M. J. C., & Reitsma, J. B. (2016). No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. *BMC Medical Research Methodology, 16*, 163.
23. Behrsin, R. F., da Silva, C. T., Cardoso, G. P., Barillo, J. L., de Souza, J. B. S., & de Araújo, E. G. (2015). Combined evaluation of adenosine deaminase level and histopathological findings from pleural biopsy with Cope's needle for the diagnosis of tuberculous pleurisy. *International Journal of Clinical and Experimental Pathology, 8*, 7239–7246.
24. Abrao, F. C., Abreu, I. R. L. B., Miyake, D. H., *et al.,*(2014). Role of adenosine deaminase and the influence of age on the diagnosis of pleural tuberculosis. *International Journal of Tuberculosis and Lung Disease, 18*(11), 1363–1369.
25. Jadhav, A. A., & Bardapurkar, J. S. (2007). Diagnostic value of adenosine deaminase to differentiate exudates and transudates. *Indian Journal of Physiology and Pharmacology, 51*, 170–174.
26. 26.Cavanaugh, J. E., & Neath, A. A. (2019). The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *Wiley Interdisciplinary Reviews: Computational Statistics, 11*(3), e1460.
27. Jakobsen, J. C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials: A practical guide with flowcharts. *BMC Medical Research Methodology, 17*, 162.
28. D'Angelo, G., & Ran, D. (2025). Tutorial on Firth's logistic regression models for biomarkers in preclinical space. *Pharmaceutical Statistics, 24*(1), e2422.
29. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
30. Vickers, A. J., van Calster, B., & Steyerberg, E. W. (2016). Net benefit approaches to the evaluation of

prediction models, molecular markers, and diagnostic tests. *BMJ, 352*, i6.
31. Steyerberg, E. W., Harrell, F. E., Jr., Borsboom, G. J. J. M., Eijkemans, M. J. C., Vergouwe, Y., & Habbema, J. D. F. (2001). Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology, 54*, 774–781.
32. Riley, R. D., Debray, T. P. A., Collins, G. S., Archer, L., Ensor, J., van Smeden, M., & others. (2021). Minimum sample size for external validation of a clinical prediction model with a binary outcome. *Statistics in Medicine, 40*(19), 4230–4251.
33. Su, X., Ambler, G., Green, N., & Pavlou, M. (2026). The use of variable selection in clinical prediction modelling for binary outcomes: A systematic review. *Journal of Clinical Epidemiology, 197*, 112371.
34. Duquenne, P., Liposits, G., Vonnes, C. O., Navarrete, E., Serrano, A. G., Canoui-Poitrine, F., *et al.,*(2026). Prediction models for overall survival and all-cause mortality risk in older adults with cancer: A systematic review. *The Lancet Healthy Longevity, 7*(3), 100829.
35. Vargas, F. S., Teixeira, L. R., & Marchi, E. (2004). *Derrame pleural* (1st ed.). Roca.