Alimentary Pharmacology & Therapeutics

WILEY

AP&T Alimentary Pharmacology & Therapeutics

RESEARCH COMMUNICATION OPEN ACCESS

Comparison Between Dynamic Models for Predicting Response to Corticosteroids in Alcohol-Associated Hepatitis: A Global Cohort Study

Francisco Idalsoaga[1,2] | Luis Antonio Díaz[2,3] | Leonardo Guizzetti[4] | Winston Dunn[5] | Heer Mehta[5] | Jorge Arnold[2] | Gustavo Ayares[2] | Rokhsana Mortuza[1] | Gurpreet Mahli[1] | Alvi H. Islam[1] | Shiv K. Sarin[6] | Rakhi Maiwall[6] | Wei Zhang[7] | Steve Qian[8] | Douglas Simonetto[9] | Ashwani K. Singal[10] | Mohamed A. Elfeki[10] | Carolina Ramirez-Cadiz[11] | Joaquín Cabezas[12,13] | Victor Echavarría[12] | Meritxell Ventura Cots[14] | María Fátima Higuera-De La Tijera[15] | Juan G. Abraldes[16] | Mustafa Al-Karaghouli[16] | Prasun K. Jalal[17] | Mohamad Ali Ibrahim[17] | Guadalupe García-Tsao[18] | Daniela Goyes[18] | Lubomir Skladaný[19] | Daniel J. Havaj[19] | Karolina Sulejova[19] | Svetlana Adamcova Selcanova[19] | Diego Rincón[20,21] | Vijay H. Shah[9] | Patrick S. Kamath[9] | Marco Arrese[2] | Ramon Bataller[22,23] | Juan Pablo Arab[2,24]

[1]Division of Gastroenterology and Hepatology, Western University & London Health Sciences Centre, London, Canada | [2]Departamento de Gastroenterología, Escuela de Medicina, Pontificia Universidad Católica de Chile, Santiago, Chile | [3]MASLD Research Center, Division of Gastroenterology and Hepatology, University of California San Diego, La Jolla, California, USA | [4]Maple Statistical Consulting, Ottawa, Canada | [5]Division of Gastroenterology, Department of Medicine, University of Kansas Medical Center, Kansas City, Missouri, USA | [6]Department of Hepatology, Institute of Liver and Biliary Sciences, New Delhi, India | [7]Gastroenterology Unit, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA | [8]Division of Gastroenterology and Hepatology, University of Florida, Gainesville, Florida, USA | [9]Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, Minnesota, USA | [10]Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, University of Louisville School of Medicine, Louisville, Kentucky, USA | [11]Department of Anesthesiology, Virginia Commonwealth University School of Medicine, Richmond, Virginia, USA | [12]Gastroenterology and Hepatology Department, University Hospital Marqués de Valdecilla, Santander, Spain | [13]Research Institute Valdecilla (IDIVAL), Santander, Spain | [14]Liver Unit, Hospital Vall D'hebron, Universitat Autonoma Barcelona, Ciberehd, Barcelona, Spain | [15]Servicio de Gastroenterología, Hospital General De México, Universidad Nacional Autónoma de México, Mexico City, Mexico | [16]Division of Gastroenterology, Liver Unit, University of Alberta, Edmonton, Alberta, Canada | [17]Department of Gastroenterology and Hepatology, Baylor College of Medicine, Houston, Texas, USA | [18]Section of Digestive Diseases, Yale University School of Medicine/VA-CTt Healthcare System, New Haven/West Haven, Connecticut, USA | [19]Division of Hepatology, Gastroenterology and Liver Transplantation, Department of Internal Medicine II, Slovak Medical University, F. D. Roosevelt University Hospital, Banska Bystrica, Slovak Republic | [20]Liver Unit, Department of Digestive Diseases Hospital General, Universitario Gregorio Marañón, Madrid, Spain | [21]Ciberehd Centro De Investigación Biomédica En Red De Enfermedades Hepáticas Y Digestivas, Madrid, Spain | [22]Liver Unit, Hospital Clinic, Barcelona, Spain | [23]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain | [24]Division of Gastroenterology, Hepatology, and Nutrition, Department of Internal Medicine, Virginia Commonwealth University School of Medicine, Richmond, Virginia, USA

Correspondence: Juan Pablo Arab (juanpablo.arab@vcuhealth.org)

Received: 12 December 2024 | **Revised:** 2 January 2025 | **Accepted:** 4 February 2025

Handling Editor: Daniel Huang

Funding: Marco Arrese receives support from the Chilean government through the Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT 1241450).

Keywords: alcohol | alcohol-associated liver disease | alcoholic cirrhosis | alcohol-related hepatitis | cirrhosis | end-stage liver disease | outcome prediction

Abbreviations: AH, alcohol-associated hepatitis; ALD, alcohol-associated liver disease; ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUC, area under the curve; AUROC, area under the receiver operating characteristic; CI, confidence interval; IQR, interquartile range; Lille-4, Lille day 4 score; Lille-7, Lille day 7 score; MELD, model for end-stage liver disease; NIAAA, National Institute on Alcohol Abuse and Alcoholism; NLR, neutrophil-to-lymphocyte ratio; TSB, trajectory of serum bilirubin.

ABSTRACT

Several dynamic models predict mortality and corticosteroid response in alcohol-associated hepatitis (AH), yet no consensus exists on the most effective model. This study aimed to assess predictive models for corticosteroid response and short-term mortality in severe AH within a global cohort. We conducted a multi-national study of patients with severe AH treated with corticosteroids for at least 7 days, enrolled between 2009 and 2019. Dynamic models—Lille-4, Lille-7, trajectory of serum bilirubin (TSB), and neutrophil-to-lymphocyte ratio (NLR)—were used to estimate 30- and 90-day mortality. Lille-7 demonstrated the highest accuracy for both 30- and 90-day mortality.

1 | Introduction

Alcohol-associated liver disease (ALD) is a leading cause of chronic liver disease and death globally [1]. Severe alcohol-associated hepatitis (AH) is an acute entity characterised by high short-term mortality, up to 30%–50% at 6 months [2]. Corticosteroids are the current treatment for severe AH, improving survival at 30 days, but they also carry significant risks [3–5]. Various dynamic models, such as the Lille day 7 score, Lille day 4 score, change in the neutrophil-to-lymphocyte ratio (NLR), and the trajectory of serum bilirubin (TSB), are used to identify patients who are benefiting or not from corticosteroids. Still, there are no studies assessing a head-to-head comparison [6–8]. This study aimed to evaluate and compare dynamic scores to predict response to corticosteroid treatment and short-term survival in patients with severe AH.

2 | Material and Methods

2.1 | Study Design and Participants

We conducted a multinational cohort study that included patients admitted with severe AH, defined according to the clinical criteria established by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) [9]. Eligible patients met the NIAAA criteria, received steroid treatment within 30 days of admission, and had a MELD score ≥ 20 at admission.

We excluded participants who deceased within 7 days after admission, patients under 18 years old, pregnant individuals, those with aspartate aminotransferase (AST) and/or alanine aminotransferase (ALT) levels above 400 IU/mL, patients with prolonged alcohol abstinence (> 60 days) before presentation, individuals with drug-induced liver injury, ischaemic hepatitis, biliary duct obstruction, viral hepatitis, autoimmune hepatitis, or Wilson disease. Additionally, patients with hepatocellular carcinoma beyond Milan criteria, extrahepatic neoplasia with a life expectancy of fewer than 6 months, or a history of severe extrahepatic disease leading to a survival of fewer than 6 months were excluded.

We obtained clinical data from the patients and laboratory tests. The Lille Score was calculated using the original formula, and a score ≥ 0.45 on Day 7 indicated unresponsiveness, leading to the discontinuation of corticosteroid treatment [6]. The Lille Score on Day 4 was also calculated using the same formula but with Day 4 bilirubin instead of Day 7 bilirubin [7]. For TSB, patients categorised as "fast fallers" (bilirubin $< 0.8 \times$ admission value at day 7) were considered to have no benefit from corticosteroid treatment [8]. While an NLR value between 5 and 8 at baseline was considered indicative of patients likely to benefit from corticosteroid therapy [10]. The diagnosis of cirrhosis was determined based on medical history and imaging methods, including ultrasound, transient elastography, computed tomography, and magnetic resonance imaging. Only the research team had access to the data, and we requested a waiver of informed consent in each participating centre. The study was approved by the Ethics Committee of Pontificia Universidad Católica de Chile.

2.2 | Statistical Analysis

This study aimed to evaluate and compare dynamic scores to predict response to corticosteroid treatment. The primary and secondary outcomes were defined as 30-day and 90-day mortality, respectively, considering that the different models were designed to define mortality as the outcome. Categorical variables were analysed with Chi-squared, continuous variables with Student's *t*-test or non-parametric tests, and discrimination ability was assessed using the area under the receiver operating characteristic (AUC, also called the *c*-statistic), compared with an extension of the DeLong method [11]. Youden's J index was provided for each ROC curve. The Lille-7 score served as the gold standard. Multiple imputation was used for missing data for sensitivity analyses. Analyses were adjusted using Dunnett-type (many-to-one) simultaneous confidence intervals with the MOVER method [12], which is suitable for comparing several correlated areas under the curve (AUCs). As the AUC is meaningfully interpreted as a probability, its differences are in units of percentage points.

3 | Results

3.1 | Baseline Characteristics of the Cohort

We included a total of 289 patients with sAH from 18 centres (8 countries on three continents). The mean age was 48.3 [IQR: 41.0–56.0] years, and 34.3% (99) were women. The most frequent ethnicities were Non-Hispanic White 73% (211), Mestizo 8% (23), Hispanic or Latino 7.3% (21), and Black 3.5% (100). The 77.5% (221) had a prior history of cirrhosis. The median MELD on admission was 28.0 [26.0–30.0]. Of the total cohort, 54.2% were responders to corticosteroids according to Lille-4, 51.7% according to Lille-7, 44.4% according to TSB, and 22.9% according to Delta NLR. At admission, patients presented with a median bilirubin of 16.1 [10–25.3] mg/dL, an International Normalized Ratio (INR) of 2.1 [1.6–2.2], and albumin 2.6 [2.0–3.0] g/dL. The median creatinine at admission was 1.2 [0.7–1.2] mg/dL. In the cohort, 39% (61 patients) had documented infections; however,

data on infections were missing for 133 patients, and median steroid use was 21 (IQR: 7–31). The mortality at 30 days and 90 days was 13.5% and 25.3%, respectively. Only 5.5% (16) of patients underwent liver transplantation during the follow-up period, with 3.11% [9] receiving an early transplant. No patients in the cohort received plasma exchange. The main causes of death were multi-organ failure (46.4%, 58), infections 14.4% (18), gastrointestinal bleeding 10.4% [13], and acute kidney injury 7.2% [9].

3.2 | Performance of Score for Predicting 30-Day Mortality

In evaluating the performance of different scores, the Lille-7 score demonstrated the highest predictive accuracy for 30-day mortality (AUC = 0.724, 95% CI: 0.633–0.815). The TSB score also exhibited good predictive value (AUC = 0.703, 95% CI: 0.626–0.780), while the Lille-4 score had a slightly lower AUC (0.674, 95% CI: 0.579–0.770). In contrast, the Delta NLR showed limited predictive ability (AUC = 0.527, 95% CI: 0.428–0.626), suggesting it may be less effective as a mortality predictor in this cohort (Table 1) (Figure 1).

The differences in AUC between the scores, both unadjusted and adjusted for multiple comparisons, did not reach statistical significance. The unadjusted difference between the TSB and Lille-7 scores was −0.02 (95% CI: −0.073–0.032, $p = 0.443$), and the difference between Lille-4 and Lille-7 scores was −0.048 (95% CI: −0.11–0.014, $p = 0.128$). After adjustment for multiple comparisons, the adjusted AUC difference between the TSB and Lille-7 scores was −0.020 (95% CI: −0.080–0.050), while the adjusted difference between Lille-4 and Lille-7 scores was −0.048 (95% CI: −0.125–0.028).

3.3 | Performance of Score for Predicting 90-Day Mortality

In the analysis of risk scores predicting 90-day mortality, the Lille-7 score again exhibited the highest predictive accuracy (AUC = 0.710, 95% CI: 0.635–0.784), followed closely by the Lille-4 score (AUC = 0.698, 95% CI: 0.622–0.774). The TSB score demonstrated moderate predictive value (AUC = 0.650, 95% CI: 0.579–0.722), while the Delta NLR continued to show limited predictive capability (AUC = 0.546, 95% CI: 0.467–0.626),

TABLE 1 | Individual estimation of the AUC of risk scores to predict mortality by 30 days and 90 days.

	Mortality 30 day				Mortality 90 day			
Score	**AUC**	**95% CI**	***p***	**Youden's J**	**AUC**	**95% CI**	***p***	**Youden's J**
TSB	0.703	(0.626–0.780)	< 0.0001	0.92	0.650	(0.579–0.722)	< 0.0001	0.85
Lille-4	0.674	(0.579–0.770)	< 0.0001	0.60	0.698	(0.622–0.774)	< 0.0001	0.73
Lille-7	0.724	(0.633–0.815)	< 0.0001	0.64	0.710	(0.635–0.784)	< 0.0001	0.64
NLR	0.527	(0.428–0.626)	0.284	5.23	0.546	(0.467–0.626)	0.023	4.13

Note: Youden's J is defined as the value of the score variable that maximises the sum of sensitivity and specificity.
Abbreviations: Lille-4, Lille at Day 4; Lille-7, Lille at Day 7; NLR, delta neutrophil-to-lymphocyte ratio; TSB, trajectory serum bilirubin.

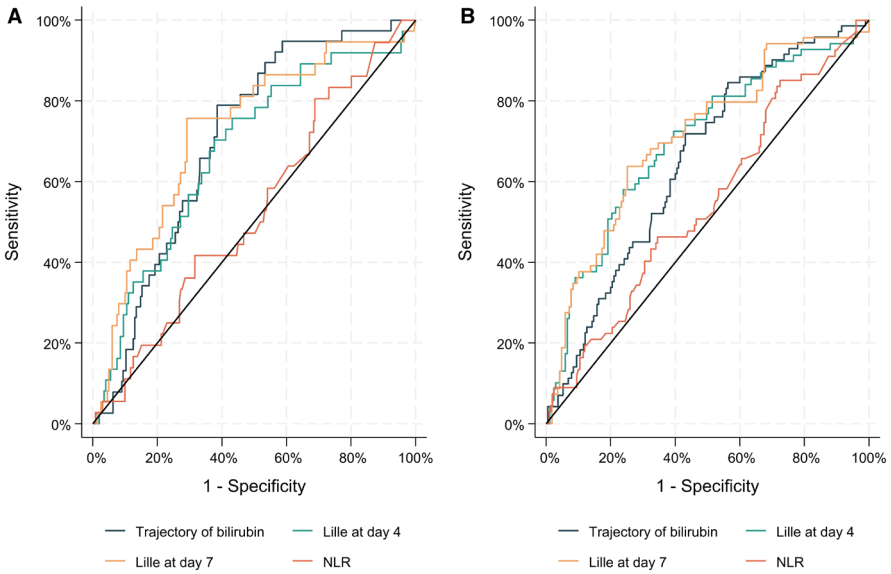

FIGURE 1 | ROC curves of risk scores to predict mortality by 30 days and 90 days. The figure displays the performance of different risk scores in predicting mortality, with panel (A) representing the 30-day mortality predictions and panel (B) illustrating the 90-day mortality predictions. Lille-4, Lille at day 4; Lille-7, Lille at day 7; NLR, delta neutrophil-to-lymphocyte ratio; TSB, trajectory serum bilirubin.

consistent with the findings for 30-day mortality (Table 1) (Figure 1).

The unadjusted AUC difference between the TSB and Lille-7 scores was statistically significant ($\Delta = -0.059$, 95% CI: −0.108 to −0.010, $p = 0.017$), whereas the difference between Lille-4 and Lille-7 scores was not significant ($\Delta = -0.014$, 95% CI: −0.058 to 0.031, $p = 0.542$). After adjusting for multiple comparisons, the adjusted AUC difference between the TSB and Lille-7 scores remained −0.059 (95% CI: −0.117 to 0.001), and the adjusted difference between Lille-4 and Lille-7 scores was −0.014 (95% CI: −0.069 to 0.041). Sensitivity analyses showed similar results to those reported.

4 | Discussion

Severe AH is associated with high short-term mortality, and corticosteroids are the only therapy shown to reduce 30-day mortality in this condition. Still, they carry significant risks of adverse effects, including infections, gastrointestinal bleeding, hyperglycaemia, and psychological disorders [13, 14]. Predictive models, including Lille-7, Lille-4, and TSB, help identify patients who may not benefit from corticosteroid therapy, reducing unnecessary exposure. Delta NLR demonstrated poor discriminatory ability, and its use is not clinically justified based on this work. This study confirms the utility of Lille-7, Lille-4, and TSB models in predicting corticosteroid response and mortality in severe AH. Lille-7 exhibited the highest predictive accuracy for both 30-day and 90-day mortality, but the differences between the models were not statistically significant. For predicting 90-day mortality, all models showed reduced predictive performance.

Despite advancements, the current models for predicting corticosteroid response are imperfect, achieving adequate discrimination with AUC values greater than 0.70. NLR, however, lacks the discriminatory power required for clinical decision-making. The use of Lille-4 models is clinically relevant as they help reduce corticosteroid exposure, potentially improving patient outcomes by minimising unnecessary risks. This is particularly important for patients with poorly controlled comorbidities or a high infection risk.

Our study has strengths, including a large, global cohort from 23 countries, providing a representative sample. It is the first to compare multiple risk models for corticosteroid response and mortality directly, adding significant value to the existing literature. However, its retrospective design limits control over data recording and co-interventions, introducing potential bias. Additionally, reasons for corticosteroid discontinuation and the presence of concomitant infections were not systematically captured. In conclusion, this multi-national cohort study demonstrates that Lille-7, Lille-4, and TSB models predict 30-day mortality effectively, with Lille-7 showing the highest accuracy. The use of Lille-4 models reduces corticosteroid exposure, particularly in high-risk patients. However, none of the models demonstrate optimal performance, highlighting the need for improved strategies for risk stratification.

Author Contributions

Francisco Idalsoaga: conceptualization, investigation, writing – review and editing, writing – original draft. **Luis Antonio Díaz:** conceptualization, writing – review and editing, writing – original draft. **Leonardo Guizzetti:** methodology, formal analysis. **Winston Dunn:** investigation, writing – review and editing. **Heer Mehta:** investigation, writing – review and editing. **Jorge Arnold:** investigation, writing – review and editing. **Gustavo Ayares:** investigation, writing – review and editing. **Rokhsana Mortuza:** writing – review and editing. **Gurpreet Mahli:** writing – review and editing. **Alvi H. Islam:** writing – review and editing. **Shiv K. Sarin:** writing – review and editing, investigation. **Rakhi Maiwall:** investigation, writing – review and editing. **Wei Zhang:** investigation, writing – review and editing. **Steve Qian:** investigation, writing – review and editing. **Douglas Simonetto:** investigation, writing – review and editing. **Ashwani K. Singal:** investigation, writing – review and editing. **Mohamed A. Elfeki:** writing – review and editing, investigation. **Juan Pablo Arab:** conceptualization, writing – review and editing. **Joaquín Cabezas:** investigation, writing – review and editing. **Victor Echavarría:** investigation, writing – review and editing. **Meritxell Ventura Cots:** writing – review and editing, investigation. **María Fátima Higuera-De La Tijera:** investigation, writing – review and editing. **Juan G. Abraldes:** investigation, writing – review and editing. **Mustafa Al-Karaghouli:** investigation, writing – review and editing. **Prasun K. Jalal:** investigation, writing – review and editing. **Mohamad Ali Ibrahim:** investigation, writing – review and editing. **Guadalupe García-Tsao:** investigation, writing – review and editing. **Daniela Goyes:** investigation, writing – review and editing. **Lubomir Skladaný:** investigation, writing – review and editing. **Daniel J. Havaj:** writing – review and editing, investigation. **Karolina Sulejova:** investigation, writing – review and editing. **Svetlana Adamcova Selcanova:** investigation, writing – review and editing. **Diego Rincón:** investigation, writing – review and editing. **Vijay H. Shah:** writing – review and editing. **Patrick S. Kamath:** writing – review and editing. **Marco Arrese:** writing – review and editing. **Ramon Bataller:** writing – review and editing, conceptualization. **Juan Pablo Arab:** conceptualization, writing – review and editing.

Conflicts of Interest

R.B.: consulting for GSK, Novo Nordisk and Boehringer-Ingelheim. J.G.A.: consulting for 89bio, Agomab, Novo Nordisk, Boehringer Ingelheim, AstraZeneca, Terumo, Boston Pharmaceuticals. Grant support: Salix, Gilead, Cook. L.S.: Consulting for Abbvie, Gilead, Astellas, Worwag, ProMed. D.J.H.: Lecturing for Abbvie, Gilead, Astellas, ProMed.

Data Availability Statement

The datasets generated and analysed during the current study are not publicly available but are available from the corresponding author upon reasonable request.

References

1. B. Gao and R. Bataller, "Alcoholic Liver Disease: Pathogenesis and New Therapeutic Targets," *Gastroenterology* 141, no. 5 (2011): 1572–1585.

2. R. Bataller, J. P. Arab, and V. H. Shah, "Alcohol-Associated Hepatitis," *New England Journal of Medicine* 387, no. 26 (2022): 2436–2448, https://doi.org/10.1056/NEJMra2207599.

3. J. P. Arab, J. P. Roblero, J. Altamirano, et al., "Alcohol-Related Liver Disease: Clinical Practice Guidelines by the Latin American Association for the Study of the Liver (ALEH)," *Annals of Hepatology* 18, no. 3 (2019): 518–535.

4. J. P. Arab, L. A. Díaz, N. Baeza, et al., "Identification of Optimal Therapeutic Window for Steroid Use in Severe Alcohol-Associated

Hepatitis: A Worldwide Study," *Journal of Hepatology* 75, no. 5 (2021): 1026–1033.

5. G. Ayares, F. Idalsoaga, L. A. Díaz, J. Arnold, and J. P. Arab, "Current Medical Treatment for Alcohol-Associated Liver Disease," *Journal of Clinical and Experimental Hepatology* 12, no. 5 (2022): 1333–1348.

6. A. Louvet, S. Naveau, M. Abdelnour, et al., "The Lille Model: A New Tool for Therapeutic Strategy in Patients With Severe Alcoholic Hepatitis Treated With Steroids," *Hepatology* 45, no. 6 (2007): 1348–1354.

7. M. Garcia-Saenz-de-Sicilia, C. Duvoor, J. Altamirano, et al., "A Day-4 Lille Model Predicts Response to Corticosteroids and Mortality in Severe Alcoholic Hepatitis," *American Journal of Gastroenterology* 112, no. 2 (2017): 306–315.

8. R. Parker, J. Cabezas, J. Altamirano, et al., "Trajectory of Serum Bilirubin Predicts Spontaneous Recovery in a Real-World Cohort of Patients With Alcoholic Hepatitis," *Clinical Gastroenterology and Hepatology* 20, no. 2 (2022): e289–e297.

9. D. W. Crabb, R. Bataller, N. P. Chalasani, et al., "Standard Definitions and Common Data Elements for Clinical Trials in Patients With Alcoholic Hepatitis: Recommendation From the NIAAA Alcoholic Hepatitis Consortia," *Gastroenterology* 150, no. 4 (2016): 785–790.

10. E. H. Forrest, N. Storey, R. Sinha, et al., "Baseline Neutrophil-To-Lymphocyte Ratio Predicts Response to Corticosteroids and Is Associated With Infection and Renal Dysfunction in Alcoholic Hepatitis," *Alimentary Pharmacology & Therapeutics* 50, no. 4 (2019): 442–453.

11. L. Zou, Y. H. Choi, L. Guizzetti, D. Shu, J. Zou, and G. Zou, "Extending the DeLong Algorithm for Comparing Areas Under Correlated Receiver Operating Characteristic Curves With Missing Data," *Statistics in Medicine* 43, no. 21 (2024): 4148–4162.

12. G. Y. Zou and L. Yue, "Using Confidence Intervals to Compare Several Correlated Areas Under the Receiver Operating Characteristic Curves," *Statistics in Medicine* 32, no. 29 (2013): 5077–5090.

13. N. Vergis, S. R. Atkinson, S. Knapp, et al., "In Patients With Severe Alcoholic Hepatitis, Prednisolone Increases Susceptibility to Infection and Infection-Related Mortality, and Is Associated With High Circulating Levels of Bacterial DNA," *Gastroenterology* 152, no. 5 (2017): 1068–1077.e4.

14. A. Louvet, F. Wartel, H. Castel, et al., "Infection in Patients With Severe Alcoholic Hepatitis Treated With Steroids: Early Response to Therapy Is the Key Factor," *Gastroenterology* 137, no. 2 (2009): 541–548.