

OPEN

# Prediction of Inhospital Mortality in Critically Ill Patients With Sepsis: Confirmation of the Added Value of 24-Hour Lactate to Acute Physiology and Chronic Health Evaluation IV

**BACKGROUND:** We previously reported the added value of 24-hour lactate concentration alone and in combination with 24-hour lactate clearance and lactate concentration at admission for the prediction of inhospital mortality in critically ill patients with sepsis. We aimed to validate this finding.

**DERIVATION COHORT:** The derivation cohort from Leiden, The Netherlands, consisted of 451 critically ill patients with sepsis.

**VALIDATION COHORT:** The validation cohort consisted of 4,440 critically ill adult patients with sepsis from the Medical Information Mart for Intensive Care cohort admitted to the ICU of Beth Israel Deaconess Medical Center, Boston, MA, between January 2006 and 2018.

**PREDICTION MODEL:** Predictors of mortality were: age, chronic comorbidities, length of stay pre-ICU, Glasgow Coma Scale, and Acute Physiology Score. Lactate concentration at 24-hour alone, in combination with 24-hour lactate clearance and in combination with lactate concentration at admission, was added to assess improvement of the prediction model. The outcome was inhospital mortality.

**RESULTS:** Inhospital mortality occurred in 160 patients (36%) in the derivation cohort and in 2,347 patients (53%) in the validation cohort. The Acute Physiology and Chronic Health Evaluation (APACHE) IV model had a moderate discriminative performance (recalibrated C-statistic, 0.62; 95% CI, 0.60–0.63). Addition of 24-hour lactate concentration increased the recalibrated C-statistic to 0.64 (95% CI, 0.62–0.66). The model with 24-hour lactate concentration and lactate concentration at admission showed the best fit as depicted by the smallest Akaike Information Criterion in both the derivation and validation data.

**CONCLUSION:** The 24-hour lactate concentration and lactate concentration at admission contribute modestly to prediction of inhospital mortality in critically ill patients with sepsis. Future updates and possible modification of APACHE IV should consider the incorporation of lactate concentration at baseline and at 24 hours.

**KEY WORDS:** Acute Physiology and Chronic Health Evaluation; critical care; external validation; lactic acid; prognosis; sepsis; septic shock

Sepsis is a life-threatening organ dysfunction caused by a dysregulated host response to infection (1). It is one of the leading causes of mortality and prolonged disability among critically ill patients (2, 3). Serum lactate level and lactate clearance are cornerstones in the management of critically ill patients with sepsis (4), after multiple studies showing the association between mortality and elevated lactate levels (>4 mmol/L) or lactate clearance at 6 or 24 hours (5–9). Early lactate-guided resuscitation in critically ill

Meryem Baysan, MD<sup>1-3</sup>

Mendi S. Arbous, MD, PhD<sup>1,2</sup>

Ewout W. Steyerberg, MD, PhD<sup>4</sup>

Johanna G. van der Bom,  
MD, PhD<sup>2,3</sup>

Copyright © 2022 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of the Society of Critical Care Medicine. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

DOI: 10.1097/CCE.0000000000000750

patients with sepsis showed evidence of mortality reduction in different randomized control studies (10, 11), which was confirmed in later meta-analyses (12, 13). Despite the important role of lactate in the management of critically ill patients with sepsis, lactate concentration and lactate clearance are thus far not taken into account in outcome prediction models for critically ill patients.

One of the most widely used prediction models for benchmarking in the ICU is the Acute Physiology and Chronic Health Evaluation for predicting inhospital mortality (APACHE) (14, 15). The APACHE IV is based on 142 variables collected in the first 24 hours after admission to the ICU (14). We previously reported the added value of lactate level and lactate clearance at 0, 6, and 24 hours after ICU admission compared with APACHE IV scores for mortality prediction in critically ill patients with sepsis (16). We found that lactate level at 24 hours had the highest added predictive value to predict inhospital mortality when added to the APACHE IV model. External validation in an independent patient population is an essential next step in prediction model development that should be performed before a model can be implemented in clinical practice (17).

We aimed to validate the role of the lactate concentration during the first 24 hours of ICU admission for predicting inhospital mortality in critically ill patients with sepsis admitted to ICUs. We hereto analyzed patients from in the Medical Information Mart for Intensive Care (MIMIC)-III study.

## MATERIALS AND METHODS

The Transparent Reporting of a multivariable prediction model for individual Prognosis Or Diagnosis checklist for prediction model development was used for the reporting of this study (**Supplemental Table 1**, <http://links.lww.com/CCX/B49>) (18, 19).

### Derivation Cohort

The derivation cohort consisted of 451 critically ill patients with sepsis admitted to the ICU of Leiden University Medical Center, The Netherlands, between January 2006 and 2018 (16). The institutional review board of Leiden University Medical Center approved the study in September 2017 (reference G17.094), which was conducted according to the 1964 Helsinki

declaration and its later amendments. A waiver for informed consents was granted by the same institutional review board. The critically ill patients with sepsis were identified by their APACHE IV admission diagnosis “sepsis.” Patients under 18 years old, patients without any lactate measurement during their ICU admission, and patients admitted for less than 24 hours in the ICU were excluded since the APACHE IV predicts from 24 hours after ICU admission. Only patients’ first ICU admissions were analyzed (14). Patients discharged to another ICU or admitted from another ICU were also excluded from the analysis (16). The regression coefficients of the three best performing models were used in the validation cohort (**Supplemental Table 2**, <http://links.lww.com/CCX/B49>).

### MIMIC Cohort

We were permitted access to the MIMIC database, a large, single-center database comprising information relating to patients admitted to critical care units at the Beth Israel Deaconess Medical Center, Boston, MA, between June 2001 and October 2012 for validation. Information regarding data collection and regulatory norms of the MIMIC database was published earlier (20). Summarized, the MIMIC database has been approved by the institutional review boards of Beth Deaconess Medical Center and the Massachusetts Institute of Technology in January 2001. Furthermore, a waiver for informed consent was granted by the institutional review board of Beth Deaconess Medical Center (reference 2001P001699). The data set was freely accessible after following an online human subjects training and signing a data user agreement. The MIMIC database has been last updated in 2016 (MIMIC-III v1.4) and has deidentified information of patients admitted to the ICU (20). The same in- and exclusion criteria were applied to the validation cohort as in the derivation cohort.

### Outcome Definition

The primary outcome was inhospital mortality, which was defined as mortality in the ICU or in another ward during the same hospital admission. This was in accordance with the inhospital definition of the MIMIC-III database and the derivation cohort. With 2,414 cases of inhospital mortality (events), the validation cohort provided ample statistical power for validation,

where the advice is to have at least 100 events and 100 nonevents (21).

## Patients and Predictors

We downloaded data from the MIMIC-III database in August 2019. We aimed to include all critically ill patients with the admission diagnosis sepsis. However, the MIMIC-III database did not contain the variable “APACHE IV admission diagnosis.” As an alternative to that variable, we identified patients with an admission diagnosis sepsis by using a combination of the variables for the *International Classification of Disease*, 9th Revision (ICD-9) codes and for the sepsis-3 criteria in the following way (22). We selected patients with the following.

One of the ICD-9 codes explicitly mentioning sepsis: 995.92 (severe sepsis) or 785.52 (septic shock); and all patients fulfilling both of the following two criteria as suggested by Seymour et al (23):

- 1) “The combination of giving antibiotics and body fluid cultures (blood, urine, cerebrospinal fluid, etc.); if the antibiotic was given first, the culture sampling had to be obtained within 24 hours; if the culture sampling was first, the antibiotic had to be within 72 hours. The onset of infection was defined as the time at which the first of these two events occurred” (23).
- 2) Organ dysfunction defined as having a Sequential Organ Functioning Assessment score of at least two points (24).

Information on the source of sepsis was not available in the MIMIC-III database. It was not possible to reliably approximate these data. Other downloaded data comprised: age, sex, weight, length, length of ICU stay and pre-ICU length of stay in days, Glasgow Coma Scale, Acute Physiology Score (APS), chronic comorbidities, inhospital mortality, and lactate value at admission to the ICU and 24 hours after ICU admission (14). Pre-ICU length of stay was calculated as the difference in days between ICU admission time and hospital admission time. A pragmatic approximation of chronic health conditions was made using ICD-9, as depicted in **Supplemental Table 3** (<http://links.lww.com/CCX/B49>). For the Glasgow Coma Scale, we used the worst total score during the first day of ICU admission. The APS was calculated as the sum of weights of the worst values during the first ICU day for pulse rate, mean arterial pressure, temperature, respiratory rate, arterial oxygen tension conditional on mechanical ventilation and  $F_{iO_2}$ , hematocrit, WBC count, serum creatinine

conditional on acute renal failure (defined as urine output <410 mL/d or chronic dialysis), urine output, blood urea nitrogen, sodium, albumin, bilirubin, glucose, and acid base abnormalities. Worst values were defined as in the original APACHE model, in which the extremes of the values were assigned additional weights and normal values were assigned a normal weight (**Supplemental Table 4**, <http://links.lww.com/CCX/B49>) (25).

## Handling of Missing Data

We defined outliers in each variable as extreme non-physiologic values after assessment of biological plausibility (**Supplemental Table 5**, <http://links.lww.com/CCX/B49>). Outliers were set as missing values. We followed a multiple imputation procedure to handle missing data among physiologic variables needed to calculate the APS. We assumed that the data were missing at random. Multiple imputation was done with the R package “Multiple Imputation by Chained Equation” (26, 27), using the predictive mean matching method. The variable lactate clearance at 24 hours was imputed using passive imputation. To improve the efficacy of the imputation model, we included auxiliary variables that were expected to be correlated with the incomplete variables (**Supplemental Table 5**, <http://links.lww.com/CCX/B49>) (26). This led to a more effective imputation procedure as more variability was explained. Since at most 15% of values were missing (**Supplemental Fig. 1**, <http://links.lww.com/CCX/B49>), we imputed 15 datasets (28) and took into account both the within- and between-imputation variabilities for the analysis by applying Rubin rules to pool the estimates and variances (29, 30). Furthermore, a sensitivity analysis was performed, using a single imputation with the last-value-carried-forward method (**Supplemental Fig. 2**, <http://links.lww.com/CCX/B49>).

## Statistical Analysis

We first calculated predicted risks using the regression coefficients from the derivation cohort (**Supplemental Table 2**, <http://links.lww.com/CCX/B49>) in the validation cohort. The underlying logistic regression model had the following predictors: age, pre-ICU length of stay, Glasgow Coma Scale, APS, and chronic comorbidities. In this simplified model, linear terms were used instead of cubic spline terms. Due to the lack of

reliable data regarding sepsis source and admission type (elective surgery, urgent surgery, or medical) in the validation cohort, these predictors and their regression coefficients were not included in the logistic regression models for the validation cohort. A total of four models from the derivation cohort were applied to our validation cohort, as depicted in **Supplemental Table 6** (<http://links.lww.com/CCX/B49>). To accommodate for differences between cohorts regarding calendar time and geographical location, we recalibrated the intercepts and slopes of the models in our validation cohort. Recalibration thereby increased generalizability of the model predictions and minimized miscalibration (31).

The performance of the models in the validation cohort was assessed according to model discrimination and calibration. Prediction model assessments were performed using the *val.prob.ci.2*-function in R (30, 32). Calibration-in-the-large and calibration slopes were assessed for each model. The Loess algorithm was used to smoothen the calibration plots. Calibration was defined as the agreement between observed outcomes and predicted outcomes. Calibration-in-the-large is a comparison of the mean of all predicted risk with the mean observed risk and should ideally be 0. It was calculated as the intercept in a logistic regression model with the log odds of the predicted risk as the only predictor. Regression slopes in each of these regression models provided the calibration slopes, which should ideally be 1 if observed and predicted risks graphically follow a 45° line (33). Discrimination was defined as differentiation of those with inhospital mortality from those without inhospital mortality. It was assessed for each model with the C-statistic, which is identical to the area under the receiver operating characteristic curve (AUROC).

Improvement in model performances in comparison to the APACHE IV model was assessed using the difference in C-statistic with a *p* value calculated with the likelihood ratio statistic. Akaike Information Criterion (AIC) was used to estimate the relative information loss in each model, as stated in our previous study (16). Furthermore, for more insight into the prediction performances, we reported the Net Reclassification Improvement (NRI) of the APACHE IV model with 24-hour lactate concentration compared with the APACHE IV model separately for events (mortality) and nonevents (34, 35). The NRI for events can be interpreted as change in sensitivity,

whereas the NRI for nonevents can be interpreted as the change in specificity (16, 36). The fraction of patients who were correctly reclassified was calculated by using formula 1:

$$\text{Fraction correctly reclassified} = (\text{Event rate} \times \text{NRI}_{(\text{event})}) + (\text{Nonevent rate} \times \text{NRI}_{(\text{nonevent})})$$

All analyses were performed in R (R foundation for Statistical Computing, Vienna, Austria) (37).

## RESULTS

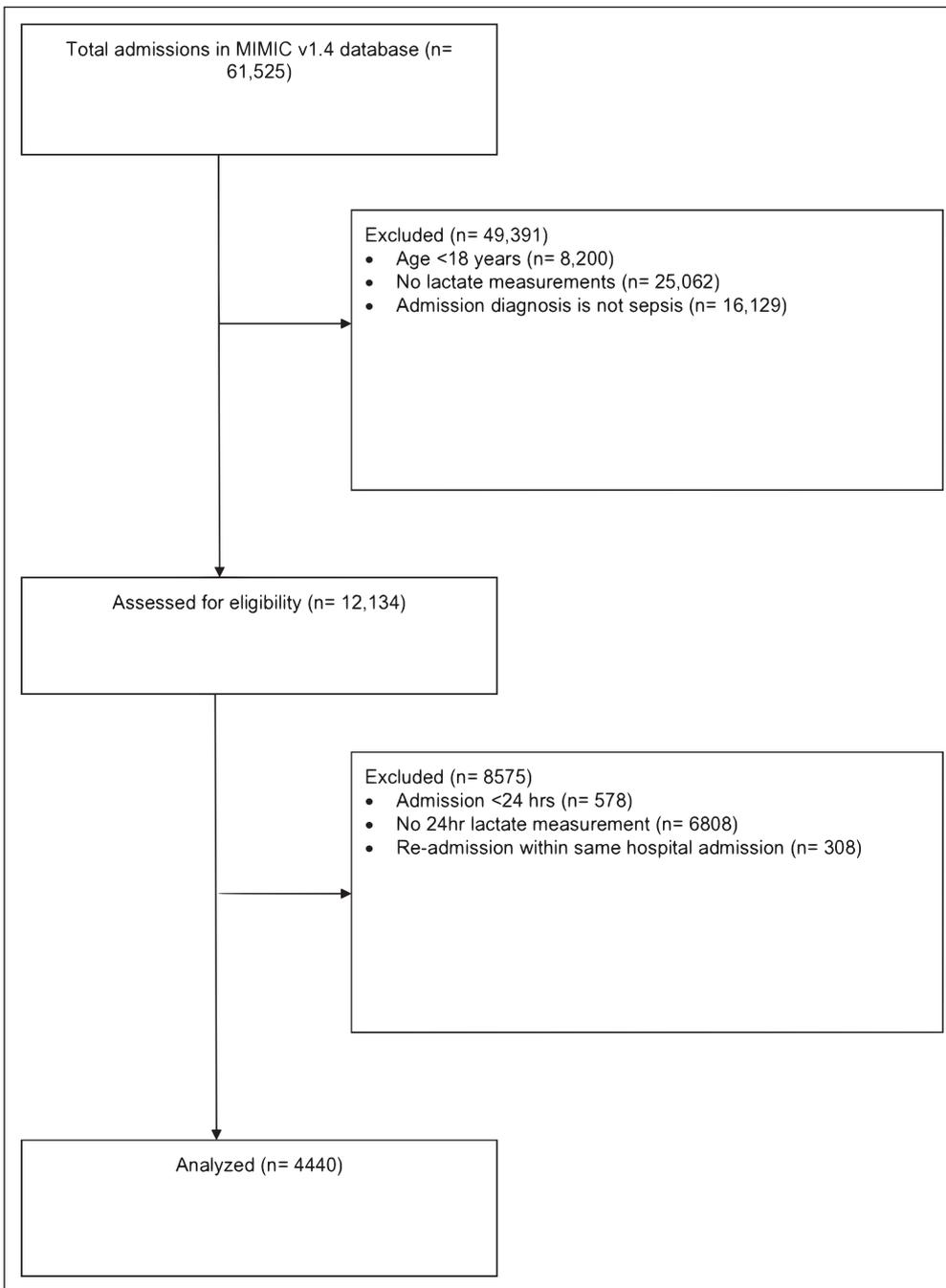
### Validation Cohort Characteristics

A total of 12,134 ICU admissions with admission diagnosis sepsis were found in the MIMIC database. We excluded admissions with a length of stay less than 24 hours at the ICU ( $n = 578$ ) and without a known lactate concentration at 24 hours ( $n = 6,808$ ). Only the first ICU admissions during the same hospital stay were included ( $n = 308$  readmissions excluded). As shown in **Figure 1**, eventually, 4,440 admissions were included in the analysis. The median age was 65 years (interquartile range [IQR], 53–76 yr), and most patients were male (57%) (**Table 1**). Median length of ICU stay was 5.5 days (IQR, 3.1–10.9 d). A total of 2,347 patients died (53%) (Table 1).

Lactate concentration at admission and after 24 hours was similar in the validation cohort and the derivation cohort (**Supplemental Table 7**, <http://links.lww.com/CCX/B49>). However, the lactate clearance fractions were 25% (IQR, –9.2 to 49.3%) and 18% (IQR, –17.0 to 46.0%), respectively.

### Predictive Performance

The C-statistic of the APACHE IV model was 0.71 (95% CI, 0.70–0.73), whereas it was 0.73 (95% CI, 0.71–0.74) in the APACHE IV model with 24-hour lactate concentration, 0.73 (95% CI, 0.71–0.74) in the APACHE IV model with 24-hour lactate concentration and baseline lactate concentration, and 0.73 (95% CI, 0.71–0.74) in the APACHE IV model with 24-hour lactate concentration and 24-hour lactate clearance in our validation cohort (**Fig. 2A**; and **Supplemental Table 8**, <http://links.lww.com/CCX/B49>), whereas the recalibrated C-statistic was, respectively, 0.62 (95% CI, 0.60–0.63), 0.64 (95% CI, 0.62–0.65), 0.64 (95% CI, 0.62–0.65), and 0.64 (95% CI, 0.62–0.66). The calibration slope was 1.02



**Figure 1.** Flow diagram of cohort selection based on ICU admissions. MIMIC = Medical Information Mart for Intensive Care.

(95% CI, 0.93–1.11) in the APACHE IV model and in the APACHE IV model with 24-hour lactate concentration, and 1.02 (95% CI, 0.93–1.10) in the APACHE IV model with 24-hour lactate concentration and baseline lactate concentration, and the slope was 1.02 (95% CI, 0.93–1.11) in the APACHE IV model with 24-hour lactate concentration and 24-hour lactate clearance. The recalibrated calibration slope was 1.00 (95% CI, 0.83–1.17) in the APACHE IV model, the slope was

1.00 (95% CI, 0.86–1.15) in the APACHE IV model with either 24-hour lactate concentration alone or in combination with baseline lactate concentration and was 1.00 (95% CI, 0.86–1.14) in the APACHE IV model with 24-hour lactate concentration and 24-hour lactate clearance. Furthermore, the unfitted and refitted calibration-in-the-large performances resulted in an intercept of 0.00 (95% CI, –0.06 to 0.06). Summary of model performances of both the validation and derivation cohort can be found in **Table 2**.

The AIC difference of the APACHE IV model with the APACHE IV model that included 24-hour lactate concentration was 111.7 (**Table 3**). Similarly, the AIC difference of the APACHE IV model with 24-hour lactate concentration and 24-hour lactate clearance was 109.8, whereas it was 112.9 in the APACHE IV model with 24-hour lactate concentration and lactate concentration at admission. The  $R^2$  in the APACHE IV model ( $R^2 = 0.18$ ) increased to 0.21 in the APACHE IV models

with lactate concentration at 24-hour alone, in combination with 24-hour lactate clearance, and in combination with lactate concentration at admission (**Table 2**).

NRI quantifies how well the model with 24-hour lactate concentration correctly reclassifies subjects compared with the APACHE IV without lactate concentration. Among the patients who died ( $n = 2,347$ ), 44% were correctly reclassified, whereas 56% were incorrectly reclassified

**TABLE 1.**

**Characteristics of Patients With Sepsis Admitted to the ICU at Leiden University Medical Center (2006–2018) of the Validation and Derivation Cohort**

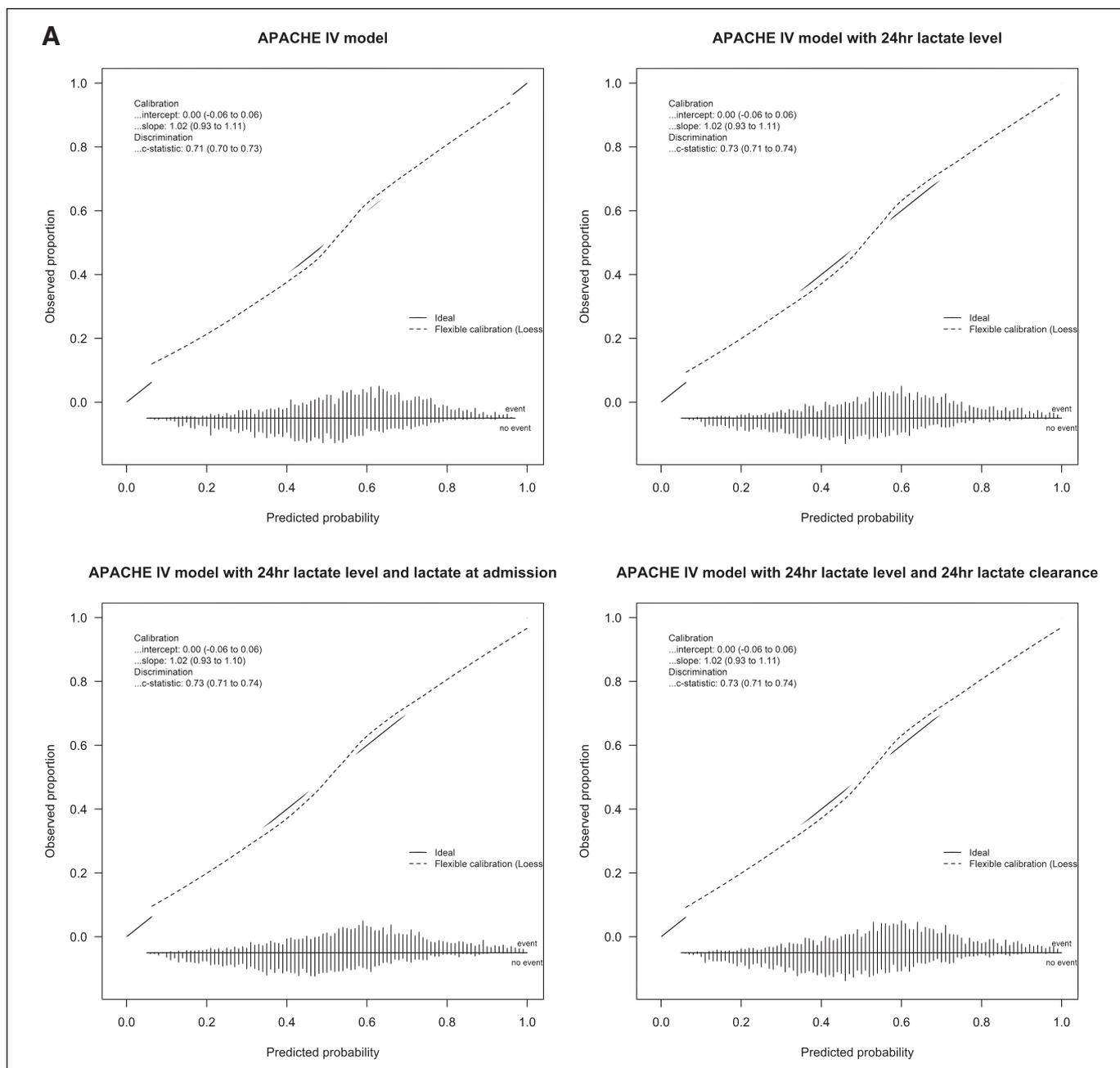
Characteristics	Medical Information Mart for Intensive Care-III (Validation Cohort) ( <i>n</i> = 4,440)	Leiden (Derivation Cohort) ( <i>n</i> = 451)
Male sex, absolute number (%)	2548 (57.4)	291 (64.5)
Age, median (IQR)	65.0 (53.0–76.0)	66.0 (56.0–74.0)
Missing	203 (4.6%)	0 (0%)
Body mass index, median (IQR)	27.5 (23.9–79.9)	24.8 (22.3–27.8)
Missing, absolute number (%)	646 (14.5)	283 (3.4)
Pre-ICU LOS hr, median (IQR)	0 (0.0–22.3)	9.6 (2.4–165.6)
Missing, absolute number (%)	1 (0.0)	
ICU LOS days, median (IQR)	5.50 (3.1–10.9)	
Chronic comorbidity, absolute number (%)		
No comorbidity	1311 (29.5)	179 (39.7)
Cardiorespiratory	1332 (30.0)	19 (4.2)
Renal	628 (14.1)	23 (5.1)
Cirrhosis	301 (6.8)	14 (3.1)
Metastatic neoplasm	149 (3.4)	23 (5.1)
Hematologic malignancy	127 (2.9)	26 (5.8)
Immunologic insufficiency	93 (2.1)	88 (19.5)
Thrombolysis	17 (0.4)	0 (0)
Diabetes mellitus	482 (10.9)	79 (17.5)
Mechanically ventilated in the first 24 hr, absolute number (%)	3416 (76.9)	201 (44.6)
FiO <sub>2</sub> in %		
Median (IQR)	60.0 (50.0–100.0)	32.0 (28.0–40.0)
Missing, absolute number (%)	2235 (50.3)	70 (15.5)
Glasgow Coma Scale, absolute number (%)		
Severe (3–8)	499 (11.2)	34 (7.5)
Moderate (9–12)	396 (8.9)	26 (5.8)
Mild (13–15)	3498 (78.8)	391 (86.7)
Missing	47 (1.1)	0 (0)
Mortality	2,347 (52.9)	160 (35.5)

IQR = interquartile range, LOS = length of stay.

(Supplemental Fig. 3, <http://links.lww.com/CCX/B49>). Among patients who survived without in-hospital mortality (*n* = 2,093), 74% were correctly reclassified, whereas 26% were incorrectly reclassified. This resulted in an NRI for events of –11% (95% CI, –15.1 to –7.1) and an NRI for nonevents of 48% (95% CI, 44.5–52.0). The fraction of patients that were correctly reclassified was, therefore, 17%.

## DISCUSSION

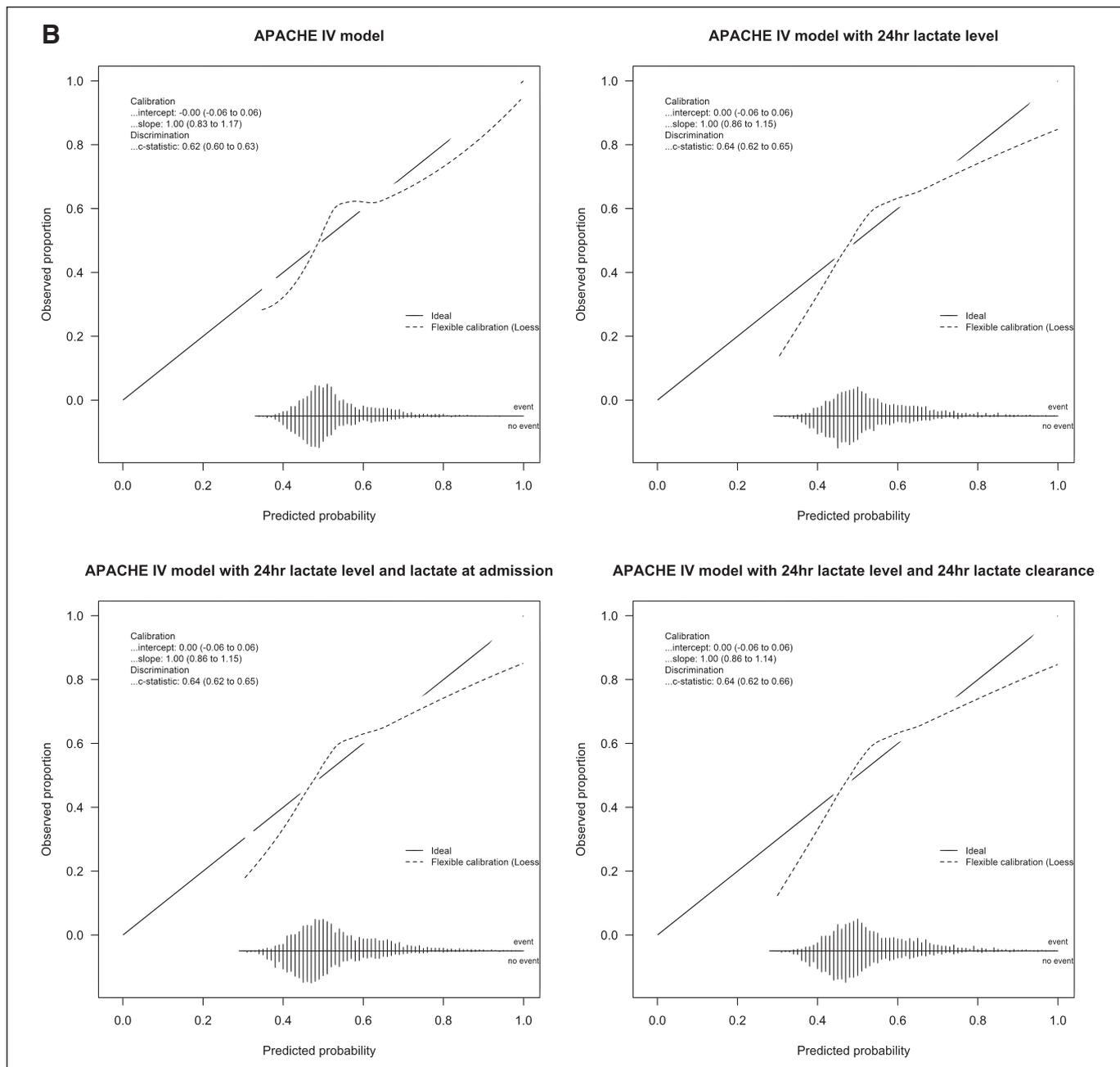
The aim of this study was to validate the added predictive value of lactate to the APACHE IV model in predicting in-hospital mortality in critically ill patients with sepsis, which our earlier study had indicated. The results of the present study confirm that lactate concentration measured 24 hours after ICU admission



**Figure 2.** Overview of calibration curves of the 4 models in the validation cohort after recalibration of intercept and slope. **A**, The calibration curves of validated models before recalibration.

modestly improved the predictive performance beyond that of the APACHE IV model predicting in-hospital mortality among critically ill patients with sepsis. Similarly, addition of lactate clearance at 24 hours or lactate concentration at admission modestly increased the predictive power beyond that of the APACHE IV model. The addition of both lactate concentration at admission and 24-hour lactate had, relatively, the strongest effect on the predictive performance of APACHE IV in critically ill patients with sepsis.

The discriminative performance of APACHE IV in the MIMIC cohort was much lower than that in our previously published derivation cohort and in other cohorts. Previous validation studies of APACHE IV in different countries (India, Brazil, The Netherlands, and Malaysia) showed overall good discrimination (AUROC between 0.78 and 0.89) (38–40). To the best of knowledge, no other study validated APACHE IV (neither with nor without additional variables) in the publicly available MIMIC database. The relatively low



**Figure 2. (Continued). B,** The calibration curves of validated models after recalibration are illustrated. APACHE IV = Acute Physiology and Chronic Health Evaluation.

discriminative performances of APACHE IV may for a large part be explained by insufficient information on two important APACHE IV predictors in the MIMIC database: admission type and sepsis source.

Missing data may also have influenced the findings. A larger proportion of missing data were seen in the MIMIC cohort compared with the derivation cohort (16). More than 40% missing in  $FiO_2$ , albumin, and in multiple auxiliary variables could be seen, possibly affecting the regression coefficients (41). However, a

little to no change in model performances in both the recalibrated models and nonrecalibrated models was seen in the sensitivity analyses, confirming robustness of the data (41, 42).

Additionally, the low discriminative performances may be explained by differences in case-mix, definitions, and standard of care (43, 44). The identification of critically ill patients with sepsis could not be done with the APACHE IV definition of sepsis, since this was not available in the MIMIC

**TABLE 2.**  
**Performances of the Models in the Recalibrated Validation and Derivation Cohort (16)**

Model	Medical Information Mart for Intensive Care (Validation Cohort)				Derivation Cohort	
	C-Statistic (95% CI)	Intercept	Regression Slope (95% CI)	R <sup>2</sup>	C-Statistic	R <sup>2</sup>
Acute Physiology and Chronic Health Evaluation IV	0.62 (0.60–0.63)	−1.34e-13	1.00 (0.83–1.17)	0.05	0.78	0.28
+ Lactate 24 hr	0.64 (0.62–0.65)	1.40e-11	1.00 (0.86–1.15)	0.07	0.79	0.30
+ Lactate 24 hr and lactate at baseline	0.64 (0.62–0.65)	1.10e-11	1.00 (0.86–1.15)	0.07	0.79	0.30
+ Lactate 24 hr and 24-hr lactate clearance	0.64 (0.62–0.66)	1.47e-11	1.00 (0.86–1.14)	0.07	0.79	0.30

**TABLE 3.**  
**Improvement in Model Performances After Adding Various Measures of Lactate Concentration During the First 24 hr After Admission to the Intensive Care Quantified as Difference in Akaike Information Criteria Compared With the Acute Physiology and Chronic Health Evaluation IV Model in the Validation and the Derivation Cohorts (16)**

Model	Medical Information Mart for Intensive Care (Validation Cohort)				Derivation Cohort		
	Observations	df	AIC	ΔAIC	Observations	df	ΔAIC
Acute Physiology and Chronic Health Evaluation IV	4,440	20	5,549.7	Reference	451	26	Reference
+ Lactate 24 hr	4,440	21	5,438.0	111.68	451	27	7.71
+ Lactate at baseline and lactate 24 hr	4,440	22	5,436.7	112.94	451	28	7.16
+ Lactate 24 hr and 24-hr lactate clearance	4,440	22	5,439.8	109.81	451	28	7.32

ΔAIC = Akaike Information Criterion difference from the Acute Physiology and Chronic Health Evaluation IV model, df = degree of freedom for the chi-squared distribution.

database. We, therefore, identified the present cohort using the sepsis-3 criteria in combination with the ICD-9 codes. The resulting difference in case-mix could have led to differences in performances between the prediction models of the derivation and validation cohorts (43). Furthermore, the definitions of chronic comorbidities and admission diagnosis in our derivation cohort accorded with the APACHE IV definitions (14, 16). In our analysis in MIMIC, we pragmatically defined these predictors using the Elixhauser coding algorithm and criteria as previously reported (22). The differences in the

definition and calculation of predictors between derivation and validation cohort could eventually influence model performances (45, 46).

Despite the observed low predictive performance of APACHE IV in the current analysis on MIMIC data, the present findings corroborate our previous observation that lactate adds predictive performance to APACHE IV. The currently observed added predictive value of lactate at admission and at 24 hours was similar to the added predictive value observed in the derivation cohort. We, therefore, infer that a next version of APACHE might profit from including lactate values to improve

the prediction of mortality among critically ill patients with sepsis. In clinical practice lactate is already regularly monitored in patients with sepsis. It is generally known that lactate and lactate clearance 24 hours after admission to the ICU are associated with mortality in critically ill patients with sepsis (5, 10). The latest Surviving Sepsis Campaign recommends lactate monitoring to support clinical management (4). Thus, the present findings are biologically and clinically plausible. Yet, before implementation in clinical practice, the models to predict mortality among critically ill patients with sepsis need to be refined and validated in critically ill patients with sepsis in other ICUs and in more recent calendar times, since both the validation and derivation study were rather historical cohorts, whereas the management and diagnosis of sepsis have changed in recent years and might influence the outcome distribution (1, 4, 47).

## CONCLUSIONS

This validation study confirms a modest but nonnegligible added value of 24-hour lactate in predicting inhospital mortality in critically ill patients with sepsis, especially if both lactate concentration at admission and 24-hour lactate concentration are added to the APACHE IV model. Future updates of APACHE IV should consider incorporating lactate at baseline and at 24 hours as predictors.

## ACKNOWLEDGMENTS

We thank PhysioNet for the access and support of the MIMIC database. Furthermore, we also thank Aad Pors with the data management of this study. The authors also thank for the helpful comments from the Center for Clinical Transfusion Research Integrity Program's Scientific committee on the statistical analysis plan and final article.

1 Department of Intensive Care, Leiden University Medical Center, Leiden, The Netherlands.

2 Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands.

3 Center for Clinical Transfusion Research, Sanquin Research, Leiden, The Netherlands.

4 Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands.

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Drs. Baysan, Arbous, and van der Bom designed the study. Dr. Baysan performed the analysis. All authors interpreted the data. The first draft of the article was written by Dr. Baysan, and all authors commented on previous versions of the article. All authors read and approved the final article.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (<http://journals.lww.com/ccejournal>).

Supported, in part, by Grant PPOC-16-31 by Sanquin Blood Supply, Amsterdam, The Netherlands.

The authors have disclosed that they do not have any potential conflicts of interest.

For information regarding this article, E-mail: [j.g.van.der.bom@lumc.nl](mailto:j.g.van.der.bom@lumc.nl)

## REFERENCES

1. Singer M, Deutschman CS, Seymour CW, et al: The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* 2016; 315:801–810
2. Fleischmann C, Scherag A, Adhikari NK, et al; International Forum of Acute Care Trialists: Assessment of global incidence and mortality of hospital-treated sepsis. Current estimates and limitations. *Am J Respir Crit Care Med* 2016; 193:259–272
3. Iwashyna TJ, Ely EW, Smith DM, et al: Long-term cognitive impairment and functional disability among survivors of severe sepsis. *JAMA* 2010; 304:1787–1794
4. Rhodes A, Evans LE, Alhazzani W, et al: Surviving sepsis campaign: International guidelines for management of sepsis and septic shock: 2016. *Intensive Care Med* 2017; 43:304–377
5. Masyuk M, Wernly B, Lichtenauer M, et al: Prognostic relevance of serum lactate kinetics in critically ill patients. *Intensive Care Med* 2019; 45:55–61
6. Vincent JL, Quintairo E, Silva A, Couto L Jr, et al: The value of blood lactate kinetics in critically ill patients: A systematic review. *Crit Care* 2016; 20:257
7. Shapiro NI, Howell MD, Talmor D, et al: Serum lactate as a predictor of mortality in emergency department patients with infection. *Ann Emerg Med* 2005; 45:524–528
8. Vink EE, Bakker J: Practical use of lactate levels in the intensive care. *J Intensive Care Med* 2018; 33:159–165
9. Nguyen HB, Rivers EP, Knoblich BP, et al: Early lactate clearance is associated with improved outcome in severe sepsis and septic shock. *Crit Care Med* 2004; 32:1637–1642
10. Jansen TC, van Bommel J, Schoonderbeek FJ, et al; LACTATE study group: Early lactate-guided therapy in intensive care unit patients: A multicenter, open-label, randomized controlled trial. *Am J Respir Crit Care Med* 2010; 182:752–761
11. Lyu X, Xu Q, Cai G, et al: [Efficacies of fluid resuscitation as guided by lactate clearance rate and central venous oxygen saturation in patients with septic shock]. *Zhonghua Yi Xue Za Zhi* 2015; 95:496–500
12. Gu WJ, Zhang Z, Bakker J: Early lactate clearance-guided therapy in patients with sepsis: A meta-analysis with trial sequential analysis of randomized controlled trials. *Intensive Care Med* 2015; 41:1862–1863

13. Simpson SQ, Gaines M, Hussein Y, et al: Early goal-directed therapy for severe sepsis and septic shock: A living systematic review. *J Crit Care* 2016; 36:43–48
14. Zimmerman JE, Kramer AA, McNair DS, et al: Acute Physiology and Chronic Health Evaluation (APACHE) IV: Hospital mortality assessment for today's critically ill patients. *Crit Care Med* 2006; 34:1297–1310
15. Knaus WA, Zimmerman JE, Wagner DP, et al: APACHE-acute physiology and chronic health evaluation: A physiologically based classification system. *Crit Care Med* 1981; 9:591–597
16. Baysan M, Baroni GD, van Boekel AM, et al: The added value of lactate and lactate clearance in prediction of in-hospital mortality in critically ill patients with sepsis. *Crit Care Explor* 2020; 2:e0087
17. Altman DG, Vergouwe Y, Royston P, et al: Prognosis and prognostic research: Validating a prognostic model. *BMJ* 2009; 338:b605
18. Collins GS, Reitsma JB, Altman DG, et al: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD statement. *Ann Intern Med* 2015; 162:55–63
19. Moons KG, Altman DG, Reitsma JB, et al: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med* 2015; 162:W1–73
20. Johnson AE, Pollard TJ, Shen L, et al: MIMIC-III, a freely accessible critical care database. *Sci Data* 2016; 3:160035
21. Peek N, Arts DG, Bosman RJ, et al: External validation of prognostic models for critically ill patients required substantial sample sizes. *J Clin Epidemiol* 2007; 60:491–501
22. Johnson AEW, Aboab J, Raffa JD, et al: A comparative analysis of sepsis identification methods in an electronic database. *Crit Care Med* 2018; 46:494–499
23. Seymour CW, Liu VX, Iwashyna TJ, et al: Assessment of clinical criteria for sepsis: For the third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* 2016; 315:762–774
24. Vincent JL, Moreno R, Takala J, et al: The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med* 1996; 22:707–710
25. Knaus WA, Wagner DP, Draper EA, et al: The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991; 100:1619–1636
26. van Buuren S: Flexible Imputation of Missing Data. Second Edition Vancouver, Chapman & Hall/CRC Interdisciplinary Statistics. CRC Press LLC, 2018
27. van Buuren S, Groothuis-Oudshoorn K: Mice: Multivariate imputation by chained equations in R. *J Stat Soft* 2011; 45:1–67
28. Graham JW, Olchowski AE, Gilreath TD: How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prev Sci* 2007; 8:206–213
29. Marshall A, Altman DG, Holder RL, et al: Combining estimates of interest in prognostic modelling studies after multiple imputation: Current practice and guidelines. *BMC Med Res Methodol* 2009; 9:57
30. Rubin DB, Schenker N: Multiple imputation in health-care databases: An overview and some applications. *Stat Med* 1991; 10:585–598
31. Moons KG, Kengne AP, Grobbee DE, et al: Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* 2012; 98:691–698
32. Van Calster B, Nieboer D, Vergouwe Y, et al: A calibration hierarchy for risk models was defined: From utopia to empirical data. *J Clin Epidemiol* 2016; 74:167–176
33. Steyerberg EW, Vergouwe Y: Towards better clinical prediction models: Seven steps for development and an ABCD for validation. *Eur Heart J* 2014; 35:1925–1931
34. Leening MJ, Vedder MM, Witteman JC, et al: Net reclassification improvement: Computation, interpretation, and controversies: A literature review and clinician's guide. *Ann Intern Med* 2014; 160:122–131
35. Pencina MJ, D'Agostino RB, Pencina KM, et al: Interpreting incremental value of markers added to risk prediction models. *Am J Epidemiol* 2012; 176:473–481
36. Steyerberg EW: Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. 2nd Ed. New York, NY, Springer Science and Business Media, 2019
37. R Core Team: R: A Language and Environment for Statistical Computing. Vienna, Austria, R Foundation for Statistical Computing, 2019
38. Brinkman S, Bakhshi-Raiez F, Abu-Hanna A, et al: External validation of acute physiology and chronic health evaluation IV in dutch intensive care units and comparison with acute physiology and chronic health evaluation II and simplified acute physiology score II. *J Crit Care* 2011; 26:105.e11–105.e18
39. Wong RS, Ismail NA, Tan CC: An external independent validation of APACHE IV in a Malaysian intensive care unit. *Ann Acad Med Singap* 2015; 44:127–132
40. Salluh JI, Soares M: ICU severity of illness scores: APACHE, SAPS and MPM. *Curr Opin Crit Care* 2014; 20:557–565
41. Thabane L, Mbuagbaw L, Zhang S, et al: A tutorial on sensitivity analyses in clinical trials: The what, why, when and how. *BMC Med Res Methodol* 2013; 13:92
42. Marshall A, Altman DG, Royston P, et al: Comparison of techniques for handling missing covariate data within prognostic modelling studies: A simulation study. *BMC Med Res Methodol* 2010; 10:1–16
43. Riley RD, Ensor J, Snell KI, et al: External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: Opportunities and challenges. *BMJ* 2016; 353:i3140
44. Debray TP, Vergouwe Y, Koffijberg H, et al: A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J Clin Epidemiol* 2015; 68:279–289
45. Luijken K, Wynants L, van Smeden M, et al; Collaborators: Changing predictor measurement procedures affected the performance of prediction models in clinical examples. *J Clin Epidemiol* 2020; 119:7–18
46. Ramspek CL, Jager KJ, Dekker FW, et al: External validation of prognostic models: what, why, how, when and where? *Clin Kidney J* 2021; 14:49–58
47. Dellinger RP, Levy MM, Rhodes A, et al; Surviving Sepsis Campaign Guidelines Committee including The Pediatric Subgroup: Surviving Sepsis Campaign: International guidelines for management of severe sepsis and septic shock, 2012. *Intensive Care Med* 2013; 39:165–228