



## Original Article

# In-hospital Outcomes of Infective Endocarditis from 1978 to 2015: Analysis Through Machine-Learning Techniques

Plinio Resende, Jr, MD, MSc,<sup>a</sup> Claudio Querido Fortes, MD, PhD,<sup>b</sup>

Emília Matos do Nascimento, PhD,<sup>c</sup> Catarina Sousa, MD,<sup>d</sup>

Natalia Rodrigues Querido Fortes, MD,<sup>e</sup> Diego Centenaro Thomaz, MD,<sup>f</sup>

Basilio de Bragança Pereira, PhD,<sup>g</sup> Fausto J. Pinto, MD, PhD,<sup>d</sup> and

Gláucia Maria Moraes de Oliveira, MD, PhD<sup>a</sup>

<sup>a</sup>Department of Cardiology/ICES, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

<sup>b</sup>Department of Infectious Diseases, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

<sup>c</sup>Department of Statistics/UEZO, State University Center of West Zone, Rio de Janeiro, Brazil

<sup>d</sup>Faculty of Medicine, University of Lisbon, Lisbon, Portugal

<sup>e</sup>Clementino Fraga Filho Hospital, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

<sup>f</sup>School of Medicine, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

<sup>g</sup>Department of Statistics/ICES, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

## ABSTRACT

**Background:** Early identification of patients with infective endocarditis (IE) at higher risk for in-hospital mortality is essential to guide management and improve prognosis.

**Methods:** A retrospective analysis was conducted of a cohort of patients followed up from 1978 to 2015, classified according to the modified Duke criteria. Clinical parameters, echocardiographic data, and blood cultures were assessed. Techniques of machine learning, such as the classification tree, were used to explain the association

## RÉSUMÉ

**Contexte :** Le dépistage précoce des patients atteints d'endocardite infectieuse (EI) présentant un risque élevé de mortalité à l'hôpital est essentiel pour orienter la prise en charge et améliorer le pronostic.

**Méthodologie :** Une analyse rétrospective a été réalisée sur une cohorte de patients suivis de 1978 à 2015 et classés selon les critères de Duke modifiés. Les paramètres cliniques, les données des échocardiographies et les hémocultures ont été évalués. Des techniques d'apprentissage automatique, comme l'arbre de classification,

Infective endocarditis (IE) is a relatively uncommon pathology, and several questions related to it remain unanswered, such what the impact is of its initial presentation on in-hospital mortality, what the risk of surgical treatment is, especially in the presence of neurologic complications, and how complications interact to explain disease prognosis.<sup>1,2</sup> The varied clinical spectrum of IE, and the great potential

for complications, in addition to the scarcity of randomized clinical trials on IE, ensures that it remains a prognostic challenge.

Despite the unquestionable diagnostic and therapeutic advances in the management of IE in past decades, its prognosis remains unclear, with high mortality rates.<sup>1-3</sup> According to some studies,<sup>3-6</sup> the epidemiologic profile of IE evolved in the past decade, tending toward more severe clinical presentations. Most studies<sup>3,4,7-9</sup> of IE have assessed the influence of clinical complications on the course of the disease. One of those studies, from Duke University, has identified the following independent risk factors for death: diabetes, *Staphylococcus aureus*, Apache II score, and embolic events. More recently, a score has been designed to predict mortality, and 7 parameters (systolic blood pressure, heart failure, age, renal function, pneumonia, elevated peak C-reactive protein, and non-intravenous drug abuse [SHARPEN]) were

Received for publication June 1, 2021. Accepted August 2, 2021.

**Ethics Statement:** This study was submitted to and approved by the Clementino Fraga Filho Hospital Committee on Ethics and Research (CONEP # 996 2001).

Corresponding author: Dr Plinio Resende C., Jr, Av Lucio Costa, 3602 Bloco 2/1304, Rio de Janeiro- CEP 22630-010, Brazil. Tel.: (55) 21996086978; fax: (55)2124969072.

E-mail: [plinio.resende@rededor.com.br](mailto:plinio.resende@rededor.com.br)

See page 171 for disclosure information.

between clinical characteristics and in-hospital mortality. Additionally, the log-linear model and graphical random forests (GRaFo) representation were used to assess the degree of dependence among in-hospital outcomes of IE.

**Results:** This study analyzed 653 patients: 449 (69.0%) with definite IE; 204 (31.0%) with possible IE; mean age,  $41.3 \pm 19.2$  years; 420 (64%) men. Mode of IE acquisition: community-acquired (67.6%), nosocomial (17.0%), undetermined (15.4%). Complications occurred in 547 patients (83.7%), the most frequent being heart failure (47.0%), neurologic complications (30.7%), and dialysis-dependent renal failure (6.5%). In-hospital mortality was 36.0%. The classification tree analysis identified subgroups with higher in-hospital mortality: patients with community-acquired IE and peripheral stigmata on admission; and patients with nosocomial IE. The log-linear model showed that surgical treatment was related to higher in-hospital mortality in patients with neurologic complications.

**Conclusions:** The use of a machine-learning model allowed identification of subgroups of patients at higher risk for in-hospital mortality. Peripheral stigmata, nosocomial IE, absence of vegetation, and surgery in the presence of neurologic complications are predictors of fatal outcomes in machine learning-based analysis.

independently associated with inpatient mortality.<sup>9</sup> However, there is a scarcity of studies<sup>6,10</sup> on the impact of patient characteristics at the time of diagnosis, which are associated with complications and in-hospital mortality.

Recently, artificial intelligence computation techniques have been used in healthcare to assess databases, helping in the interpretation of the association of variables.<sup>11,12</sup> These techniques, known as machine learning, use computational tools that can build algorithms based on the frequency and association of observed data, determining patterns of a disease in relation to its prognosis.<sup>11</sup> These algorithms are built using data extracted from traditional statistical models, determining patterns of associated variables that facilitate understanding of complex diseases and can help in interpretation of traditional statistical data.<sup>11-13</sup> The identification of these variables together can reveal causal dependence among variables not detected by traditional statistics. A current approach extends the traditional machine-learning concept of automated decision-making to learning through data properties.<sup>13</sup> This methodology can help physicians anticipate the prognosis of a disease,<sup>11</sup> especially in IE, for which some questions remain unanswered. To the best of our knowledge, this technique has been applied very little in the field of IE stratification.

The objective of this study is to assess the in-hospital outcomes, complications, and mortality of IE, based on its initial clinical presentations, by using machine-learning techniques.<sup>11-13</sup>

## Materials and methods

### Population studied

Retrospective analysis was performed of a prospective cohort of 653 patients with IE, assessed by a dedicated

multidisciplinary IE team, from August 1978 to March 2015, at the Clementino Fraga Filho Hospital (HUCFF) of the Rio de Janeiro Federal University, Rio de Janeiro, Brazil. The HUCFF is a tertiary, university-affiliated hospital, which, over the years, kept an average of 400 hospital beds, with an intensive care unit specializing in cardiology, and a cardiovascular surgery service, providing comprehensive care to patients diagnosed with IE. All patients admitted with the diagnosis of IE were included in this study, except those with a second episode of IE. The diagnosis of IE was based on the modified Duke criteria (definite or possible).<sup>14</sup> Patients included in the study before 2000 were reclassified according to the modified Duke criteria.

**Résultats :** Cette étude a permis d'analyser 653 patients : 449 (69,0 %) avec une EI avérée; 204 (31,0 %) avec une EI possible; âge moyen de  $41,3 \pm 19,2$  ans; 420 (64 %) étaient des hommes. Mode d'acquisition de l'EI : communautaire (67,6 %), nosocomial (17,0 %), indéterminé (15,4 %). Des complications sont survenues chez 547 patients (83,7 %), les plus fréquentes étant l'insuffisance cardiaque (47,0 %), les complications neurologiques (30,7 %) et l'insuffisance rénale dépendante de la dialyse (6,5 %). La mortalité hospitalière était de 36,0 %. L'analyse de l'arbre de classification a permis d'identifier des sous-groupes présentant une mortalité hospitalière plus élevée : les patients présentant une EI communautaire et des stigmates périphériques à l'admission; et les patients présentant une EI nosocomiale. Le modèle log-linéaire a montré que le traitement chirurgical était lié à une mortalité hospitalière plus élevée chez les patients présentant des complications neurologiques.

**Conclusions :** L'utilisation d'un modèle d'apprentissage automatique a permis d'identifier des sous-groupes de patients présentant un risque plus élevé de mortalité à l'hôpital. Les stigmates périphériques, l'EI nosocomiale, l'absence de végétation et la chirurgie en présence de complications neurologiques sont des prédicteurs d'issue fatale dans l'analyse basée sur l'apprentissage automatique.

multidisciplinary IE team, from August 1978 to March 2015, at the Clementino Fraga Filho Hospital (HUCFF) of the Rio de Janeiro Federal University, Rio de Janeiro, Brazil. The HUCFF is a tertiary, university-affiliated hospital, which, over the years, kept an average of 400 hospital beds, with an intensive care unit specializing in cardiology, and a cardiovascular surgery service, providing comprehensive care to patients diagnosed with IE. All patients admitted with the diagnosis of IE were included in this study, except those with a second episode of IE. The diagnosis of IE was based on the modified Duke criteria (definite or possible).<sup>14</sup> Patients included in the study before 2000 were reclassified according to the modified Duke criteria.

The variables were selected with consideration of the population epidemiologic profile, predisposing factors for IE, clinical manifestations, etiologic agents, and main complications. The following variables were analyzed: age, sex, mode of IE acquisition (community-acquired or nosocomial), presence of previous comorbidities, predisposing heart disease, causative microorganism. The clinical signs assessed were fever, cardiac murmur, and presence of peripheral stigmata. Because of the low frequency of embolic phenomena (peripheral emboli, subconjunctival hemorrhages, Janeway lesion) and immune phenomena (Osler's nodules, Roth's spots) in isolation, these were all grouped under the same category of peripheral stigmata. The echocardiographic findings consisted of heart valve disease diagnosis, presence of vegetation on the first transthoracic echocardiography (TTE), and valvar regurgitation  $\geq$  moderate on the first TTE. Ejection fraction was not included in the initial list of collected variables and, therefore, was not retrievable in all cases. Heart failure (HF) was diagnosed based on clinical findings, response to specific therapy, and qualitative echocardiographic analysis of ventricular function. Renal failure consisted of deterioration of

renal function requiring dialysis treatment. All cases of neurologic complications were clinically suspected and confirmed by computed tomography. The dedicated multidisciplinary team collected all data by review of physical charts and confirmed the diagnosis of peripheral stigmata.

The cohort was analyzed as a total dataset and as divided into 2 groups, according to time period (1978-1999 and 2000-2015), because of the modification to the Duke criteria for IE diagnosis<sup>14</sup> and the development of new technologies, such as automated microbiology techniques,<sup>15</sup> new culture media, and the advent of echocardiography devices with more diagnostic resources.<sup>16</sup>

This study was submitted to and approved by the HUCFF Committee on Ethics and Research (CONEP # 996 2001). Anonymization hinders the tracking of patients.

### Statistical analysis

The characteristics of the population were described as mean and standard deviation for continuous variables, and as frequency and percentage for categorical variables. Logistic regression analysis was used to determine the independent variables associated with in-hospital mortality; this analysis was followed by elastic-net regularization to initially identify the variables related to the outcome. Final logistic regression was then estimated by use of maximum likelihood. Classification tree,<sup>17,18</sup> a nonparametric statistical tool and a machine-learning technique, was used to analyze factors related to outcome. To assess the dependence degree between the variables (complications and death), the log-linear model and graphical random forests (GRaFo)<sup>19</sup> representation, another technique of machine learning, were applied. The weight of the association of outcomes was determined by Cramer's V coefficient (modification of the contingency coefficient). The 5% significance level and 95% confidence interval were adopted. For statistical analysis and implementation of the classification trees, the Partykit package in R software: R Core Team (2019) was used. For implementation of the GRaFo representation, the igraph software package for complex network research was used.<sup>19</sup>

### Results

This study population comprised 653 patients, 449 of whom (69.0%) met the diagnostic criteria for definite IE, and 204 of whom (31.0%) met the diagnostic criteria for possible IE. From the total sample, 13 patients were excluded because they were transferred to another hospital, which hindered their follow-up. Table 1 shows the general characteristics of this study population.

The patients' mean age was  $41.3 \pm 19.2$  years, and 420 (64%) were male. Community-acquired IE was diagnosed in 442 patients (67.6%), and nosocomial IE in 113 (17%). Comorbidities were present in 189 patients (29%); the most prevalent were dialysis-dependent renal failure (DRF), affecting 53 patients (8%), and diabetes, affecting 49 patients (7.5%). Predisposing heart disease was identified in 348 patients (53%); the most frequent types were congenital heart disease (bicuspid aortic valve and mitral prolapse) in 134 patients (20%) and rheumatic heart valve disease in 52 patients (8%). In this study, all cases of mitral prolapse were

classified as congenital disease, although most of them could be categorized as degenerative disease.

Regarding clinical presentation, fever was present in 561 patients (86%), cardiac murmur in 544 (83%), and peripheral stigmata in 188 (28.7%). Blood cultures were positive in 361 patients (55%), negative in 260 (40%), and inconclusive in 32 (5%). The distribution of the causative microorganisms was as follows: *S. aureus*, 114 patients (17.5%); *Streptococcus viridans*, 99 (15%); *Enterococcus faecalis*, 39 (6%); *Streptococcus gallolyticus*, 20 (3%); and others, 90 (13.7%). The distribution of affected heart valves was as follows: mitral valve, 271 patients (41%); aortic valve, 148 patients (22%); tricuspid valve, 60 patients (9%). Prosthetic IE was diagnosed in 75 patients (11%), and 97 patients (14.8%) had more than one valve affected.

Complications were identified in 547 (83.7%) patients, and the most frequent were as follows: HF, 309 patients (47%); neurologic complications, 201 patients (30.7%), such as embolic stroke (14%), infectious neurologic complications (10.5%), and hemorrhagic stroke (6.2%); embolism to other organs, 90 patients (13.7%); and DRF, 43 patients (6.5%). Surgical treatment was performed in 196 patients (30%), and the in-hospital mortality rate was 36%. Table 2 shows the complications and outcomes of the cohort studied.

Differences in the general characteristics of the study population divided into 2 periods (Table 1) were that the 2000-2015 patients were older, had a higher prevalence of nosocomial IE, had a higher frequency of comorbidities and prosthetic IE, and had more than one heart valve affected. In addition, in the 2000-2015 group, a greater number of positive blood cultures and definite IE diagnoses according to Duke criteria were identified, and *E. faecalis* was the causative microorganism more often. Regarding complications and outcomes in the whole cohort (Table 2), the prevalence of DRF, the frequency of surgical treatment, and the in-hospital mortality showed an increase after 2000.

Logistic regression after elastic-net regularization evidenced the following independent variables associated with in-hospital mortality: fever, age ( $> 53$  years), mode of IE acquisition, presence of peripheral stigmata, absence of vegetation on the first TTE, and IE diagnosis based on Duke criteria. The classification tree selected subgroups of patients with similar characteristics, building an algorithm that identified patients with distinct mortality rates (Fig. 1). The presence of fever, community-acquired IE, and peripheral stigmata ( $n = 165$ ) identified a significant subgroup of patients with fatal outcome (36.4%). The classification tree was able to immediately distinguish the group with peripheral stigmata as having high mortality. In the absence of peripheral stigmata, other variables selected 2 subgroups with high mortality as follows: 43.5% mortality for patients with community-acquired IE, definite IE diagnosis according to Duke criteria, and age  $> 53$  years ( $n = 62$ ); and 38.9% mortality for patients with community-acquired IE, definitive IE diagnosis, age  $< 53$  years, and absence of vegetation on the first TTE ( $n = 36$ ). Another subgroup presenting high mortality (51.5%) was that of patients with fever and nosocomial endocarditis ( $n = 68$ ).

Figure 2 (graphical log-linear model and GRaFo representation) shows the degree of dependence between the variables (complications and death). The weights on the edges

**Table 1.** General characteristics of the whole population studied (1978-2015), and divided into 2 groups according to time periods (1978-1999 and 2000-2015)

Variable	1978–2015	1978–1999	2000–2015	<i>P</i> ( $\neq$ in the 2 periods)
Number of patients	653 (100)	452 (69.2)	201(30.8)	< 0.0001
Male sex	420 (64.3)	289 (63.9)	131 (65.1)	0.7610
Age, y	41.3 $\pm$ 19.2	37.5 $\pm$ 18.6	49.7 $\pm$ 17.8	< 0.0001
Hospital length of stay, d	48.3 $\pm$ 51.1	41.4 $\pm$ 57.1	61.2 $\pm$ 69.7	0.6390
$\Delta$ t surgery,* d	18.4 $\pm$ 21.0	15.2 $\pm$ 18.6	23.8 $\pm$ 23.3	0.0060
<b>Mode of acquisition</b>				
Community-acquired	443 (67.8)	314 (69.4)	129 (64.1)	0.1816
Nosocomial	113 (17.3)	43 (17.0)	70 (34.8)	< 0.0001
Undetermined	97 (14.8)	95 (21.0)	2 (1.0)	< 0.0001
<b>Comorbidities</b>				
Diabetes	189 (28.9)	108 (23.8)	81 (40.2)	< 0.0001
DRF	49 (7.5)	22 (4.8)	27 (13.0)	< 0.0001
HIV	53 (8.1)	15 (3.3)	38 (18.9)	< 0.0001
HIV	23 (3.5)	16 (3.5)	7 (3.4)	0.9710
<b>Predisposing heart disease</b>				
Congenital heart disease	348 (53.2)	237 (52.4)	111 (55.2)	0.5090
Rheumatic valve disease	134 (20.5)	98 (21.6)	36 (17.9)	0.2710
Rheumatic valve disease	52 (7.9)	31 (12.3)	21 (10.4)	0.1180
<b>Clinical presentation</b>				
Fever	561 (85.9)	404 (89.3)	157 (78.1)	< 0.0001
Cardiac murmur	544 (83.3)	389 (86.0)	155 (77.1)	0.0050
Peripheral stigmata	188 (28.7)	144 (31.8)	44 (21.8)	0.0090
<b>Blood cultures</b>				
Positive	361 (55.2)	226 (50.0)	135 (67.1)	< 0.0001
Negative	260 (39.8)	197 (43.5)	63 (31.3)	0.0032
Unconclusive	32 (5.3)	29 (6.4)	3 (1.5)	0.0071
<b>Microorganisms</b>				
<i>Staphylococcus aureus</i>	114 (17.4)	80 (17.6)	34 (16.9)	0.8076
<i>Streptococcus viridans</i>	99 (15.1)	70 (15.4)	29 (14.4)	0.7277
<i>Enterococcus faecalis</i>	39 (5.9)	15 (3.3)	24 (11.9)	< 0.0001
<i>Streptococcus gallolyticus</i>	20 (3.0)	12 (2.6)	8 (4.0)	0.3643
Others	90 (13.7)	50 (11.0)	40 (19.9)	0.0025
<b>Echocardiography</b>				
Valve regurgitation $\geq$ moderate	356 (54.5)	239 (52.8)	117 (58.2)	0.2070
Vegetation on the 1st TTE	452 (69.2)	302 (66.8)	150 (74.6)	0.0460
<b>Endocarditis site</b>				
Mitral valve	271 (41.5)	186 (41.1)	85 (42.2)	0.7853
Aortic valve	148 (22.6)	110 (24.3)	38 (18.9)	0.1260
Tricuspid valve	60 (9.1)	48 (10.6)	12 (5.9)	0.0576
Prosthesis	75 (11.4)	31 (6.8)	44 (21.8)	< 0.0001
Intravascular device	6 (0.9)	1 (0.2)	5 (2.4)	0.0051
More than 1 valve	97 (14.8)	54 (11.9)	43 (21.3)	0.0020
<b>IE diagnosis based on Duke criteria</b>				
Definite	449 (68.7)	293 (64.8)	156 (77.3)	0.0013
Possible	204 (31.2)	159 (35.1)	45 (22.3)	0.0013

Values are n (%) or mean ( $\pm$  SD), unless otherwise indicated.

DRF, dialysis-dependent renal failure; HIV, human immunodeficiency virus; TTE, transthoracic echocardiography.

\*  $\Delta$ t surgery: time elapsed from admission to surgery.

correspond to Cramer's V coefficients, a measure of dependence between discrete variables. Figure 2 shows a 2-by-2 interaction (blue lines) and a 3-by-3 interaction (red lines) of dependent variables. In the analysis of the whole cohort (Fig. 2A), HF was associated with DRF and surgical treatment. Neurologic complications, HF, and DRF were associated with mortality. On the 3-by-3 interaction of mortality, surgical treatment, and neurologic complications, a greater dependence between mortality and neurologic complications was observed. Separate analysis for the 1978-1999 group (Fig. 2B) and the 2000-2015 group (Fig. 2C) shows, for the former, a 3-by-3 dependence relationship of surgical treatment, mortality, and neurologic complications. The highest degree of dependence was observed between mortality and neurologic complications. In addition, the dependence relationships found in Figure 2A and Figure 2B are the same, differing only in regard to the Cramer's V coefficients. In the 2000-2015 group (Fig. 2C), there is a dependence

relationship between surgical treatment and mortality, conditioned to the presence of HF and of the other variables.

## Discussion

IE is a relatively uncommon cardiovascular disease. Despite all the diagnostic and therapeutic advances regarding IE, its prognosis remains unclear, with a mean in-hospital mortality of 20% in developed countries<sup>1-3</sup> and as much as 40% in developing countries.<sup>20</sup> In this cohort of patients, the hospital mortality rate was 36%. It was higher than that reported for developed countries, but within the range reported for developing countries,<sup>20</sup> and it was similar to that of a recently published Italian cohort (37.5%).<sup>21</sup> The elevated mortality rate in this cohort can be explained by the severity of the patients' conditions, in comparison to the population profiles from the studies conducted in developed countries. Still, being a tertiary reference centre for infectious disease management

**Table 2. Complications of infective endocarditis in the population studied (1978-2015) and divided into 2 groups according to time period (1978-1999 and 2000-2015)**

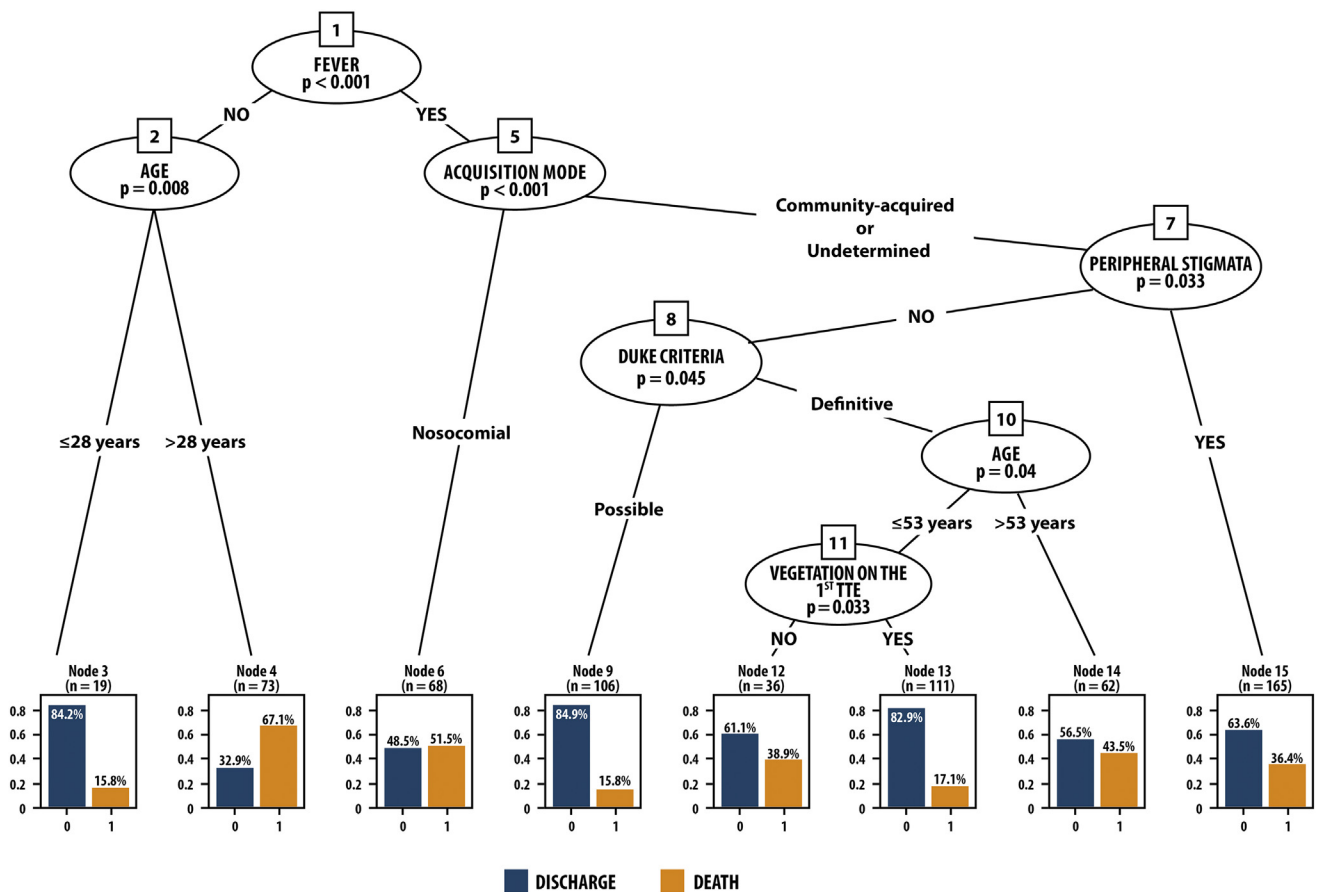
Variable	1978-2015 N = 653	1978-1999	2000-2015	P (≠ in the 2 periods)
Presence of complication	547 (83.7)	379 (83.8)	168 (83.5)	0.9320
Heart failure	309 (47.3)	219 (48.4)	90 (44.7)	0.3850
Neurological complication	201 (30.7)	147 (32.5)	54 (26.8)	0.1480
Embolic stroke	91 (13.9)	56 (12.3)	35 (17.4)	0.0870
Hemorrhagic stroke	41 (6.2)	25 (5.5)	16 (7.9)	0.2380
Embolization to other organs	90 (13.7)	63 (13.9)	27 (13.4)	0.8630
Dialysis-dependent renal failure	43 (6.5)	17 (3.7)	26 (12.9)	< 0.0001
Surgical treatment	196 (30.0)	123 (27.2)	73 (36.3)	0.0190
Outcome (in-hospital mortality)	235 (36.0)	157 (34.7)	78 (38.8)	0.0240

Values are n (%), unless otherwise indicated.

could explain this higher mortality rate, owing to the presence of an inherent selection bias toward more severe disease.

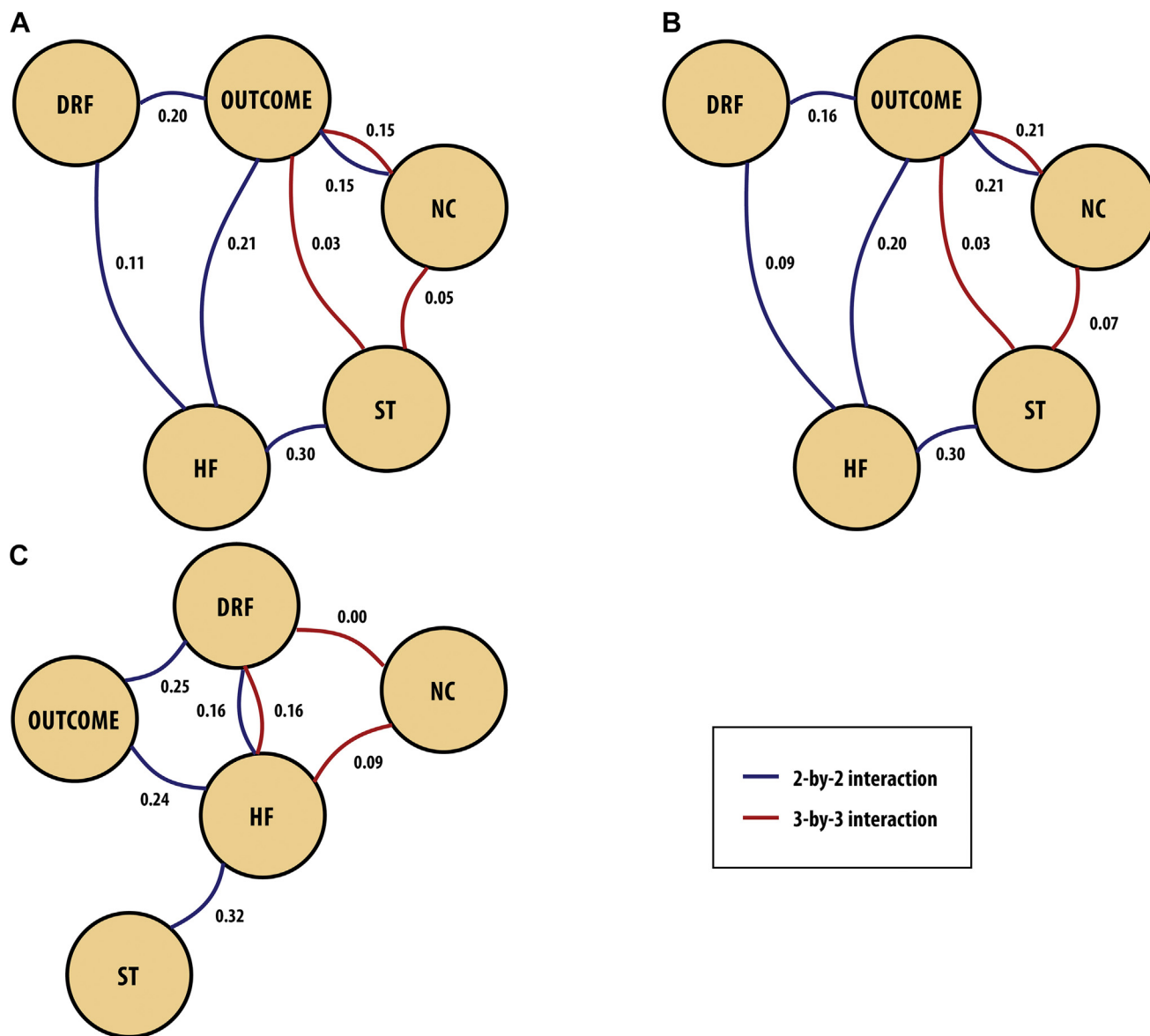
Some findings in this study were similar to those of other studies,<sup>2,22</sup> such as fever and cardiac murmur as frequent clinical presentations, predominance of male patients, and presence of predisposing conditions, the most frequent being impairment of the mitral valve followed by impairment of the aortic valve, as published in the International Collaboration on Endocarditis—Prospective Cohort Study (ICE-PCS)<sup>1</sup> and

in a systematic review recently published.<sup>23</sup> However, this study population was characterized by the involvement of younger individuals and a significant prevalence of rheumatic disease, similar to the findings of other Brazilian cohorts.<sup>7-9</sup> Regarding blood culture, a positive result was found less frequently than in the recent report from the Euro-Endo (European infective endocarditis) registry.<sup>3</sup> The possible causes for that difference were the previous use of antimicrobials and causative microorganisms that did not grow in



**Figure 1.** Classification tree. Interaction between variables: fever, age, peripheral stigmata, vegetation on the first transthoracic echocardiography (TTE), mode of acquisition, Duke criteria. Hospital discharge = 0; death = 1 (terminal nodes). The variables were grouped in the classification tree, allowing the identification of subgroups of patients with similar characteristics, for whom different mortality rates were found: Node 15 = fever + community-acquired infective endocarditis (IE) + peripheral stigmata (n = 165), 36.4% mortality; Node 14 = fever + community-acquired IE + definite IE diagnosis + age > 53 years (n = 62), 43.5% mortality; Node 12 = fever + community-acquired IE + definite IE diagnosis + age < 53 years + absence of vegetation on the first TTE (n = 36), 38.9% mortality; Node 6 = fever + nosocomial IE (n = 68), 51.5% mortality.





**Figure 2.** Graphical log-linear model and graphical random forests (GRaFo) representation. **(A)** 1978-2015, **(B)** 1978-1999, and **(C)** 2000-2015. The weights on the edges correspond to Cramer’s V coefficients (V). Outcome = death. This representation assesses the degree of dependence between the variables (complications and death). In the analysis of **(A)** the whole cohort, for the 2-by-2 interaction (**blue lines**), heart failure (HF) was associated with dialysis-dependent renal failure (DRF) ( $V = 0.11$ ) and with surgical treatment (ST) ( $V = 0.30$ ). Neurologic complications (NC), HF, and DRF were associated with mortality ( $V = 0.21, 0.20, 0.15$ , respectively). For the 3-by-3 interaction (**red lines**) of mortality, ST, and NC, a greater dependence between mortality and NC ( $V = 0.15$ ) was observed. The dependence relationships found in **(A)** and **(B)** are the same. In the 2000-2015 group **(C)**, there is a dependence relationship between surgical treatment and mortality, conditioned to the presence of HF and of the other variables.

the routine culture media.<sup>15,24</sup> In this cohort, the most prevalent microorganisms were *S. aureus* and *S. viridans*, similar to results reported in the literature.<sup>1,3,7-9,23</sup> A smaller number of patients were diagnosed with definite IE, based on the modified Duke criteria,<sup>14</sup> than the number reported in the literature.<sup>1,3</sup> This difference might be explained by the fact that our cohort was initiated when neither high-definition echocardiography nor more efficient microbiological blood culture techniques were available, and those methods are the cornerstone of the diagnosis of definitive IE.<sup>14</sup> New imaging criteria, such as radiolabeled white blood cell single-photon

emission computed tomography (SPECT)/computed tomography and computed tomography, have been included in the 2015 European Society of Cardiology Guidelines for the Management of IE, aimed at increasing diagnostic sensitivity.<sup>2</sup> For this cohort, those imaging techniques were not available, so they could not be applied.

The rates of the major complications reported in this study, such as HF, neurologic complications, and DRF, are similar to those reported in other studies,<sup>1,2,22,23</sup> and HF was the most prevalent (47% of the cases). In this study, 30% of the patients underwent surgery, whereas, in some cohorts, up

to 50% of the patients have undergone surgery.<sup>25-27</sup> The literature shows a trend that early indication of surgery to correct valvular dysfunction in IE is a factor in a better prognosis.<sup>25-28</sup>

The early recognition of factors associated with mortality is paramount to improve IE prognosis. The use of artificial intelligence methodologies, such as the classification tree<sup>17,18</sup> and the GRaFo representation,<sup>19</sup> allowed the identification of subgroups at higher risk for mortality, based on specific characteristics in the initial presentation of IE, and for development of complications over the course of disease. Few studies have assessed the impact of the initial clinical presentation of IE on prognosis.<sup>6,10</sup> Therefore, in addition to the parametric methodology, a nonparametric analysis—the classification tree—was used to identify factors on hospital admission that could be associated with unfavourable outcome during hospitalization. Decision and classification trees represent a methodology capable of dividing the sample studied into subgroups with characteristics in common.<sup>17,18</sup> Considering the wide variety of IE clinical manifestations, use of a classification tree to group those characteristics, facilitating their association with the IE outcome, can be helpful in clinical practice.

The presence of peripheral stigmata on the initial clinical presentation identified a subgroup of patients with high in-hospital mortality. On the classification tree analysis, the subgroup of patients with community-acquired IE who, on hospital admission, already had fever and peripheral stigmata, had a high in-hospital mortality rate (36.4%). In the absence of peripheral stigmata, other variables, such as the diagnosis of definite IE, age > 53 years, and absence of vegetation on the first TTE, were associated with higher mortality (43.5%). On the other hand, patients with the diagnosis of possible IE, based on the Duke criteria, had a lower mortality rate (15.1%), similar to that of patients under the age of 53 years who had a vegetation on the first TTE (17.1%). Probably the absence of vegetation on the first TEE delays the beginning of antibiotic treatment, with prognosis consequences. The subgroup of patients with nosocomial IE had the highest mortality rate (51.5%), probably because they had more severe IE. The Duke criteria used to diagnose IE might indicate prognosis, a finding not reported in the literature. Favouring this hypothesis, and based on the classification tree, the presence of peripheral stigmata, considered a minor criterion for the diagnosis of IE, was an early marker of in-hospital mortality.

Analysis of the association of IE complications and in-hospital mortality, using the GRaFo representation, showed that HF, neurologic complications, and DRF were related to in-hospital mortality. Surgical treatment is related to mortality, conditioned to the presence of HF, neurologic complications, and DRF, and the degree of dependence on neurologic complications was greater. Considering that early surgical indication has been increasingly recommended in the literature,<sup>25-31</sup> our finding suggests that surgery should be postponed in patients with neurologic complications. In this study, embolic, infectious, and hemorrhagic neurologic complications were analyzed together, which did not allow us to determine their individual influence on surgical mortality. In addition, the current literature is ambiguous regarding the safety of cardiopulmonary bypass in patients with recent neurologic injury, particularly in those with hemorrhagic

lesions or with ischemic lesions at high risk for hemorrhagic transformation.<sup>31</sup> Furthermore, other series show that early cardiac surgery after ischemic stroke alone is not contraindicated.<sup>29,32</sup> Future studies are needed to better assess the impact of stroke on the surgical prognosis of patients with IE.

In this cohort, the surgical treatment isolated did not increase in-hospital mortality. One cannot rule out the possibility that the higher severity of the disease selected the patients who might have a better outcome with surgical treatment. Despite the 2 great advances in diagnosis of IE—that is, 2-dimensional/transesophageal echocardiography and automated blood cultures—mortality from IE has not substantially decreased over the years. This lack of change might be due to the higher complexity of the patients hospitalized, higher prevalence of invasive devices, and the selection of resistant pathogens owing to the widespread use of antimicrobials, with higher probability of complications.

One of the greatest challenges in the treatment of IE is beginning antimicrobial therapy based on signs and symptoms; however, for any more severe infectious pathology, the early use of antimicrobials impacts the outcome.<sup>1-3,22,33,34</sup> This study can recommend the early start of antimicrobial therapy, according to the guidelines<sup>2,22</sup> for the management of IE, based on the finding of peripheral stigmata, because their association with increased in-hospital mortality has been well demonstrated. The classification tree provided an image of the problem and allowed the identification of a minor diagnostic criterion of IE as a predictor of mortality. The need to increase the sensitivity of the diagnostic criteria of IE is evident, because of the high mortality still associated with this pathology.<sup>1-3,20,21,33-35</sup>

This study identified the epidemiologic profile change of patients with IE from 2000 on, which has been reported in other studies.<sup>3-5,33,36-38</sup> This study found higher prevalence of comorbidities and nosocomial IE, higher frequency of prosthetic valve IE, and an increase in both positive blood cultures and number of surgeries. Regarding the causative microorganisms, *E. faecalis* was increasingly identified, being more often associated with nosocomial IE and having a greater potential for complications, both systemic and related to valvular lesion.<sup>39</sup>

### Limitations

Some limitations of this study should be considered. The long follow-up period for this cohort did not allow data on some echocardiographic parameters to be recovered, such as vegetation size, presence/absence of abscess, and left ventricular ejection fraction. The rate of positive blood cultures was lower than that reported in the literature, but this has been questioned recently in a Spanish study<sup>37</sup> that found no relation between negative blood culture and mortality. Bias of selection of patients with surgical indication might have occurred, as some patients were not submitted to surgery because of their clinical severity. A cohort with a duration this long and from one single centre is influenced by diagnostic limitations and therapeutic availabilities in each follow-up period. However, it better reflects the systematization of treatment by a multidisciplinary team from one single centre. This study, of those published to date, is one of the largest cohorts of patients with IE from a single centre who were prospectively followed up by a multidisciplinary team

dedicated to IE. Because of the complex nature of IE, these factors are much valued in the major international guidelines.<sup>2,22</sup>

Machine-learning techniques are promising tools that, through the recognition of common patterns of the characteristics of a disease, can assist in improving understanding of complex pathologies, such as EI.

## Conclusions

In this large cohort of patients with IE, by using machine learning-based analysis, subgroups were identified with the following common characteristics that correlated with increased in-hospital mortality: peripheral stigmata, nosocomial IE, absence of vegetation, and surgery in the presence of neurologic complications. Machine learning can add value to better our clinical judgement in the field of rare diseases.

## Acknowledgement

We are grateful to Stela Maris Costalonga, MD, who helped translate this article.

## Funding Sources

The authors have no sources of funding to declare.

## Disclosures

The authors have no conflicts of interest to disclose.

## References

1. Murdoch DR, Corey GR, Hoen B, et al. Clinical presentation, etiology, and outcome of infective endocarditis in the 21st century: The International Collaboration on Endocarditis-Prospective Cohort Study. *Arch Intern Med* 2009;169:463-73.
2. Habib G, Lancellotti P, Antunes Manuel J, et al. 2015 ESC guidelines for the management of infective endocarditis (ESC). *Eur Heart J* 2015;36:3075-23.
3. Habib G, Erba B, Lung B, et al. Clinical presentation, aetiology and outcome of infective endocarditis. Results of the ESC-EORP EURO-ENDO (European infective endocarditis) registry: a prospective cohort study. *Eur Heart J* 2019;40:3222-33.
4. Fernandez-Hidalgo N, Almirante B, Tornos P, et al. Contemporary epidemiology and prognosis of health care-associated infective endocarditis. *Clin Infect Dis* 2008;47:1287-97.
5. Wang A, Gaca J, Chu VH, et al. Management considerations in infective endocarditis: a review. *JAMA* 2018;320:72-83.
6. Chu VH, Cabell CH, Benjamin DK, et al. Early predictors of in-hospital death in infective endocarditis. *Circulation* 2004;109:1745-9.
7. Marques A, Cruz I, Caldeira D, et al. Risk factors for hospital mortality from infective endocarditis [Fatores de risco para mortalidade hospitalar na endocardite infecciosa]. *Arq Bras Cardiol* 2020;114:1-8.
8. Damasco PV, Correal J, Campos A, et al. Epidemiological and clinical profile of infective endocarditis at a Brazilian tertiary care center: an eight-year prospective study. *Rev Soc Bras Med Trop* 2019;52:e2018375.
9. Chee QZ, Tan YB, Ngiam JN, et al. The SHARPEN clinical risk score predicts mortality in patients with infective endocarditis: an 11-year study. *Int J Cardiol* 2015;191:273-6.
10. Nunes MCP, Guimarães- Junior MH, Pinto PH, et al. Outcomes of infective endocarditis in the current era: early predictors of a poor prognosis. *Int J Infect Dis* 2018;68:102-7.
11. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med* 2019;380:1347-58.
12. Chen C. Ascent of machine learning in medicine. *Nat Mater* 2019;18:407.
13. Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA* 2018;319:1317-8.
14. Li JS, Sexton DJ, Mick N, et al. Proposed modifications to the Duke criteria for the diagnosis of infective endocarditis. *Clin Infect Dis* 2000;30:633-8.
15. Liesman R, Pritt B, Maleszewski J, Patel R. Laboratory diagnosis of infective endocarditis. *J Clin Microbiol* 2017;55:2599-608.
16. Habib G, Badano L, Tribouilloy C, et al. Recommendations for the practice of echocardiography in infective endocarditis. *Eur J Echocardiogr* 2010;11:202-19.
17. Venkatasubramanian A, Wolfson J, Mitchell N, et al. Decision trees in epidemiological research. *Emerg Themes Epidemiol* 2017;14:11.
18. Lemon S, Roy J, Clark M, Friedmann P. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Ann Behav Med* 2003;26:172-81.
19. Csardi G, Nepusz T. The igraph software package for complex network research. Available at: <https://igraph.org>. Accessed December 22, 2021.
20. Njuguna B, Gardner A, Karwa R, Delahaye F. Infective endocarditis in low- and middle-income countries. *Cardiol Clin* 2017;35:153-63.
21. Serra N, Colomba C, Di Carlo P, et al. Infective endocarditis: preliminary results of a cohort study in the southern Italian population. *Cureus* 2020;12:e8338.
22. Nishimura RA, Otto CM, Bonow RO, et al. 2014 AHA/ACC guideline for the management of patients with valvular heart disease: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* 2014;129:2440-92.
23. Sousa C, Ribeiro R, Pinto F. The burden of infective endocarditis in Portugal in the last 30 years—a systematic review of observational studies. *Revista Portug Cardiol* 2021;40:205-17.
24. Houpiquian P, Raoult D. Blood culture-negative endocarditis in a reference center: etiologic diagnosis of 348 cases. *Medicine* 2005;84:162-73.
25. Cahil TI, Prendergast B. Infective endocarditis. *Lancet* 2016;387:882-93.
26. Martínez A, Calderón-Parra J, Miró J, et al. Effect of the type of surgical indication on mortality in patients with infective endocarditis who are rejected for surgical intervention. *Int J Cardiol* 2019;282:24-30.
27. Lamas C. Infective endocarditis still a deadly disease [Endocardite infecciosa: ainda uma doença mortal]. *Arq Bras Cardiol* 2020;114:9-11 [in Spanish].
28. Kang DH, Kim YJ, Kim SH, et al. Early surgery versus conventional treatment for infective endocarditis. *N Engl J Med* 2012;366:2466-73.
29. Barsic B, Dickerman S, Krajinovic V, et al. Influence of the timing of cardiac surgery on the outcome of patients with infective endocarditis and stroke. *Clin Infect Dis* 2013;56:209-17.
30. Tleyjeh I, Ghomrawi H, Steckelberg J, et al. The impact of valve surgery on 6-month mortality in left-sided infective endocarditis. *Circulation* 2007;115:1721-8.



31. Bonaros N, Czerny M, Pfausler B, et al. Infective endocarditis and neurologic events: indications and timing for surgical interventions. *Eur Heart J Suppl* 2020;22(Suppl M):M19-25.
32. Gaca JG, Sheng S, Daneshmand MA, et al. Outcomes for endocarditis surgery in North America: a simplified risk scoring system. *J Thorac Cardiovasc Surg* 2011;141:98-106.
33. Chambers H, Bayer A. Native-valve infective endocarditis. *N Engl J Med* 2020;383:567-76.
34. Nunes M, Gelape C, Ferrari T. Profile of infective endocarditis at a tertiary care center in Brazil during a seven-year period: prognostic factors and in-hospital outcome. *Int J Infect Dis* 2010;14:e394-8.
35. Cahill T, Baddour L, Habib G, et al. Challenges in infective endocarditis. *J Am Coll Cardiol* 2017;69:325-44.
36. San Román JA, Vilacosta I, López J, Sarriá C. Critical questions about left-sided infective endocarditis. *J Am Coll Cardiol* 2015;66:1068-76.
37. Ferrera C, Vilacosta I, Fernandez C, et al. Reassessment of blood culture-negative endocarditis: its profile is similar to that of blood culture-positive endocarditis. *Rev Esp Cardiol* 2012;65:891-900.
38. Selton-Suty C, Marie Célaré M, Moing V, et al. Preeminence of *Staphylococcus aureus* in infective endocarditis: a 1-year population-based survey. *Clin Infect Dis* 2012;54:1230-9.
39. Dahl A, Rasmussen R, Bundgaard H, et al. *Enterococcus faecalis* infective endocarditis. A pilot study of the relationship between duration of gentamicin treatment and outcome. *Circulation* 2013;127:1810-7.