

Peer Review File

Article Information: <https://dx.doi.org/10.21037/atm-22-1737>

First Round of Peer Review

Reviewer A:

Comment 1: Based on the data presented in Table 1, patients who died were sicker, as indicated by the higher severity of illness scores, higher ALP and AST levels, higher proportion requiring invasive mechanical ventilation, etc. The higher bilirubin levels could simply be a natural consequence of patients being sicker and having more severe organ dysfunction – rather than it being an “independent predictor of mortality in ARDS” (line 247).

At a minimum, the Introduction and the Discussion sections should be edited to more carefully delineate this relationship. In order to better justify the conclusion that bilirubin is a predictor of mortality in ARDS, a more robust mechanistic explanation is needed – for example, consider expanding the second-to-last paragraph of the Discussion (lines 275-289). In addition, the authors may want to alter their modeling to control for the severity of illness – please see below.

Reply 1: We sincerely thank the reviewer for this comment. We agree with the reviewer. To better justify the conclusion that bilirubin is a predictor of mortality in ARDS, we have delineated this relationship in the **Introduction** section and give a more robust mechanistic explanation in the **Discussion** Section as follows:

Changes in the text:

1. Line 93-100 in the Introduction Section:

“Liver-lung interactions have been well documented previously, suggesting that hepatic dysfunction is a relevant clinical condition that affects the development and progression of ARDS (23, 24). A British prospective cohort study found that elevated serum TBIL levels on ICU admission correlated with ARDS development in sepsis (25), and admission TBIL was independently correlated with 60-day ARDS mortality in sepsis-related patients (26). Similarly, a French retrospective study found that TBIL in the initial phase of ARDS was associated with 90-day mortality (27).”

2. Line 299-318 in the Discussion Section:

“In support of our results, a recent study found that higher circulating cell-free hemoglobin (CFH), one of the precursors of bilirubin, is an independent risk factor for acute kidney injury in ARDS patients (37). Similarly, Shaver et al. reported that CFH contributed to ARDS severity by enhancing lung permeability and inflammation in an experimental ARDS mouse model (38). Additionally, in the context of inflammatory diseases such as acute lung injury, excess heme, or heme released in certain pathophysiological contexts may have adverse effects, partly through mechanisms of vascular endothelial dysfunction and activation of programmed cell death pathways (39, 40). The functions of bilirubin are considered something akin to a “Janus face” (11). It is reported that mildly elevated levels of TBIL have potent antioxidant and other positive benefits (12). However, hyperbilirubinemia exerts a detrimental effect on organs and causes severe and irreversible damage, especially in ARDS (19, 23). Several mechanisms may be involved in this damaging effect. First, available data have suggested that elevated bilirubin results in oxidative stress and the activation of local inflammatory responses in lung tissues, including the infiltration of alveolar macrophages, neutrophils, and the release of cytokines (IL-1, IL-6, and tumor necrosis factor [TNF]- α), which are the pathophysiological features of ARDS (23, 41, 42). Second, in vitro and in vivo studies have shown that bilirubin directly contributes to alveolar epithelial cell injury (43), leading to the interruption of cell cycles and cell apoptosis (16, 44, 45).”

Comment 2: There are some issues and questions arising from the variable selection and modeling decisions related to the multivariable logistic regression models.

First, did Model 3 adjust for all the listed covariates IN ADDITION TO all the covariates in Models 1 & 2 (that is, did Model 3 adjust for a total of 18 covariates other than bilirubin)? If so, that seems like a lot of variables to be included in a model. Can you explain the reasoning for including all those variables? Were model fit statistics or other decision-making processes used to determine if adding all these variables was necessary and/or beneficial?

In addition, logistic regression assumes that the relationship between a continuous variable and the outcome is unidirectional (the higher the value, the more likely

the outcome) – however, this is not true for some of the physiologic parameters included in Model 3. For example, both high AND low extremes of temperature, heart rate, and glucose levels would result in higher odds of mortality physiologically. Were these variables treated as simple continuous variables (which is inappropriate), or categorized (something like low, normal, and high categories with the normal being the reference group)?

In order to address the covariate/modeling issues as well as my above comments from #1, the authors may want to use one of the severity of illness scores as a covariate in Model 3, rather than adjusting for all the physiologic parameters individually. This will reduce the total number of covariates to a more reasonable number, address the issues related to some the continuous physiologic parameters (since severity of illness scores will assign higher scores to both low and high extremes), and potentially support the conclusion that bilirubin is independently associated with mortality, if the association between bilirubin and mortality remains significant after controlling for the severity of illness.

Reply 2: We sincerely thank the reviewer for the suggestion. We agree with the reviewer and have revised Model 3. Briefly, the Oxford acute severity of illness score (OASIS), containing 10 parameters^[1], was adjusted for Model 3. Sentences in the **Method** section, **Results** section, and **Table 3** were revised correspondingly as follows:

Changes in the text:

1. Line 160-164 in the Method Section:

“Besides the unadjusted model, potential covariates were progressively adjusted in three models: Model 1 was adjusted for age and gender; Model 2 was additionally adjusted for pneumonia, hypertension, COPD, renal failure, cancer, and diabetes mellitus; and Model 3 was further adjusted for admission type, SpO₂, PEEP, ALP, ALT, AST, serum glucose, and OASIS score.”

2. Line 202-209 in the Result Section:

“After adjustment for multiple covariates, the association remained significant in the multivariate logistic regression models, indicating that each 1 mg/dL increase resulted in a 4% increase in 30-day mortality (Model 3: OR = 1.04; 95% CI: 1.01 to 1.08). When the TBIL levels were treated as categorical variables, patients with TBIL levels ≥ 2 mg/dL had a higher risk of 30-day mortality in the main model (OR = 1.51; 95% CI: 1.02 to 2.22), while those with TBIL levels between 1.2 and 2 mg/dL showed a

statistically significant difference according to the lowest TBIL levels in the main model (OR = 1.30; 95% CI: 0.85 to 1.99).”

“Table 1. Relationship between total serum bilirubin and outcomes in different models

Outcomes	Non-adjusted	Model 1	Model 2	Model 3
30-day mortality	ICU OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
TBIL per 1 mg/dL	1.04 (1.01, 1.06)*	1.07 (1.04, 1.11)*	1.07 (1.03, 1.11)*	1.04 (1.01, 1.08)*
TBIL category				
<1.2 mg/dL	Reference	Reference	Reference	Reference
≥1.2, <2 mg/dL	1.14 (0.77, 1.70)	1.14 (0.77, 1.70)	1.13 (0.75, 1.69)	1.30 (0.85, 1.99)
≥2 mg/dL	1.45 (1.03, 2.03)*	1.64 (1.16, 2.32)*	1.51 (1.06, 2.15)*	1.51 (1.02, 2.22)*
In-hospital mortality				
TBIL per 1 mg/dL	1.03 (1.01~1.06)*	1.06 (1.03~1.10)*	1.05 (1.02~1.09)*	1.04 (1.01~1.07)*
TBIL category				
<1.2 mg/dL	Reference	Reference	Reference	Reference
≥1.2, <2 mg/dL	1.54 (1.12, 2.13)*	1.54 (1.11, 2.13)*	1.53 (1.10, 2.13)*	1.65 (1.17, 2.33)*
≥2 mg/dL	1.38 (1.03, 1.86)*	1.57 (1.16, 2.14)*	1.47 (1.07, 2.01)*	1.41 (1.01, 1.87)*

Model 1: adjusted for age and gender.

Model 2: further adjusted for pneumonia, hypertension, COPD, renal failure, cancer, and diabetes mellitus.

Model 3: further adjusted for admission type, SpO₂, PEEP, ALP, ALT, AST, serum glucose, and OASIS score.

*: p <0.05.

OR, odds ratio; CI, confidence interval; TBIL, total bilirubin; COPD, chronic obstructive pulmonary disease; SpO₂, pulse oxygen saturation; PEEP, positive end-expiratory pressure; ALP, alkaline phosphatase; ALT, alanine transaminase; AST, aspartate aminotransferase.”

References:

[1] Johnson AE, Kramer AA, Clifford GD. A new severity of illness scale using a subset of Acute Physiology And Chronic Health Evaluation data elements shows comparable predictive accuracy. Crit Care Med. 2013 Jul;41(7):1711-8.

Comment 3. In Figure 1, please include the numbers of patients who were excluded at each step, for completeness.

Reply 3: Thank you so much for your comment. We have revised Figure 1 according to your suggestion and the new one will be uploaded for the manuscript as follows:

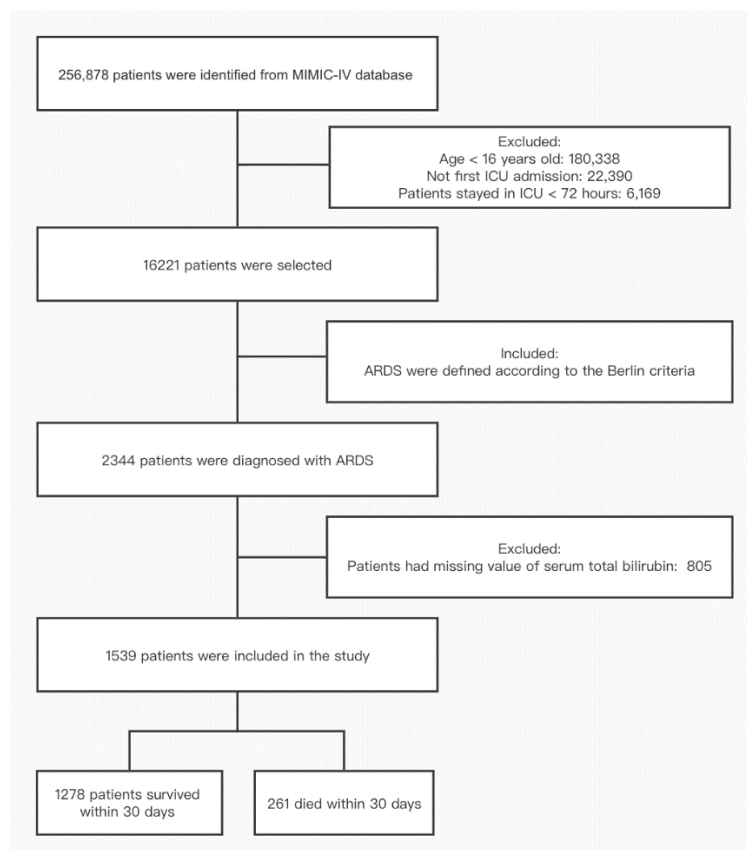


Figure 1. Overview of the current study. MIMIC-IV, Medical Information Mart for Intensive Care-IV; ICU, intensive care unit; ARDS, acute respiratory distress syndrome.

Comment 4: Also for completeness, in addition to the data shown in Table 3, it would be helpful to have another table (perhaps in a supplement) showing all the OR's and 95%CI's for each of the covariates.

Reply 4: We sincerely thank the reviewer for this comment. The new tables showing all OR's and 95%CI's for each of the covariates will be uploaded as appendix for the manuscript as follows:

Changes in the text:

1. Line 209-211 and Line 231-233 in the revised manuscript:

“The OR and 95% CI for each of the covariates in the relationship between TBIL and 30-day ICU mortality are shown in Table A.1 (Supplementary Material).”

“The OR and 95% CI for each of the covariates in the relationship between TBIL and in-hospital mortality are shown in Table A.2 (Supplementary Material).”

“

Appendixes

Table A.2 Relationship between serum total bilirubin and 30-day ICU mortality

	Univariate Analyses		Logistic Multivariate Analyses*	
	OR	95% CI	OR	95% CI
Gender	1.096	0.838-1.432	1.017	0.763-1.354
Age	1.021	1.012-1.03	1.023	1.013-1.033
Pneumonia	0.905	0.691-1.183	0.815	0.609-1.090
Hypertension	1.161	0.888-1.518	1.044	0.781-1.395
COPD	1.117	0.555-2.25	0.767	0.354-1.660
Renal failure	1.669	1.169-2.383	1.612	1.060-2.451
Cancer	1.874	1.326-2.647	1.727	1.189-2.507
Diabetes mellitus	0.725	0.532-0.988	0.700	0.505-0.969
PEEP	1.028	1.004-1.053	1.038	1.001-1.067
Admission type	2.104	1.333-3.322	2.276	1.412-3.670
SpO ₂	0.951	0.924-0.978	0.965	0.936-0.995
OASIS	1.048	1.032-1.064	1.038	1.020-1.056
ALT	1.000	0.999-1.000	1.000	0.999-1.001
ALP	1.001	1.000-1.002	1.000	0.999-1.002
AST	1.000	0.999-1.000	1.000	0.999-1.000
Serum glucose	1.000	0.999-1.000	1.000	0.999-1.000

*: Model 3.

COPD, chronic obstructive pulmonary disease; SpO₂, pulse oxygen saturation; PEEP, positive end-expiratory pressure; ALP, alkaline phosphatase; ALT, alanine transaminase; AST, aspartate aminotransferase. OASIS, Oxford Acute Severity of Illness Score.

Table A.3 In-hospital mortality

	Univariate		Logistic	Multivariate		Logistic
	Analyses			Analyses*		
	OR	95% CI		OR	95% CI	
Gender	1.025	0.817-1.285		0.971	0.762-1.236	
Age	1.023	1.015-1.030		1.025	1.017-1.034	
Pneumonia	0.936	0.747-1.174		0.871	0.681-1.112	
Hypertension	1.051	0.838-1.319		0.937	0.734-1.197	
COPD	0.937	0.504-1.740		0.664	0.336-1.311	
Renal failure	1.345	0.977-1.852		1.352	0.932-1.961	
Cancer	2.493	1.844-3.371		2.445	1.775-3.368	
Diabetes mellitus	0.921	0.717-1.182		0.824	0.621-1.093	
PEEP	1.023	1.001-1.045		1.045	1.019-1.071	
Admission type	1.511	1.074-2.126		1.573	1.101-2.248	
SpO ₂	1.008	0.979-1.037		1.023	0.992-1.055	
OASIS	1.026	1.013-1.039		1.02	1.005-1.034	
ALT	1.000	0.999-1.000		1.000	0.999-1.001	
ALP	1.001	1.000-1.002		1.001	0.999-1.002	
AST	1.000	0.999-1.000		0.999	0.999-1.000	
Serum glucose	1.000	0.999-1.001		1.000	0.999-1.002	

*: Model 3.

COPD, chronic obstructive pulmonary disease; SpO₂, pulse oxygen saturation; PEEP, positive end-expiratory pressure; ALP, alkaline phosphatase; ALT, alanine transaminase; AST, aspartate aminotransferase. OASIS, Oxford Acute Severity of Illness Score.”

Comment 5: A large number of patients were excluded because of missing bilirubin values. Do the authors suspect any differences in this group compared to the study population? A brief discussion of this under limitations would be helpful.

Reply 5: Thank you so much for your comment. We agree with the review that this is one of the limitations of our study. And a brief discussion was added under limitations as follows (see Line 332-333 in the revised manuscript):

Changes in the text:

“Fourth, a large number of patients were excluded because of missing bilirubin values, which may cause some potential bias to the results.”

Comment 6: For patients who were not on invasive mechanical ventilation, did they all receive non-invasive ventilation such as BPAP and all have mild ARDS? Did some patients receive other modalities of oxygen support (which may not completely satisfy the Berlin definition)?

Reply 6: Thank you so much for your comment. The non-invasive ventilation included non-invasive positive pressure ventilation, high flow nasal oxygen/cannula, and supplemental oxygen. For patients who were not on invasive mechanical ventilation, they all received positive pressure ventilation with the positive end-expiratory pressure (PEEP) ≥ 5 cmH₂O (see **Table 1**), meeting the Berlin definition. Detailed modalities of oxygen support were added in the **Method** section (see Line 136-147 in the revised manuscript):

Changes in the text:

“We extracted or calculated the following variables: age, gender, admission type, ventilation status (invasive [including tracheostomy and positive pressure ventilation via endotracheal tube] and non-invasive [including non-invasive positive pressure ventilation, high flow nasal oxygen/cannula, and supplemental oxygen]), baseline measured parameters (including total bilirubin, platelets, white blood cell, alkaline phosphatase [ALP], alanine transaminase [ALT], aspartate aminotransferase [AST], and serum glucose), comorbidity (including pneumonia, hypertension, chronic obstructive pulmonary disease [COPD], renal failure, rheumatic disease, cancer, and diabetes mellitus), vital data taken within 24 hours of ICU admission (heart rate, temperature, respiratory rate, systolic blood pressure [SBP], diastolic blood pressure [DBP]), SpO₂, PaO₂/FiO₂, PEEP, and the Oxford Acute Severity of Illness Score (OASIS) at diagnosis (32).”

Comment 7: In the results section of the abstract (line 38), the adjusted odds ratio should be 1.08, not 0.08.

Reply 7: Thank you very much for your careful review of our paper. This mistake was corrected and the revised sentence was shown below (see Line 39-42 in the revised manuscript):

Changes in the text:

“In the multivariable logistic analysis, each 1 g/dL increase in TBIL levels led to a 4% increase in the odds of 30-day ICU mortality (adjusted odds ratios [OR] =1.04; 95% confidence interval [CI]: 1.01 to 1.08) and a 4% increase in the odds of in-hospital mortality (adjusted OR=1.04; 95% CI: 1.01 to 1.07).”

Comment 8: Please edit the manuscript to correct the typos and grammatical errors.

Reply 8: Thank you very much for your comment. We have got professional help from native speakers to edit the paper for grammar, phrasing, and punctuation (Order ID: AESE20220151). Many edits were made to further improve the flow and readability of the text. Below, we list some revised examples from our manuscript in the chart. We hope the revised version will reach your standards.

Changes in text:

<i>Line</i>	<i>Original version</i>	<i>Revised version</i>
27-29	<i>Serum total bilirubin (TBIL), an end-product of hemoglobin catabolism in mammals reflecting liver dysfunction, has been demonstrated to be an independent indicator for the risk of critically ill patients.</i>	<i>Total bilirubin (TBIL), an end-product of hemoglobin catabolism in mammals reflecting liver dysfunction, has been demonstrated as an independent risk indicator for critically ill patients.</i>
45-47	<i>Consistently, associations between serum level of TBIL and 30-day mortality were found in all subgroups stratified by comorbidities, the severity of ARDS, and other status.</i>	<i>Similarly, associations between serum TBIL levels and 30-day mortality were found in all subgroups stratified by comorbidities, the severity of ARDS, and other variables.</i>
118-122	<i>Our inclusion criteria were as follows: age ≥ 18 years, stay in ICU for more than 72 h, ARDS diagnosis meeting Berlin criteria at the time of ICU admission, had serum total bilirubin levels within 24 h of ICU admission.</i>	<i>Our inclusion criteria were as follows: age ≥ 18 years, an ICU stay of more than 72 hours, an ARDS diagnosis meeting the Berlin criteria at the time of ICU admission, and TBIL levels taken within 24 hours of ICU admission.</i>
155-157	<i>We tested differences in characteristics between groups (survive or dead within 30-day after admission to ICU with Student t-tests for continuous variables and with χ^2 tests for categorical variables.</i>	<i>We tested differences in characteristics between the groups (survival or death within 30 days of admission to the ICU) with Student's t-tests for the continuous variables and χ^2 tests for the categorical variables.</i>
244-246	<i>In 1989, Schwartz and colleagues reported that the levels of TBIL in non-survivors were significantly higher than those in non-survivors for the first week after ARDS diagnosis in 24 ARDS patients (28).</i>	<i>In 1989, Schwartz et al. investigated 24 ARDS patients during the first week after diagnosis and found that TBIL levels in non-survivors were significantly higher than those in survivors (33).</i>
256-258	<i>The level of TBIL ≥ 33 $\mu\text{mol/L}$ was considered as an effective threshold to identify those with a higher mortality</i>	<i>A TBIL level ≥ 33 $\mu\text{mol/L}$ was considered an effective threshold to identify moderate to severe ARDS</i>

	<i>risk in moderate to severe ARDS patients (30).</i>	<i>patients with a higher mortality risk (27).</i>
--	---	--

Reviewer B:

Comment 1: As a general remark, proof reading and editing of English language and grammar are required. Especially the Results section suffers from major grammar problems that lead to misinterpretation of statements and appropriate report of the data.

Reply 1: We sincerely thank the reviewer for this comment. We have got professional help from native speakers for language editing (Order ID: AESE20220151). Many edits were made to further improve the flow and readability of the text. Below, we list some revised examples from our manuscript in the chart. We hope the revised version will reach your standards.

Changes in text:

<i>Line</i>	<i>Original version</i>	<i>Revised version</i>
70-71	<i>However, these parameters are not routine lab indices and obtained from invasive procedure.</i>	<i>However, these parameters are not routine lab indices and are obtained via invasive procedures.</i>
190-191	<i>Of 1539 participants, 258 (16.8%) patients had a level of TBIL\geq2 mg/dL.</i>	<i>Of the 1,539 participants, 258 (16.8%) cases had a TBIL level \geq 2 mg/dL.</i>
191-193	<i>Overall, there were slightly higher rates of 30-day mortality and in-hospital mortality than those in patients with TBIL <2 mg/mL (21.3% vs. 16.1%, P=0.051; 31.0% vs 26.3%, P=0.131, respectively), whereas both did not reach statistically significance.</i>	<i>Overall, there were slightly higher rates of 30-day mortality in cases with higher TBIL levels compared to those with TBIL levels <2 mg/mL (p=0.008).</i>
205-209	<i>When the TBIL levels were treated as categorical variables based on the cut-off value of 2 mg/dL, those with higher TBIL levels had higher odds for 30-day mortality in four models (Un-adjusted model: OR 1.41, 95% CI: 1.01-1.97 Model 1: OR 1.60, 95% CI: 1.14-2.26; Model 2: OR 1.48, 95% CI: 1.04-2.09; Model 3: OR 1.50, 95% CI: 1.03-2.19).</i>	<i>When the TBIL levels were treated as categorical variables, patients with TBIL levels \geq 2 mg/dL had a higher risk of 30-day mortality in the main model (OR = 1.51; 95% CI: 1.02 to 2.22), while those with TBIL levels between 1.2 and 2 mg/dL showed a statistically significant difference according to the lowest TBIL levels in the main model (OR = 1.30; 95% CI: 0.85 to 1.99).</i>

251-253	<i>Also, Sheu and colleagues have included 586 ARDS patients according to American-European Consensus Committee (AECC) criteria, and showed that levels</i>	<i>Also, Sheu et al. investigated 586 ARDS patients using the American-European Consensus Committee (AECC) criteria and showed that levels of TBIL in ICU admission were associated with 60-day mortality.</i>
259-263	<i>In agreement with Sheu’s findings (29), we found that per 1 mg/dL increase in serum TBIL resulted in a slight increase in 30-day ICU mortality of ARDS after adjustments for multiple covariates and a nearly linear dose-dependent association.</i>	<i>In agreement with Sheu et al.’s findings (26), we found that each 1 mg/dL increase in serum TBIL resulted in a slight increase in 30-day ICU mortality in ARDS patients after adjusting for multiple covariates, showing a nearly linear dose-dependent association.</i>
280	<i>Subgroup analyses showed consistent results seen by those from the main findings.</i>	<i>The subgroup analysis results were consistent with the main findings.</i>

Comment 2 : Introduction: - In their Abstract, the authors refer to serum total bilirubin as the end product of hemoglobin metabolism. In contrast, in the introduction, they state, it is derived primarily from haem catabolism. Besides correct spelling, hemoglobin and heme (American English) are two different molecules. Maybe the authors should elaborate with a short paragraph on the background of bilirubin and hemoglobin/heme pathophysiology basics.

Reply 2: Thank you very much for your careful review of our paper. We have corrected the mistakes and added a short paragraph on the background of bilirubin and hemoglobin/heme pathophysiology basics as follows (see Line 77-83 in the revised manuscript):

Changes in the text:

“Bilirubin, an endogenous bile pigment, is derived primarily from hemoglobin catabolism (10). Briefly, when old or damaged erythrocytes are engulfed and degraded by macrophages, hemoglobin will be released and broken down into two main components, the globin, and the hemes. The globin is reused for erythropoiesis, while the hemes are degraded into unconjugated bilirubin (UCB). Next, UCB, a lipid-soluble molecule, is carried to the liver and converted to conjugated bilirubin (CB) with the addition of glucuronic acid (11).”

Comment 3 : Methods: - The rationale for separating into two groups using a serum total bilirubin concentration of 2 mg/dl as a break point needs clarification. 2mg/dl is already a significantly elevated serum total bilirubin concentration. In addition, according to Table 2 more than 80% of patients are included in the “lower” group, which could bias the group comparison.

How do the results change when a serum total bilirubin concentration of 1.2 mg/dl (normal upper limit value) is considered for separation. Moreover, three groups could be compared with concentrations of 1.2 mg/dl and 2 mg/dl as break points.

Reply 3: We sincerely thank the reviewer for this comment. We agree with the reviewer. Accordingly, the patients were stratified into three groups using serum TBIL concentration with the cut-off value of 1.2 mg/dL and 2 mg/dL. Table 2 and Table 3 were revised correspondingly as follows:

“Table 4. Outcomes of ARDS patients across different concentrations of total serum bilirubin

Outcomes	All patients	total serum bilirubin (mg/dL)			P-value
		<1.2	≥1.2, <2	≥2	
n (%)	1539	1077 (70.0)	204 (13.3)	258 (16.8)	-
30-day mortality	ICU 261 (17.0)	170 (15.8)	36 (17.6)	55 (21.3)	0.1
In-hospital mortality	418 (27.2)	268 (24.9)	69 (33.8)	87 (31.0)	0.008
LOS in ICU	6.8 (4.5, 11.6)	6.8 (4.5, 11.7)	7.4 (4.7, 11.9)	6.4 (4.6, 10.2)	0.375
LOS in hospital	14.7 (9.0, 22.9)	14.1 (8.7, 22.1)	15.5 (9.9, 22.2)	15.6 (10.0, 25.0)	0.026

Data are displayed as the median (IQR) and n (%). P-values comparing groups are from the Student’s t-test for continuous data and the chi-squared test for categorical variables. ICU, intensive care unit. LOS, length of stay.”

“Table 5. Relationship between total serum bilirubin and outcomes in different models

Outcomes	Non-adjusted	Model 1	Model 2	Model 3
30-day mortality	ICU OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
TBIL per 1 mg/dL	1.04 (1.01, 1.06)*	1.07 (1.04, 1.11)*	1.07 (1.03, 1.11)*	1.04 (1.01, 1.08)*
TBIL category				

<1.2 mg/dL	Reference	Reference	Reference	Reference
≥1.2, <2 mg/dL	1.14 (0.77, 1.70)	1.14 (0.77, 1.70)	1.13 (0.75, 1.69)	1.30 (0.85, 1.99)
≥2 mg/dL	1.45 (1.03, 2.03)*	1.64 (1.16, 2.32)*	1.51 (1.06, 2.15)*	1.51 (1.02, 2.22)*
In-hospital mortality	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
TBIL per 1 mg/dL	1.03 (1.01~1.06)*	1.06 (1.03~1.10)*	1.05 (1.02~1.09)*	1.04 (1.01~1.07)*
TBIL category				
<1.2 mg/dL	Reference	Reference	Reference	Reference
≥1.2, <2 mg/dL	1.54 (1.12, 2.13)*	1.54 (1.11, 2.13)*	1.53 (1.10, 2.13)*	1.65 (1.17, 2.33)*
≥2 mg/dL	1.38 (1.03, 1.86)*	1.57 (1.16, 2.14)*	1.47 (1.07, 2.01)*	1.41 (1.01, 1.87)*

Model 1: adjusted for age and gender.

Model 2: further adjusted for pneumonia, hypertension, COPD, renal failure, cancer, and diabetes mellitus.

Model 3: further adjusted for admission type, SpO₂, PEEP, ALP, ALT, AST, serum glucose, and OASIS score.

*: p <0.05.

OR, odds ratio; CI, confidence interval; TBIL, total bilirubin; COPD, chronic obstructive pulmonary disease; SpO₂, pulse oxygen saturation; PEEP, positive end-expiratory pressure; ALP, alkaline phosphatase; ALT, alanine transaminase; AST, aspartate aminotransferase.”

Comment 4: - Line 133: Were adjustments made for multiple testing if required?

Reply 4: Thank you so much for your comment. In the present study, we did not perform multiple testing, because we set the group with TBIL <1.2 mg/dL as the reference, with which the other two groups (the group with 1.2 mg/dL ≤ TBIL <2 mg/dL and group with TBIL ≥2 mg/dL) were compared. Therefore, multiple testing was not required in our study.

Comment 5: Statistical analysis: What is the rationale for using 3 different model of adjustment?

Reply 5: Thank you so much for your comment. The present study aimed to examine whether TBIL on intensive care unit (ICU) admission is independently associated with ARDS mortality. According to **Table 1**, multiple independent variables were significantly different between 30-day survivors and 30-day ICU deaths. Therefore, age, gender, pneumonia, hypertension, COPD, renal failure, cancer, diabetes mellitus, admission type, SpO₂, PEEP, ALP, ALT, AST, serum glucose, and OASIS score were adjusted in multivariate logistic regression models. According to the previous study^[1-3], we adjusted age and gender in model 1, comorbidities were adjusted in model 2, and the rest of the variates were adjusted in model 3.

References:

- [1] Ruan Z, Lu T, Chen Y, et al. Association Between Psoriasis and Nonalcoholic Fatty Liver Disease Among Outpatient US Adults. *JAMA Dermatol.* 2022 May 25:e221609.
- [2] Qiu Z, Chen X, Geng T, et al. Associations of Serum Carotenoids With Risk of Cardiovascular Mortality Among Individuals With Type 2 Diabetes: Results From NHANES. *Diabetes Care.* 2022 Jun 2;45(6):1453-1461. doi: 10.2337/dc21-2371. PMID: 35503926.
- [3] Chen Y, Chang Z, Zhao Y, et al. Association between the triglyceride-glucose index and abdominal aortic calcification in adults: A cross-sectional study. *Nutr Metab Cardiovasc Dis.* 2021 Jun 30;31(7):2068-2076. doi: 10.1016/j.numecd.2021.04.010. Epub 2021 Apr 19. PMID: 34053833.

Comment 6: Results: - Complete makeover of English language and grammar necessary. Some parts are not understandable

Line 160: “of these”: male only?

Line 163: cancer (20.3% vs. 112%) A value >100%?, What is the P-value?

Line 165: What is body temperature respiratory rate?

Line 166: positive end-expiratory pressure was already named PEEP

Line 168 ff: Why is no Standard Error used for ALP and AST? What is used?

Line 176: unclear from text what mortality refers to what numbers. In addition, the only real difference between groups is ICU LOS.

Reply 6: Thank you so much for your comments. As we mentioned above (See **Reply 1**), we have got professional help from native speakers for language editing. Many edits were made to further improve the flow and readability of the text. We have corrected the sentences in the first paragraph of the **Results** section as follows:

Changes in the text:

1. Line 177-187 in the revised manuscript:

“The total group had a mean age of 62.3 ± 16.3 years, and most were male (56.7%). Of these patients, 261 (17.0%) died in the ICU within 30 days. Compared to the survivors, cases who died within 30 days were older (66.6 ± 16.0 vs. 61.4 ± 16.2 , $p < 0.001$) and more likely to have comorbidities, such as renal failure (18.4% vs. 11.9%, $p = 0.006$) and cancer (20.3% vs. 12%, $p < 0.001$), but not diabetes mellitus (23.8% vs. 30%, $p = 0.049$). Regarding the baseline vital indexes, non-survivors had lower heart rate levels, temperature, respiratory rate, blood pressure, SpO₂, and PaO₂/FiO₂ but higher PEEP levels. Compared with survivors, those who died showed a higher level of TBIL, ALP, and AST. The rate of invasive ventilation was significantly higher in non-survivors than in survivors.”

2. Line 191-194 in the revised manuscript:

“Overall, there were slightly higher rates of 30-day mortality in cases with higher TBIL levels compared to those with TBIL levels < 2 mg/mL ($p = 0.008$). However, the 30-day ICU mortality did not reach statistical significance ($p = 0.1$).”

Comment 7: - Line 179 ff: Complicated to understand why three different but successive models for adjustment of covariants are used. Especially because the results are very similar among all three models. In addition, very complicated for the reader to follow.

Reply 7: Thank you so much for your comment. As we mentioned in **Reply 5**, multiple variates were adjusted in different models following previous studies. The original purpose was to clearly show every step of the adjustments with successive models, because the results may not consistent in different models. However, we found that our study showed a consistent result among all the models, indicating that a higher serum TBIL on ICU admission was associated with mortality in ARDS patients after adjustments for multiple covariants.

Comment 8: - Line 211: Table 2 does not present in-hospital mortality with Model 3 adjustment.

Reply 8: Thank you so much for your suggestion. We have corrected the error in the **Results** section as follows (see Line 229-231 in the revised manuscript):

Changes in the text:

“A linear relationship was present between TBIL levels and in-hospital mortality after adjustment in Model 3 (Table 3) (p for nonlinear = 0.441, Fig. 2B).”

Comment 9: - Line 264 ff: “the strength of associations in subgroup analyses may be modified and attenuated” – Please clarify

Reply 9: Thank you so much for your comments. This sentence has been corrected. And the revised version of the sentence is shown down below (see Line 284-286 in the revised manuscript):

Changes in the text:

“Given that older age, female gender, lower blood pressure, renal failure, and severe ARDS impact ARDS development and mortality, [the strength of associations in the subgroup analyses may have been attenuated \(1, 3\).](#)”

Comment 10: - Line 275 ff: The authors discuss only liver dysfunction as major TBIL-associated determinant for increased ARDS mortality. However, in the introduction they highlight TBIL as endproduct of hemoglobin/haem catabolism. Recently, some studies have addressed the association between hemolysis and outcome in ARDS patients (DOI: 10.1186/s13054-022-03894-5, DOI: 10.1152/ajplung.00155.2015). This should be discussed.

Reply 10: We sincerely thank the reviewer for this comment. We agree with the reviewer and this comment is valuable and helpful in revising and improving our manuscript. These two references were discussed in the **Discussion** section as follows (see Line 299-303 in the revised manuscript):

Changes in the text:

“In support of our results, a recent study found that higher circulating cell-free hemoglobin (CFH), one of the precursors of bilirubin, is an independent risk factor for acute kidney injury in ARDS patients (37). Similarly, Shaver et al. reported that CFH contributed to ARDS severity by enhancing lung permeability and inflammation in an experimental ARDS mouse model (38).”

Comment 11: - LFTs are grossly not different between groups. How does this impact the hypothesis on liver dysfunction?

Reply 11: Thank you so much for your comment. ALP, AST, and ALT are the parameters of liver function. As shown in **Table 1**, ALP ($p < 0.001$) and AST ($p = 0.02$) were significantly higher in 30-day ICU deaths as compared with 30-day survivors, while ALT ($p = 0.259$) is not different between these two groups. To evaluate whether TBIL on intensive care unit (ICU) admission is associated with ARDS mortality, multiple variates were adjusted in the logistic regression model. Our results indicated that TBIL is associated with ARDS mortality even after adjustment for multiple variates (including ALP, ALT, and AST). Suppose the liver function tests were not different between groups, this may indicate that LFTs do not impact ARDS death, supporting our hypothesis that TBIL is associated with ARDS mortality.

Comment 12: - Line 297 – Sample sizes of subgroups are another significant limitation

Reply 12: Thank you very much for your careful review of our paper. We agree with the reviewer. This limitation was added in the corresponding part in the **Discussion** section as follows (see Line 334-336 in the revised manuscript):

Changes in the text:

“Last, the small size of our subgroups was another significant limitation, and a larger population of such subgroups is recommended in the future to draw a more robust conclusion.”

Comment 13: Figure 1:- How many patients had missing values of serum total bilirubin. Where there other patients excluded due to missing values?

Were potential ARDS patients excluded due to missing values?

Reply 13: Thank you so much for your comments. To make every step clear in Figure 1, we revised the flowchart with a much more thorough illustration of the missing values in each step. Overall, 805 ARDS patients were excluded due to missing value. Given this fact, a brief discussion was added under limitations as follows:

Changes in the text:

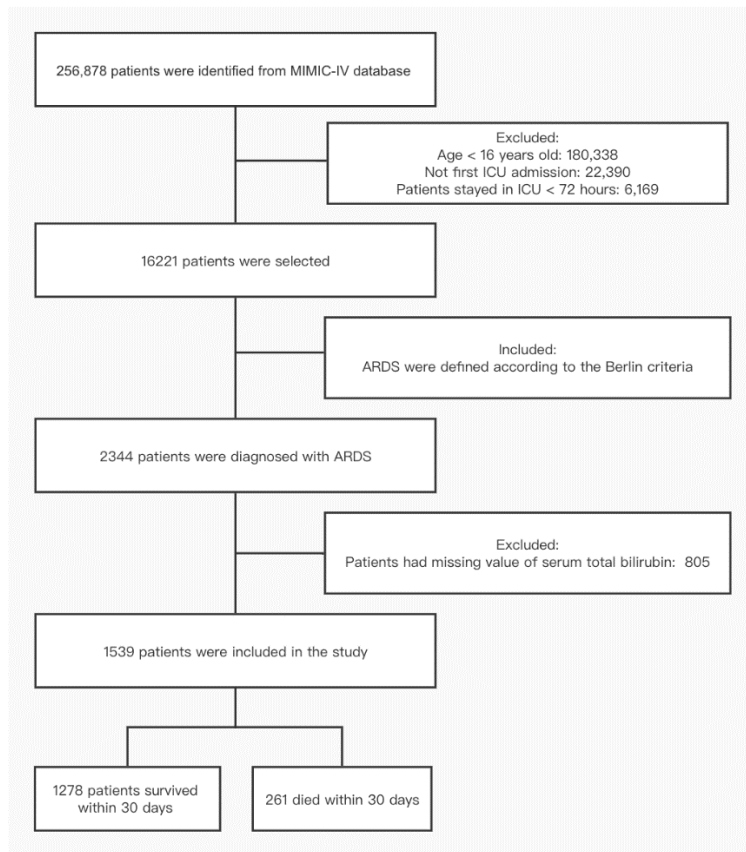


Figure 1. Overview of the current study. MIMIC-IV, Medical Information Mart for Intensive Care-IV; ICU, intensive care unit; ARDS, acute respiratory distress syndrome.

Line 332-333 in the Discussion Section:

“Fourth, a large number of patients were excluded because of missing bilirubin values, which may cause some potential bias to the results.”

Second Round of Peer Review

Reviewer A

Comment 1: The authors have satisfactorily addressed the questions and comments from the reviewers. Copy editing for typos and grammatical errors is still needed.

Reply 1: Thank you so much for your appreciation and comments. We have corrected the grammatical errors in the revised manuscript.

Reviewer B

Comment 1: Revised Figure 1: In their flow chart, the authors included that 805 patients were excluded due to missing serum total bilirubin levels. I believe, these were TBILs from the admission lab since the authors specify in their introduction that they study TBILs on admission as prognostic marker. Please clarify in the flow chart or figure legend.

Reply 1: Thank you so much for your comments. We have added a sentence in the figure legend (please kindly see Line 544 in the revised manuscript).

Changes in text:

“Notes: The values of serum total bilirubin were from the admission lab.”

Comment 2: Introduction: Line 182: wording. “but not diabetes mellitus“ sounds like there was no difference between groups. However, patients who died within 30 days had less frequently a diabetes mellitus.

Reply 2: We sincerely thank the reviewer for this comment. We have corrected the sentence in the revised manuscript (please kindly see Line 190-193 in the revised manuscript).

Changes in text:

“Compared to the survivors, cases who died within 30 days were older (66.6 ± 16.0 vs. 61.4 ± 16.2 , $p < 0.001$) and more likely to have comorbidities, such as renal failure (18.4% vs. 11.9%, $p = 0.006$) and cancer (20.3% vs. 12%, $p < 0.001$), but less likely to have diabetes mellitus (23.8% vs. 30%, $p = 0.049$).”

Comment 3: Line 191 ff: „Overall, there were slightly higher rates of 30-day mortality in cases with higher TBIL levels compared to those with TBIL levels <2 mg/mL (p=0.008)“ This statement appears not correct since in-hospital mortality rate is 33.8% for patients with TBIL levels ≥ 1.2 , <2 and “only“ 31% in patients with TBIL levels ≥ 2 (Table 2).

Reply 3: Thank you very much for your careful review of our paper. We have revised the sentence in the updated manuscript (please kindly see Line 202-204 in the revised manuscript).

Changes in text:

“Overall, there were lower rates of in-hospital mortality in cases with higher TBIL levels compared to those with TBIL levels <2 mg/mL (p=0.008).”

Comment 4: Line 194 ff: Patients with higher TBIL levels had a longer length of stay in the ICU and hospital (p=0.375 and p=0.026, respectively). ICU-LOS appears not different between groups (p=0.375). Furthermore, ICU-LOS in patients with TBIL ≥ 2 is shorter (6.4) compared to all other groups (6.8 and 7.4, respectively). Please revise.

Furthermore, statistical analysis of data presented in Table 2 raises questions. Now the authors compare 3 groups instead of 2. A student’s t-test is for pairwise comparison. Here, the authors state, they are comparing three groups. Please revise statistical data analysis.

Reply 4: Thank you so much for your comment and suggestion. We have added the sentences in the updated manuscript. We did use the chi-squared test only in Table 2, and the statistical data analysis was revised in Table 2 legend.

Changes in text:

Line 206-209:

“ICU-LOS appears not different between groups (p=0.375). Furthermore, ICU-LOS in patients with TBIL ≥ 2 is shorter (6.4 days) compared to all other groups (6.8 days and 7.4 days, respectively).”

Line 571-572:

“P-values comparing groups are the chi-squared test for categorical variables.”

Comment 5: Line 250: „sepsis-related 28-day mortality and 60-day mortality increased by 20% and 18%, respectively“. Mortality increase in this study on

ARDS patients is only 4%. Do the authors have an explanation for the only small increase per 1 mg/dl TBIL in ARDS patients given most ARDS patients suffer from sepsis too? Did the previous studies use adjustment models for statistical analysis?

Reply 5: Thank you very much for your careful review of our paper. A previous study found that sepsis-related 28-day mortality and 60-day mortality increased by 20% and 18% after adjustment for multiple covariates. However, this study recruited only 326 ARDS patients with sepsis (there were 1539 participants in our study), which may cause potential bias and lead to an overestimation of the conclusion. On the other hand, although sepsis is one of the major causes of ARDS, more than 30% of ARDS patients result from non-sepsis conditions, including aspiration, trauma, blood transfusion, Pulmonary contusion, and others. Therefore, it may be reasonable that the value in our result is lower than those in the previous study.

Comment 6: Line 268: "...although statistical significance was not reached in either group". Table 3 indicates statistical differences for the different models. Please specify this statement with regard to the analysis you are referring to.

Reply 6: We sincerely thank the reviewer for this comment. This statement was derived from **Table 2**. To make it clear, we have revised the sentence in the updated manuscript (please kindly see Line 302-308 in the revised manuscript).

Changes in text:

"Also, we found that 30-day ICU mortality in patients with serum TBIL levels ≥ 2 mg/dL was higher than those with serum TBIL between 1.2 to 2 mg/dL or those with TBIL < 1.2 mg/dL (25), although statistical significance was not reached."