Article

Not peer-reviewed version

Machine Learning Models to Establish the Risk of Being a Carrier of Multi-Drug-Resistant Bacteria upon Admission to the ICU

Sulamita Carvalho-Brugger * , Mar Miralbés Torner , Gabriel Jimenez Jimenez , Montserrat Vallverdú Vidal , Begoña Balsera Garrido , Francisco Xavier Nuvials , Mercedes Palomar Martínez , Jose Javier Trujillano Cabello

Posted Date: 27 August 2025

doi: 10.20944/preprints202508.1877.v1

Keywords: multi-drug resistant bacteria; ICU; risk factors; machine learning

 doi:10.20944/preprints202508.1877.v1

Article

Machine Learning Models to Establish the Risk of Being a Carrier of Multi-Drug-Resistant Bacteria Upon Admission to the ICU

Sulamita Carvalho-Brugger [1,2,*], Mar Miralbés Torner [1,2], Gabriel Jiménez Jiménez [1,2], Montserrat Vallverdú Vidal [1,2], Begoña Balsera Garrido [1], Francisco Xavier Nuvials [3], Mercedes Palomar Martínez [1] and Jose Javier Trujillano Cabello [1,2]

[1] Arnau de Vilanova University Hospital, Lleida, Spain
[2] IRB-Lleida. Institut de Recerca Biomédica, Lleida, Spain
[3] Intensive Care Unit, Vall d'Hebron Hospital, Barcelona, Spain
* Correspondence: sulabrugger@gmail.com

Abstract

Objectives: To establish the risk of being a carrier of multidrug-resistant bacteria (MDR) upon ICU admission, according to the risk factors (RF) from the Spanish "Resistencia Zero" (RZ) project checklist, using machine learning methodology. Methods: A retrospective cohort study, conducted with a consecutive sample of patients admitted to the ICU between 2014 and 2016. The study analyzed the RZ RF for MDR, as well as other pathological variables and comorbidities. The study group was randomly divided into a Development Group (70%) and a Validation Group (30%). Several machine learning models were used: binary logistic regression, CHAID-type decision tree, and the XGBOOST methodology with SHAP analysis. Results: Data from 2459 patients were analyzed, of whom 210 (8.2%) were carriers of MDR. The risk grew with the accumulation of RF. Binary logistic regression identified colonization or previous infection by MDR, prior antibiotic treatment, living in nursing home, recent hospitalization, and renal failure as the most influential factors. The CHAID tree detected MDR in 56% of patients with previous colonization or infection, a figure that increased to almost 74% if they had also received antibiotic therapy. The XGBOOST model determined that variables related to antibiotic treatment were the most important. Conclusions: The RZ RF have limitations in predicting MDR upon ICU admission, and machine learning models offer certain advantages. Not all RF have the same importance, but their accumulation increases the risk. A group of patients with no identifiable RF, complicating decisions on preventive isolation.

Keywords: multi-drug resistant bacteria; ICU; risk factors; machine learning

1. Introduction

Multidrug-resistant bacteria (MDR) are an escalating concern and are considered a priority in public health. They have been linked to severe infections with poor therapeutic outcomes, increased hospital stays, and higher healthcare costs [1–6]. Intensive Care Units (ICUs) have been identified as major contributors to the emergence of MDR within hospitals, making them key targets for controlling these pathogens [7]. However, epidemiological surveillance studies have shown that a considerable proportion of patients are already carriers of MDR upon ICU admission, either as an infection or colonization [8–10].

In response, the Spanish Society of Intensive and Critical Care Medicine and Coronary Units (SEMICYUC), supported by the Ministry of Health, launched the "Resistencia Zero" (RZ) project in 2014. This initiative aimed to reduce the incidence of ICU-acquired infections caused by MDR pathogens by 20%. Secondary objectives included understanding the epidemiology of MDR

 doi:10.20944/preprints202508.1877.v1

infections in Spanish ICUs, distinguishing between imported and ICU-acquired MDR cases, promoting and strengthening safety policies within ICUs, and implementing evidence-based safe practices [11,12].

RZ project recommendations include active screening for MDR in all patients upon ICU admission and at least weekly thereafter. Contact precaution measures—hand hygiene, disposable gloves and gowns, and, if possible, single-patient rooms—should be applied to patients at high risk of carrying MDR according to a risk factor (RF) checklist. This checklist includes six recognized RFs: recent hospital admission and antibiotic therapy, prior MDR carriage, renal replacement therapy, and certain chronic conditions that increase colonization risk, such as chronic ulcers and cystic fibrosis.

However, as demonstrated in some studies [8,9], the sensitivity and specificity of these RFs, recognized in the literature and the RZ checklist, are insufficient to efficiently prevent MDR transmission and guide empirical antibiotic therapy.

The application of machine learning (ML) techniques, using variables present at ICU admission, may provide better risk stratification for MDR carriage. Among ML techniques, we aim to use those with clinical interpretability, an appealing feature for their implementation in clinical settings [13].

The goal of this study is to determine the risk of MDR carriage at ICU admission based on the RF checklist from the RZ project, using machine learning methodology.

2. Materials and Methods

This was a retrospective, observational study conducted at a single center, the Intensive Care Unit of Arnau de Vilanova University Hospital in Lleida (HUAV), a 22-bed multidisciplinary ICU. Data were collected from April 2014 to December 2016, during the implementation of the RZ program in Spain.

Patients included were those admitted to the ICU who underwent active MDR screening through mucosal swabs (nasal, pharyngeal, axillary, rectal) within the first 48 hours, as per RZ recommendations, in addition to diagnostic cultures from clinical samples (blood, urine, sputum, tracheal or bronchoalveolar aspirates, surgical wound swabs, or others) based on medical criteria. Patients under 15 years old and those without microbiological cultures performed were excluded.

Patients and/or their families were informed about the microbiological procedures and preventive isolation policy. The HUAV ethics committee (CEIC-3025) approved the study. The development of the models followed the recommendations from the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) initiative [14].

The study group was randomly divided into a Development Group (DG) and Validation Group (VG) (70:30). A patient was considered an MDR carrier at admission if any of the surveillance cultures or clinical samples collected within the first 48 hours tested positive for MDR. MDR bacteria included methicillin-resistant Staphylococcus aureus (MRSA), vancomycin-resistant Enterococcus spp. (VRE), extended-spectrum beta-lactamase-producing Enterobacterales (ESBL), carbapenemase-producing gram-negative bacteria, multidrug-resistant Pseudomonas aeruginosa (resistant to more than three common antibiotic families), and carbapenem-resistant Acinetobacter baumannii. Other gram-negative bacteria resistant to three or more antibiotic families or producing other resistance mechanisms, such as AMP-C or Stenotrophomonas maltophilia, were classified as “others” [11,12].

Variables collected at ICU admission included patient data from the ENVIN-HELICS registry [15] (available at http://hws.vhebron.net/envin-helics/): age, sex, diabetes mellitus (DM), acute or chronic renal failure, immunosuppression, previous malignancy, liver cirrhosis, chronic obstructive pulmonary disease (COPD), malnutrition, and organ transplant. Other data included the source of admission (community, nursing home, other institution, hospital ward, or another ICU), reason for admission (medical, elective surgery, emergency surgery, trauma, or coronary), and whether antibiotic treatment was indicated at ICU admission.

The RZ RFs included hospitalization for more than five days in the last three months, institutionalization, history of MDR carriage, antibiotic therapy for more than seven days in the month before admission, hemodialysis or peritoneal dialysis, and chronic conditions with a high

 doi:10.20944/preprints202508.1877.v1

incidence of colonization/infection by MDR (cystic fibrosis, bronchiectasis, chronic ulcers, etc.) [11,12].

Models were developed in the DG and validated in the VG.

Binary logistic regression (LR) was used for variable selection, including those with a univariable p-value <0.1 in the multivariable model. A stepwise approach selected significant variables, and coefficients were rounded to the nearest integer for a scoring system (simple score).

A decision tree model using CHAID (Chi-square automatic interaction detection) employed cross-validation (five partitions) and a stopping rule with a minimum terminal node size of 10 records.

An XGBoost model was also developed, using gradient-boosted classification trees to improve total error. Parameters included maximum tree depth of 4 and a learning rate of 0.05. The SHAP (SHapley Additive exPlanations) analysis was used to interpret variable importance in the XGBoost model [16].

The models' accuracy was assessed in both DG and VG by calculating sensitivity (S), specificity (E), positive predictive value (PPV), negative predictive value (NPV), and correct classification percentage. Discrimination was assessed using ROC curves and the area under the curve (AUC), while calibration was evaluated with calibration curves.

Statistical analysis described categorical variables as percentages and continuous variables as medians (interquartile range) due to non-normal distribution (Kolmogorov-Smirnov test). The Mann-Whitney test was used for continuous variables, and the Chi-square test for categorical variables, with $p<0.05$ considered significant. Analyses were performed using SPSS software (version 29.0) and R statistics 4.0.3 with the lrm and SHAPforxgboost packages.

3. Results

During the study period, 2,484 patients aged between 15 and 98 years (mean age 59.4 years) were admitted to our ICU. A total of 25 patients were excluded due to loss to follow-up or lack of microbiological cultures within the first 48 hours of admission, leaving 2,459 patients for the study group. Figure 1 illustrates the patient selection diagram for the Development (GD) and Validation (GV) groups.

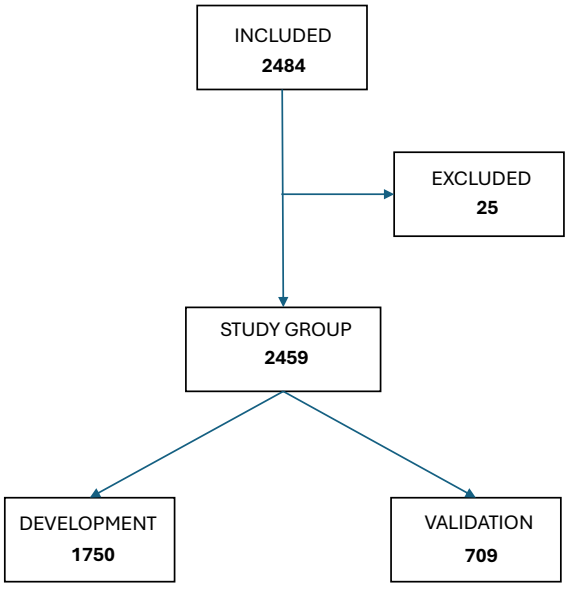


Figure 1. Flowchart of study group selection.

 doi:10.20944/preprints202508.1877.v1

Of the patients, 62.9% (1,547) were male. Among all admissions, 803 (32.7%) met one or more criteria from the RZ checklist, leading to the application of contact precautions for these patients. Table 1 describes the demographic characteristics of the GD and GV groups. Approximately one-quarter of ICU admissions presented at least one risk factor for MDR colonization, with prior hospitalization being the most frequent RF, followed by prior antibiotic use.

Table 1. Demographic characteristics of ICU patients (n = 2459), by DEVELOPMENT and VALIDATION groups.

Variable	ALL PATIENTS N = 2459	DEVELOPMENT N = 1750	VALIDATION N = 709	p-value
Ege (years)	59,4 ± 16	59,0 ± 16	60,2 ± 16	0,089
Gender (% male)	1547 (62,9)	1092 (62,4)	455 (64,2)	0,409
Medical history				
DM	646 (26,3)	450 (25,7)	196 (27,6)	0,325
COPD	328 (13,3)	229 (13,1)	99 (14,0)	0,562
CKD	729 (29,6)	532 (30,4)	197 (27,8)	0,198
Neutropenia	33 (1,3)	16 (2,3)	17 (1,0)	0,121
Immunosuppression	241 (9,8)	177 (10,1)	64 (9,0)	0,411
Neoplasia	820 (33,3)	572 (32,7)	248 (35,0)	0,275
Cirrhosis	94 (3,8)	73 (4,2)	21 (3,0)	0,157
Solid organ transplant	21 (0,9)	16 (1,0)	17 (2,3)	0,120
Origin				0,833
ED	1864 (75,8)	1332 (76,1)	532 (75,0)	
Nursing home	24 (1,0)	18 (1,0)	6 (0,8)	
Hospital ward	540 (22,0)	377 (21,5)	163 (23,0)	
Other ICU	31 (1,3)	23 (1,3)	8 (1,1)	
Diagnostic				0,763
Medical	1063 (43,2)	764 (43,7)	299 (42,2)	
Elective surgery	800 (32,5)	572 (32,7)	228 (32,2)	
Urgent surgery	261 (10,6)	182 (10,4)	79 (11,1)	
Trauma	335 (13,6)	232 (13,3)	103 (14,5)	
TYPE (Medical + urgent surgery)	1324 (53,8)	946 (54,1)	378 (53,3)	0,738
Any RZ criterion	803 (32,7)	575 (32,9)	228 (32,2)	0,738
RZ isolation criterion				
Prior hospital admission	635 (25,8)	460 (26,3)	175 (24,7)	0,411
Institutionalized	51 (2,1)	37 (2,1)	14 (2,0)	0,826
Previous colonization	65 (2,6)	41 (2,3)	24 (3,4)	0,144
Previous ATB treatment	325 (13,2)	232 (13,3)	93 (13,1)	0,926
Dyalisis	25 (1,0)	19 (1,1)	6 (0,8)	0,592
Chronic patient	60 (2,4)	40 (2,3)	20 (2,8)	0,436
ATB at ICU ADMISSION	1326 (53,9)	955 (54,6)	371 (52,3)	0,312
APACHE II (score)	12 (7-20)	12 (7-20)	13 (7-19)	0,290
ICU length of stay (days)	3 (2-8)	3 (2-7)	3 (2-9)	0,677
Hospital mortality n (%)	191 (7,8%)	135 (7,7%)	56 (7,9%)	0,877

Values are expressed as percentages, mean ± standard deviation, or median (interquartile range). MDR: multidrug-resistant bacteria. DM: diabetes mellitus. COPD: chronic obstructive pulmonary disease. CKD: chronic kidney disease. ED: emergency department. RZ: Zero Resistance program. ATB: antibiotic therapy. APACHE II: Acute Physiology and Chronic Health Evaluation II. p-value: from chi-square test or Mann–Whitney test.

A total of 210 patients (8.2%) were identified as MDR carriers, with 222 MDRs isolated (some patients carried more than one MDR). Among these, 92 patients (43.8%) carried extended-spectrum beta-lactamase-producing Enterobacteriaceae (ESBL), 78 (37.1%) carried methicillin-resistant Staphylococcus aureus (MRSA), 24 (11.4%) carried MDR Pseudomonas aeruginosa, 13 (6.2%) carried Acinetobacter baumannii, 5 (2.4%) carried carbapenemase-producing bacteria, and 10 (4.8%) carried other MDR Gram-negative bacteria. In 57 cases (27.1%), MDR presence was associated with infection;

 doi:10.20944/preprints202508.1877.v1

in the remainder, it represented colonization. Table 2 details the characteristics of patients with isolated MDRBs in the GD.

Table 2. Demographic characteristics of ICU patients according to the isolated MDR organism (n = 210).

Variable	ESBL N = 92	MRSA N = 78	*P. aeruginosa* N = 24	*A. baumannii* N = 13	OTHERS N = 15
Age (yeas)	61,7 ± 16	59,6 ± 16	63,9 ± 13	63,6 ± 13	67,0 ± 12
Gender (% male)	53 (57,6)	55 (70,5)	20 (83,3)	11 (84,6)	13 (86,7)
Medical hostory					
DM	39 (42,4)	30 (38,5)	5 (20,8)	3 (23,1)	3 (20,0)
COPD	22 (23,9)	17 (21,8)	6 (25,0)	0	3 (20,0)
CKD	40 (43,5)	35 (44,9)	13 (54,2)	6 (46,2)	10 (66,7)
Neutropenia	0	1 (1,3)	3 (12,5)	0	0
Inmunosuppression	13 (14,1)	3 (3,8)	10 (41,7)	2 (15,4)	3 (20,0)
Neoplasia	31 (33,7)	16 (20,5)	11 (45,8)	3 (23,1)	5 (33,3)
Cirrhosis	8 (8,7)	3 (3,8)	2 (8,3)	2 (15,4)	2 (13,3)
Solid organ transplant	0	2 (2,6)	2 (8,3)	3 (23,1)	1 86,7)
Origin					
ED	52 (56,5)	50 (64,1)	9 (37,5)	8 (61,5)	5 (33,3)
Nursing home	2 (2,2)	0	2 (8,3)	0	0
Hospital ward	34 (37,0)	26 (33,3)	13 (54,2)	5 (38,5)	10 (66,7)
Other ICU	4 (4,3)	2 (2,6)	0	0	0
Diagnostic					
Medical	63 (68,5)	45 (57,7)	15 (62,5)	6 (46,2)	8 (53,3)
Elective surgery	14 (15,2)	18 (23,1)	4 (16,7)	1 (7,7)	0
Urgent surgery	10 (10,9)	8 (10,3)	4 (16,7)	3 (23,1)	6 (40,0)
Trauma	5 (5,4)	7 (9,0)	1 (4,2)	3 (23,1)	1 (6,7)
TYPE (Medical + urgent surgery)	73 (79,3)	53 (67,9)	19 (79,2)	9 (69,2)	14 (93,3)
Any RZ criterion	61 (66,3)	40 (51,3)	23 (95,8)	6 (46,2)	13 (86,7)
RZ isolation criterion					
Prior hospital admission	45 (48,9)	30 (38,5)	18 (75,0)	6 (46,2)	13 (86,7)
Institucionalized	8 (8,7)	5 (6,4)	2 (8,3)	0	0
Previous colonization	18 (19,6)	17 (21,8)	8 (33,3)	0	1 (6,7)
Previous ATB treatment	40 (43,5)	20 (25,6)	14 (58,3)	4 (30,8)	8 (53,3)
Dyalisis	2 (2,2)	4 (5,1)	0	1 (7,7)	0
Chronic patient	4 (4,3)	4 (5,1)	2 (8,3)	0	0
ATB at ICU ADMISSION	67 (72,8)	52 (66,7)	20 (83,3)	11 (84,6)	14 (93,3)
APACHE II (score)	17 (10-26)	16 (11-23)	19 (14-23)	21 (10-27)	23 (14-30)
Hospital mortality n (%)	9 (9,8)	9 (11,5)	3 (12,5)	2 (15,4)	3 (20,0)

Values are expressed as percentages, mean ± standard deviation, or median (interquartile range). ESBL: Extended-Spectrum Beta-Lactamase. MRSA: Methicillin-Resistant *Staphylococcus aureus*. COPD: Chronic Obstructive Pulmonary Disease. CKD: Chronic Kidney Disease. ED: emergency department. RZ: Zero Resistance program. ATB: Antibiotic. therapy. APACHE II: Acute Physiology and Chronic Health Evaluation II.

Of the MDR carriers, 132 patients (62.8%) had risk criteria according to the RZ checklist. Among these, 50 patients (38%) had one RF, 53 (40%) had two, and 29 (22%) had three or more RFs, indicating an accumulation of risk. The presence of more RFs correlated with a higher probability of MDR carriage ($p<0.001$). In 78 MDR carriers (37%), none of the RFs from the RZ project were identified. Table 3 also provides univariable risk values (odds ratios with 95% confidence intervals) and highlights candidate variables for inclusion in multivariable models.

Using binary logistic regression, we identified that prior MDRB colonization or infection, prior antibiotic use, institutionalization, recent hospitalization, and renal failure were the most influential factors for MDR presence upon ICU admission (Table 4). The simple score methodology assigns weighted scores to these variables, reflecting their importance. Figure 2 shows MDR prevalence in GD and GV groups according to the simple score.

 doi:10.20944/preprints202508.1877.v1

Table 3. Demographic characteristics of the DEVELOPMENT group (n = 1750), according to detection of multidrug-resistant (MDR) bacteria carriers.

Variable	NO MRD N = 1600	MRD N = 150	p-value	OR (IC 95%)
Ege (years)	58,9 ± 16	60,2 ± 17	0,270	1,00 (0,90-1,01)
Gender (% male)	985 (61,6)	107 (71,3)	0,018	1,55 (1,07-2,24)
Medical history				
DM	398 (24,9)	52 (34,7)	0,009	1,60 (1,12-2,28)
COPD	199 (12,4)	30 (20,0)	0,009	1,76 (1,15-2,70)
CKD	462 (28,9)	70 (46,7)	< 0,001	2,15 (1,53-3,02)
Neutropenia	15 (0,9)	2 (1,3)	0,636	1,42 (0,32-6,30)
Inmunosuppression	154 (9,6)	23 (15,3)	0,027	1,70 (1,05-2,73)
Neoplasia	528 (33,0)	44 (29,3)	0,360	0,84 (0,58-1,22)
Cirrhosis	60 (3,8)	13 (8,7)	0,004	2,44 (1,30-4,55)
Solid organ tarnsplant	11 (0,7)	6 (4,0)	< 0,001	6,01 (2,19-16,5)
Origin			< 0,001	
ED	1246 (77,9)	86 (57,3)		
Nursing home	15 (0,9)	3 (2,0)		
Hospital ward	322 (20,1)	55 (36,7)		
Other ICU	17 (1,1)	6 (4,0)		
Diagnostic			< 0,001	
Medical	674 (42,1)	90 (60,0)		
Elective surgery	550 (34,4)	22 (14,7)		
Urgent surgery	160 (10,0)	22 (14,7)		
Trauma	216 (13,5)	16 (10,7)		
TYPE (Medical + Urgent surgery)	834 (52,1)	112 (74,7)	< 0,001	2,70 (1,85-3,96)
Any RZ criterion	481 (30,1)	94 (62,7)	< 0,001	3,90 (2,76-5,53)
RZ isolation criterion				
Prior hospital admission	384 (24,0)	76 (50,7)	< 0,001	3,25 (2,31-4,57)
Institucionalized	28 (1,8)	9 (6,0)	< 0,001	3,58 (1,66-7,74)
Previous colonization	18 (1,1)	23 (15,3)	< 0,001	15,9 (8,37-30,2)
Previous ATB treatment	174 (10,9)	58 (38,7)	< 0,001	5,17 (3,59-7,44)
Dyalisis	16 (1,0)	3 (2,0)	0,258	2,02 (0,58-7,01)
Chronic patient	35 (2,2)	5 (3,3)	0,369	1,54 (0,59-3,58)
ATB at ICU ADMISSION	845 (52,8)	110 (73,3)	< 0,001	2,46 (1,69-3,58)
APACHE II (*score*)	12 (7-19)	16 (11-25)	< 0,001	
ICU length of stay (days)	3 (2-7)	5 (2-9)	0,001	
Hospital mortality n (%)	119 (7,4)	16 (10,7)	0,156	

Values expressed as percentages, mean ± standard deviation, or median (interquartile range). MDR: Multidrug-resistant bacteria. DM: Diabetes mellitus. COPD: Chronic obstructive pulmonary disease. CKD: Chronic kidney disease. ED: emergency department. RZ: Zero Resistance program. ATB: antibiotic therapy. APACHE II: Acute Physiology and Chronic Health Evaluation II. p-value: Using chi-square test or Mann–Whitney test. OR: Odds ratio. (CI): Confidence interval. (C): Candidate variables.

 doi:10.20944/preprints202508.1877.v1

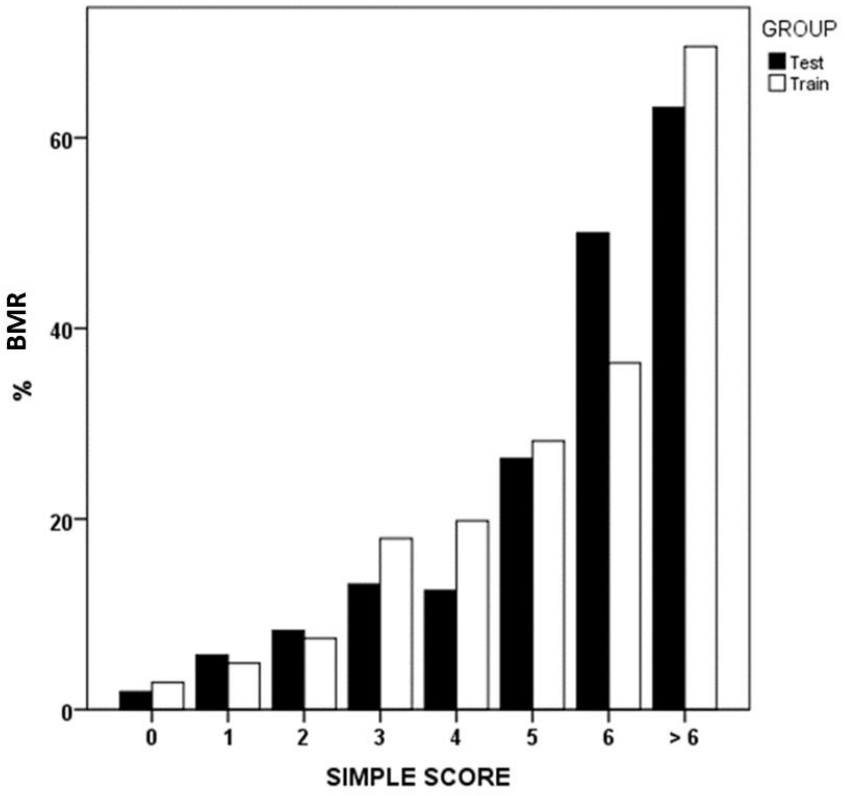


Figure 2. Simple Score values and percentage of multidrug-resistant bacteria (MDR). Development (test) and Validation (train) groups.

Table 4. Logistic regression model of factors influencing the presence of MDR bacteria.

Variable	β coefficient	OR (CI 95 %)	p-value	*Score*
Previous colonization	2,267	9,7 (4,8-19,3)	< 0,001	4
Previous ATB treatment	1,034	2,8 (1,8-4,4)	< 0,001	2
Institucionalizad patient	1,014	2,8 (1,2-6,5)	0,019	2
ATB at ICU admission	0,575	1,8 (1,2-2,7)	0,006	1
Prior hospital admission	0,519	1,7 (1,1-2,5)	0,016	1
Renal failure	0,427	1,5 (1,1-2,2)	0,025	1

MDR: multidrug resistant bacteria. ATB: Antibiotic. OR: Odds Ratio. CI: Confidence interval. *Score*: Score calculated based on the β coefficient value.

A CHAID classification tree (Figure 3) identified four key variables: prior colonization, prior antibiotic use, renal failure, and indication for antibiotic therapy at ICU admission. Among patients with prior colonization or infection, 56% had MDR upon ICU admission. This percentage increased to nearly 74% with prior antibiotic use. Patients without prior colonization showed 7.4% MDR prevalence, which rose to 20.7% with prior antibiotic use and 27.5% with additional renal failure history.

Table 4 highlights the SHAP methodology, ranking variables by importance. Notably, antibiotic use before or at ICU admission emerged as the most significant RFs.

 doi:10.20944/preprints202508.1877.v1

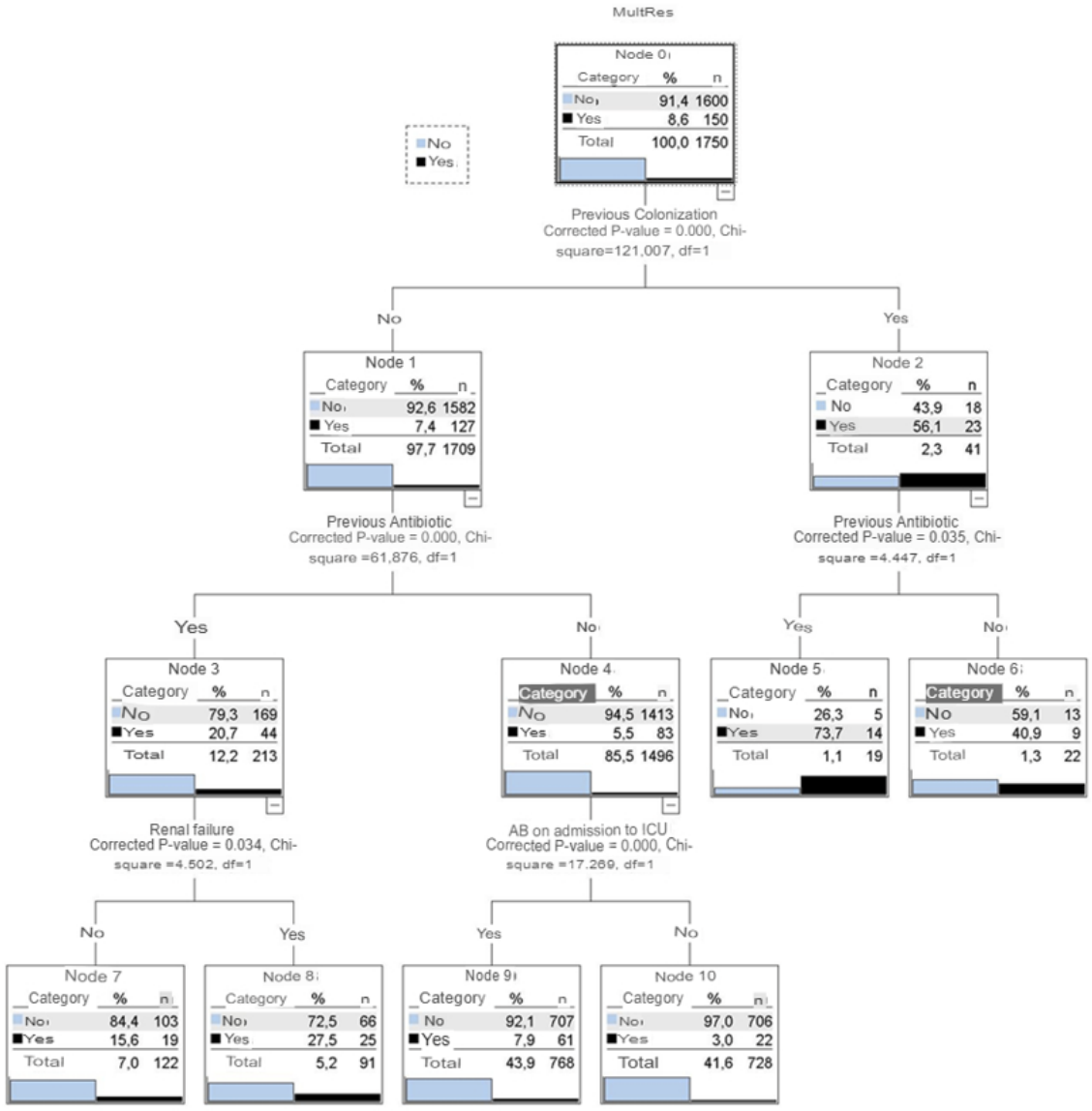


Figure 3. CHAID classification tree. 6 classification rules. MultiRes: multidrug resistant bacteria. AB: antibiotic treatment.

Figure 5 illustrates the discriminatory capacity of the models in GD and GV, measured by the area under the ROC curve (AUROC). The XGBoost model achieved the highest values, maintaining discriminatory power in GV. Figure 6 presents calibration curves, which show acceptable values, albeit with some loss of calibration in GV.

Applying the RZ criteria, isolating 32.7% of patients effectively captured 62.8% of MDR carriers. A simple score >1 improved MDR identification to 70.5% while isolating 37.1% of patients. Using the CHAID tree without the antibiotic indication variable, isolating 14.4% of patients achieved 56% MDR identification. Adding this variable increased isolation to 57.3% and detection to 83.1%. This left 1,051 patients undetected by these criteria, of whom 35 were MDR carriers. Annex 1 describes these 1,051 patients (35 MDR vs. 1,016 non-MDR). Undetected MDRs were associated with male gender, unplanned admissions, higher APACHE II scores, and no significant difference in mortality.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 August 2025 doi:10.20944/preprints202508.1877.v1

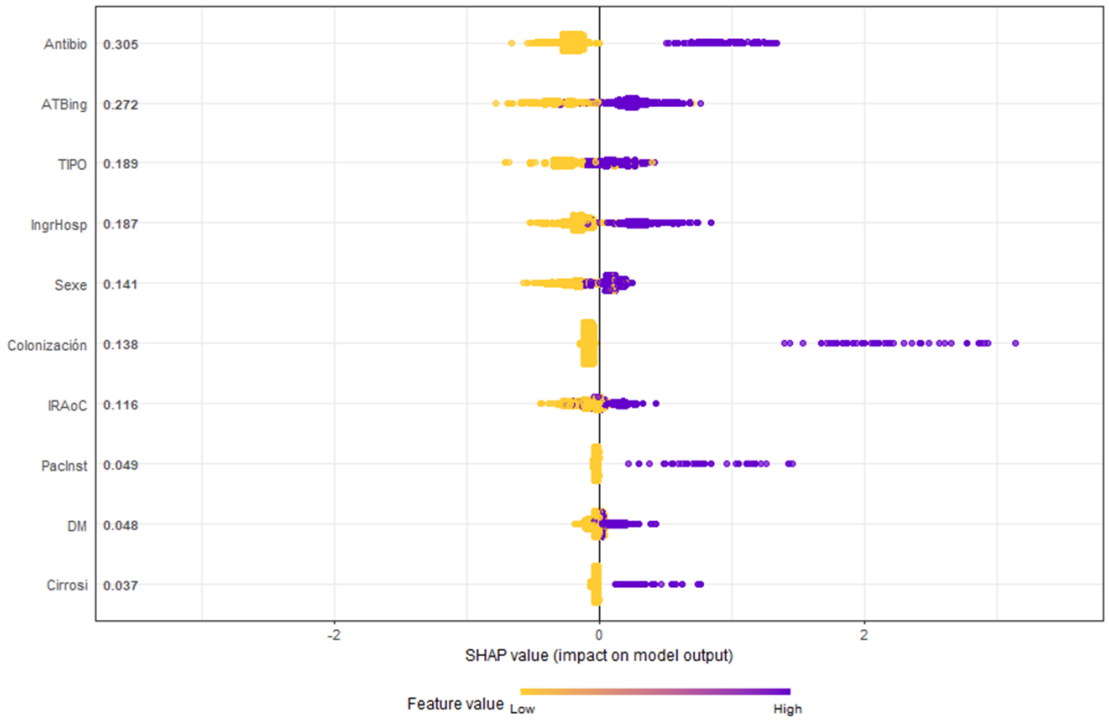


Figure 4. Honeycomb chart showing the evaluation of the variables of the GXBOOST model with SHAP analysis. ATB: previous antibiotic therapy (7 days in prior month); ATB AD: antibiotic therapy upon ICU admission; TYPE: urgent medical or surgical diagnosis; Prev hosp adm: previous hospital admission (5 days in prior 3 months); Gender: male; Renal failure: acute or chronic renal failure; Nursing home: residence in a nursing home; DM: diabetes mellitus; cirrhosis: hepatic cirrhosis.

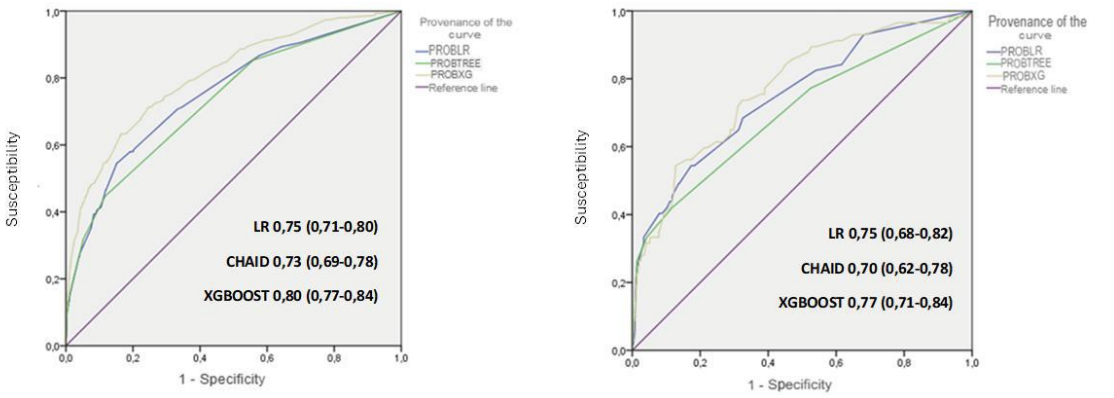
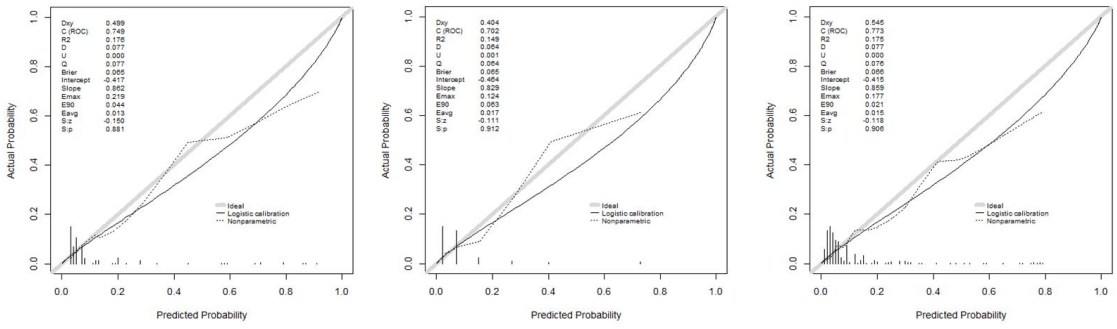


Figure 5. ROC and ABC curves of the models used. According to Development and Validation group.



 doi:10.20944/preprints202508.1877.v1

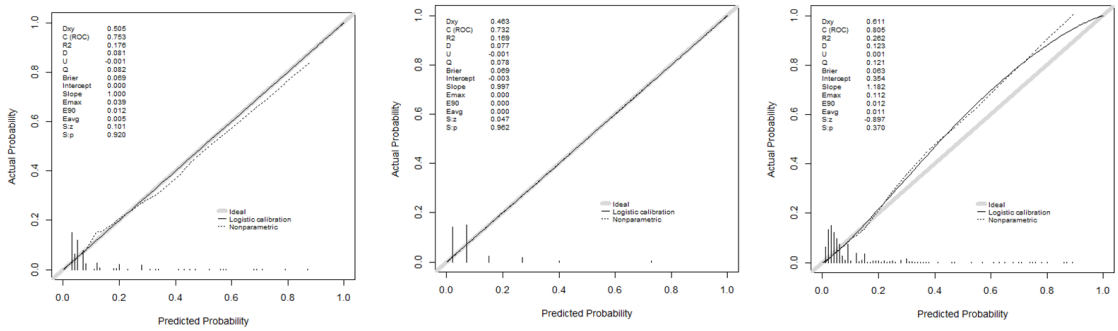


Figure 6. Calibration curves of the models used. Development and Validation group.

4. Discussion

ICU isolation for potential MDR carriers can exacerbate patient problems. While isolation and contact precautions reduce MDR transmission and outbreaks, they also have adverse effects, including psychological stress, compromised care quality, adverse reactions, and increased costs related to staff, materials, and logistics [17–19]. Conscious patients may experience social and emotional isolation, perceiving care as depersonalized. Thus, isolating all ICU admissions is suboptimal; identifying high-risk patients optimizes healthcare and improves risk-benefit ratios [20,21].

The RZ isolation criteria achieve acceptable but suboptimal performance, missing one-third of MDR carriers. A 2021 Spanish study [8] found unnecessary isolation in nearly 70% of patients based on RZ criteria, identifying prior MDR colonization/infection as the sole RF linked to MDR presence upon ICU admission. Padilla-Serrano et al [10] reported prior antibiotic use and post-surgical admissions as major RFs for ESBL Enterobacteriaceae colonization.

Our models improved MDR identification but required isolation of a higher patient percentage. MDR RFs vary; prior colonization, prolonged hospitalization, and recent antibiotics were significant in our results. However, a notable proportion of MDR carriers lacked identifiable RFs from the RZ checklist. For example, nearly half of MRSA and A. baumannii carriers had no identified RFs.

Specific ICU profiles and infection data are crucial for guiding isolation decisions [22]. Tools like the ENVIN-HELICS program track nosocomial infections and MDRs in Spanish ICUs [23].

Our study underscores that not all RFs are equally significant. For instance, prior MDR carriage and RF accumulation notably increase MDR likelihood. CHAID decision trees provide interpretable rules to identify patients for effective preventive isolation. XGBoost offers better discrimination but is less interpretable. SHAP analysis helps clarify variable importance, particularly for antibiotic use.

In conclusion, we can state that the risk factors included in the checklist for preventive isolation of patients upon ICU admission, as outlined in the RZ project, have limitations, while machine learning models offer certain advantages. Not all risk factors hold the same importance, and the decision rules provided by classification trees identify patient groups with specific characteristics. Antibiotic use, both prior to and at the time of ICU admission, is a risk factor to consider. There is a group of patients, whose specific characteristics may vary across ICUs, for whom no identifiable risk factors warranting preventive isolation are found.

Author Contributions: Conceptualization, MMP, JTC and SCB.; methodology, MMP, JTC; software, JTC, SCB.; validation, MMP, JTC, XNC.; formal analysis, JTC.; investigation, SCB, JTC.; data curation, SCB, MMT, GJJ, MVV, BBG; writing—original draft preparation, SCB.; writing—review and editing, MMP, JTC, XNC. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

 doi:10.20944/preprints202508.1877.v1

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of Arnau de Vilanova University Hospital (protocol code CEIC 3025 approved in March 2024).

Informed Consent Statement: Informed consent was not required according to the ethics committee due to the characteristics of the study, but all patients and/or family members were informed.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to privacy and ethical reasons.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ATB	antibiotic
MDR	multidrug resistant bacteria
ICU	Intensive Care Unit
RZ	Resistencia Zero Project
RF	risk factors
ESBL	extended-spectrum beta-lactamase-producing Enterobacterales
MRSA	meticilin resistant Staphylococcus aureus
SEMICYUC	Sociedad Española de Medicina Intensiva y Unidades Coronarias
ML	machine learning
HUAV	Arnau de Vilanova University Hospital
DG	Development Group
VG	Validation Group
VRE	vancomycin-resistant *Enterococcus* spp.
DM	diabetes mellitus
COPD	chronic obstructive pulmonary disease
CKD	Chronic Kidney Disease
ED	emergency department
LR	logistic regression
CHAID	Chi-square automatic interaction detection
SHAP	SHapley Additive exPlanations
S	sensitivity
E	specificity
PPV	positive predictive value
NPV	negative predictive value
AUC	area under the curve

References

1. Cağlayan Ç, Barnes SL, Pineles LL, Harris AD, Klein EY. A Data-Driven Framework for Identifying Intensive Care Unit Admissions Colonized With Multidrug-Resistant Organisms. Front Public Health. 2022 Mar 17;10.
2. Cosgrove SE, Sakoulas G, Perencevich EN, Schwaber MJ, Karchmer AW, Carmeli Y. Comparison of Mortality Associated with Methicillin-Resistant and Methicillin-Susceptible Staphylococcus aureus Bacteremia: A Meta-analysis [Internet]. Vol. 36, Clinical Infectious Diseases. 2003. Available from: https://academic.oup.com/cid/article/36/1/53/283567
3. DiazGranados CA, Zimmer SM, Klein M, Jernigan JA. Comparison of Mortality Associated with Vancomycin-Resistant and Vancomycin-Susceptible Enterococcal Bloodstream Infections: A Meta-analysis. Clinical Infectious Diseases [Internet]. 2005;41:327–33. Available from: https://academic.oup.com/cid/article/41/3/327/337451
4. Cosgrove SE, Qi Y, Kaye KS, Harbarth S, Karchmer AW, Carmeli Y. The Impact of Methicillin Resistance in Staphylococcus aureus Bacteremia on Patient Outcomes: Mortality, Length of Stay, and Hospital Charges . Infect Control Hosp Epidemiol. 2005 Feb;26(2):166–74.

 doi:10.20944/preprints202508.1877.v1

5. Song X, Srinivasan A, Plaut D, Perl TM. Effect of Nosocomial Vancomycin-Resistant Enterococcal Bacteremia on Mortality, Length of Stay, and Costs. Infect Control Hosp Epidemiol. 2003 Apr;24(4):251–6.
6. Maragakis LL, Perencevich EN, Cosgrove SE. Clinical and economic burden of antimicrobial resistance. Expert Rev Anti Infect Ther. 2008 Oct;6(5):751–63.
7. Strich JR, Palmore TN. Preventing Transmission of Multidrug-Resistant Pathogens in the Intensive Care Unit. Infect Dis Clin North Am. 2017 Sep 1;31(3):535–50.
8. Abella Álvarez A, Janeiro Lumbreras D, Lobo Valbuena B, Naharro Abellán A, Torrejón Pérez I, Enciso Calderón V, et al. Analysis of the predictive value of preventive isolation criteria in the intensive care unit. Medicina Intensiva (English Edition). 2021 May;45(4):205–10.
9. Carvalho-Brugger S, Miralbés Torner M, Jiménez Jiménez G, Badallo O, Álvares Lerma F, Trujillano J, et al. Preventive isolation criteria for the detection of multidrug-resistant bacteria in patients admitted to the Intensive Care Unit: A multicenter study within the Zero Resistance program. Medicina Intensiva (English Edition). 2023 Nov;47(11):629–37.
10. Padilla-Serrano A, Serrano-Castañeda J, Carranza-González; R, García-Bonillo M. Factores de riesgo de colonización por enterobacterias multirresistentes e impacto clínico. Rev Esp Quimioter. 2018;31(3):257–62.
11. Ministerio de Sanidad SS e I. Prevención de la Emergencia de Bacterias Multirresistentes en el Paciente Crítico - "PROYECTO RESISTENCIA ZERO" (RZ) [Internet]. 2014 [cited 2024 Apr 4]. Available from: https://seguridaddelpaciente.sanidad.gob.es/practicasSeguras/seguridadPacienteCritico/docs/PROYECTO_RZ_-_VERSION_FINAL_26MARZO_2014.pdf
12. Garnacho Montero J, Lerma FÁ, Galleymore PR, Martínez MP, Rocha LÁ, Gaite FB, et al. Combatting resistance in intensive care: The multimodal approach of the Spanish ICU "Zero Resistance" program. Vol. 19, Critical Care. BioMed Central Ltd.; 2015.
13. Yi F, Yang H, Chen D, Qin Y, Han H, Cui J, et al. XGBoost-SHAP-based interpretable diagnostic framework for alzheimer's disease. BMC Med Inform Decis Mak. 2023 Dec 1;23(1).
14. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. BMJ (Online). 2015 Jan 7;350.
15. GTEIS - SEMICYUC. SOCIEDAD ESPAÑOLA DE MEDICINA INTENSIVA, CRÍTICA Y UNIDADES CORONARIAS GRUPO DE TRABAJO DE ENFERMEDADES INFECCIOSAS "ESTUDIO NACIONAL DE VIGILANCIA DE INFECCIÓN NOSOCOMIAL EN UCI". ENVIN-HELICS 2023 [Internet]. [cited 2024 Mar 31]. Available from: https://hws.vhebron.net/envin-helics/
16. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell. 2020 Jan 1;2(1):56–67.
17. Abad C, Fearday A, Safdar N. Adverse effects of isolation in hospitalised patients: A systematic review. Vol. 76, Journal of Hospital Infection. 2010. p. 97–102.
18. Marra AR, Edmond MB, Schweizer ML, Ryan GW, Diekema DJ. Discontinuing contact precautions for multidrug-resistant organisms: A systematic literature review and meta-analysis. Am J Infect Control. 2018 Mar 1;46(3):333–40.
19. Morgan DJ, Wenzel RP, Bearman G. Contact precautions for endemic MRSA and VRE: Time to retire legal mandates. Vol. 318, JAMA - Journal of the American Medical Association. American Medical Association; 2017. p. 329–30.
20. Cooper BS, Stone SP, Kibbler CC, Cookson BD, Roberts JA, Medley GF, et al. Isolation measures in the hospital management of methicillin resistant Staphylococcus aureus (MRSA): systematic review of the literature. BMJ. 2004;329:1–8.
21. Djibré M, Fedun S, Le Guen P, Vimont S, Hafiani M, Fulgencio JP, et al. Universal versus targeted additional contact precautions for multidrug-resistant organism carriage for patients admitted to an intensive care unit. Am J Infect Control. 2017 Jul 1;45(7):728–34.

 doi:10.20944/preprints202508.1877.v1

22. López-Pueyo MJ, Barcenilla-Gaite F, Amaya-Villar R, Garnacho-Montero J. Multirresistencia antibiotica en unidades de criticos. Med Intensiva. 2011 Jan;35(1):41–53.
23. Álvarez-Lerma F, Palomar M, Olaechea P, Otal JJ, Insausti J, Cerdá E, et al. Estudio nacional de vigilancia de infección nosocomial en unidades de cuidados intensivos. Informe evolutivo de los años 2003-2005. Med Intensiva. 2007;31(1):6–17.