Received: 27 March 2025 | Accepted: 9 June 2025

DOI: 10.1097/LVT.0000000000000666

ORIGINAL ARTICLE

OPEN

Prognostic and diagnostic biomarkers in liver transplantation: A systematic review and meta-analysis

Andrea Camera[1] | Tawhidul Islam[1] | Reza Parvan[2] | Søren Erik Pischke[3,4] | Gustavo Jose Justo Silva[1] | Kåre-Olav Stensløkken[1]

[1]Department of Molecular Medicine, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway

[2]Institute for Experimental Medical Research, Oslo University Hospital and University of Oslo, Oslo, Norway

[3]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

[4]Department of Immunology, Clinic for Laboratory Medicine and Department of Anaesthesiology and Intensive Care, Clinic for Emergencies and Critical Care, Oslo University Hospital, Oslo, Norway

Correspondence
Gustavo Jose Justo Silva, Institute of Basic Medical Sciences, University of Oslo, Sognsvannsveien 9, Domus Medica, 0372 Oslo, Norway.
Email: g.j.j.d.silva@medisin.uio.no

Abstract

Liver transplantation (LT) is a therapeutic option for patients suffering from end-stage liver disease. Recent research has probed the prognostic significance of biomarkers to predict graft function and mortality post-transplant, yet few candidates are recommended in clinical practice. We employed a pipeline that integrates meta-analysis (PRISMA 2020), followed by Kaplan–Meier (KM)-based individual patient data (IPD) analysis, aiming to identify potential novel prognostic biomarker panels for LT recipients. Ovid Medline, Embase, and Cochrane databases were searched. Twenty-one prognostic and 8 diagnostic studies were eligible, pooling 34,922 patients. Single biomarkers sampled at an early stage ($\leq$15 d after LT) were significantly associated with graft-related outcomes (HR/OR 0.95 [0.94–0.97]) but did not predict mortality (HR/OR 1.00 [0.97–1.04]) or composite outcomes (HR/OR 1.02 [0.98–1.07]). Biomarkers in combination (GGT/bilirubin ratio, ALT+AST or ALT+AST+bilirubin+INR) predicted composite outcomes (graft failure or mortality, aHR/aOR 4.37 [2.65–7.21]). Biomarkers assessed at late stage ($>$15) did not show association with mortality (HR/OR 1.02 [1.00–1.04]) or composite outcomes (HR/OR 1.00 [0.99–1.01]). KM-based IPD analysis showed that coagulation factor V combined with ALT predicted graft survival (HR 2.12 [1.44–3.12]), and coagulation factor V+insulin-like growth factor 1 stratified the risk of patient survival (HR 2.97 [1.79–4.91]). Therefore, we were able to compare various scoring systems in predicting graft-related outcomes and mortality following LT. Additionally, we identified novel combinations of biomarkers that

Abbreviations: (a)HR, (adjusted) hazard ratio; (a)OR, (adjusted) odds ratio; APRI, Aspartate Aminotransferase-to-Platelet Ratio Index; DBD, donation after brain death; DCD, donation after circulatory death; EAD, early allograft dysfunction; FIB-4, Fibrosis Score 4; IGF-1, insulin-like growth factor 1; INR, international normalized ratio; IPD, Individual Patient Data; KM, Kaplan–Meier; LT, liver transplantation; MEAF, Model for Early Allograft Function; NAFLD, Non-Alcoholic Fatty Liver Disease Fibrosis Score; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies 2; QUIPS, Quality in Prognostic factor Studies; vWF, von Willebrand factor.

Gustavo Jose Justo Silva and Kåre-Olav Stensløkken contributed equally to this article.

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.ltxjournal.com.

exhibited prognostic value for LT patients. Finally, we demonstrate that combined analytical tools for assessing large clinical datasets effectively evaluate multi-modal markers for risk stratification of early and late outcomes for LT.

Keywords: allograft dysfunction, graft failure, mortality, retransplantation

INTRODUCTION

Liver transplantation (LT) is a lifesaving treatment option for patients with severe liver dysfunction that has progressed to end-stage liver disease, including both chronic and acute liver failure.[1,2] In recent years, LT increased worldwide, with 41,099 livers transplanted in 2023, the majority of which still come from deceased donors.[3] In several countries, the pool of organs available for transplantation was expanded with the consolidation of donation after circulatory death (DCD), in addition to classic donation after brain death (DBD).[4] Early allograft dysfunction (EAD) is a common complication post-LT, whose definition is not yet standardized, with a frequency estimated around 15%–30% in recipients of DBD donors, and up to 68% in recipients of DCD donors.[5] The detection of the onset of post-transplant complications has been assisted by diverse biomarkers to enable treatment, as well as to stratify the patients into different risk classes to facilitate resource allocation during the long-term follow-up.[2,6,7] Recent research has probed the prognostic significance of biomarkers in terms of prediction of short-term and long-term outcomes related to graft function and/or mortality post-LT.[7,8] Biomarkers used to assess other liver disorders, such as insulin-like growth factor 1 (IGF-1) and coagulation factor V, could also be effective in prognosticating outcomes after LT, and combining these markers could potentially enhance their prognostic capacity.[9,10]

Various scoring systems have been developed to assess the graft function and identify complications following LT. One such tool is the MELD, which serves as a clinical scoring method to evaluate the severity of chronic liver disease and short-term mortality. The latest version incorporates the measurement of serum sodium alongside previous criteria that include total bilirubin and creatinine levels, as well as international normalized ratio (INR) of prothrombin time,[11] and is used to prioritize candidates on the LT waiting list.[12] Other tools employing the use of panels of clinically available biomarkers (ALT, AST, bilirubin, and INR) was used to define EAD by Olthoff et al.[2] Also, a graded scoring system, Model for Early Allograft Dysfunction (MEAF), based on a similar panel (bilirubin, INR, and ALT) has been proposed for EAD determination.[6]

With the development of high-throughput analytical techniques, candidate biomarkers, whether genetic, proteomic, or metabolomic, will potentially emerge for evaluation in the future.[13–16] Despite the advancements in the LT field, there is a lack of standardization of study populations, measures of effect, and definition of outcomes. It is therefore important to use a meta-analysis approach to consolidate recent findings. In this study, we establish a pipeline based on systematic reviewing and meta-analysis methods to appraise the current state-of-the-art of diagnostic and prognostic biomarkers for liver transplantation in non-rejecting adult recipients receiving whole-graft LT. Here, we evaluate the prognostic performance of several biomarkers, alone or in combination, in predicting the development of complications and adverse outcomes post-LT. Furthermore, we apply KM-based IPD analysis to select novel combinations of biomarkers suitable for further clinical evaluation.

METHODS

Protocol registration and search strategy

This systematic review and meta-analysis were performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020, http://links.lww.com/LVT/A946). The research protocol is registered in the International Prospective Register of Systematic Reviews (PROSPERO, ID CRD42023387074). The search strategy and query were developed with the assistance of the librarian at the University of Oslo (Supplemental Material Appendix 1, http://links.lww.com/LVT/A945). After developing a comprehensive search syntax using controlled vocabulary, 4 databases, including Ovid Medline, Embase, Cochrane Register of Controlled Trials, and Cochrane Database of Systematic Reviews, were searched from inception until February 2023.

Study selection and eligibility criteria

Study screening and selection were performed using Covidence to streamline reviewing activities

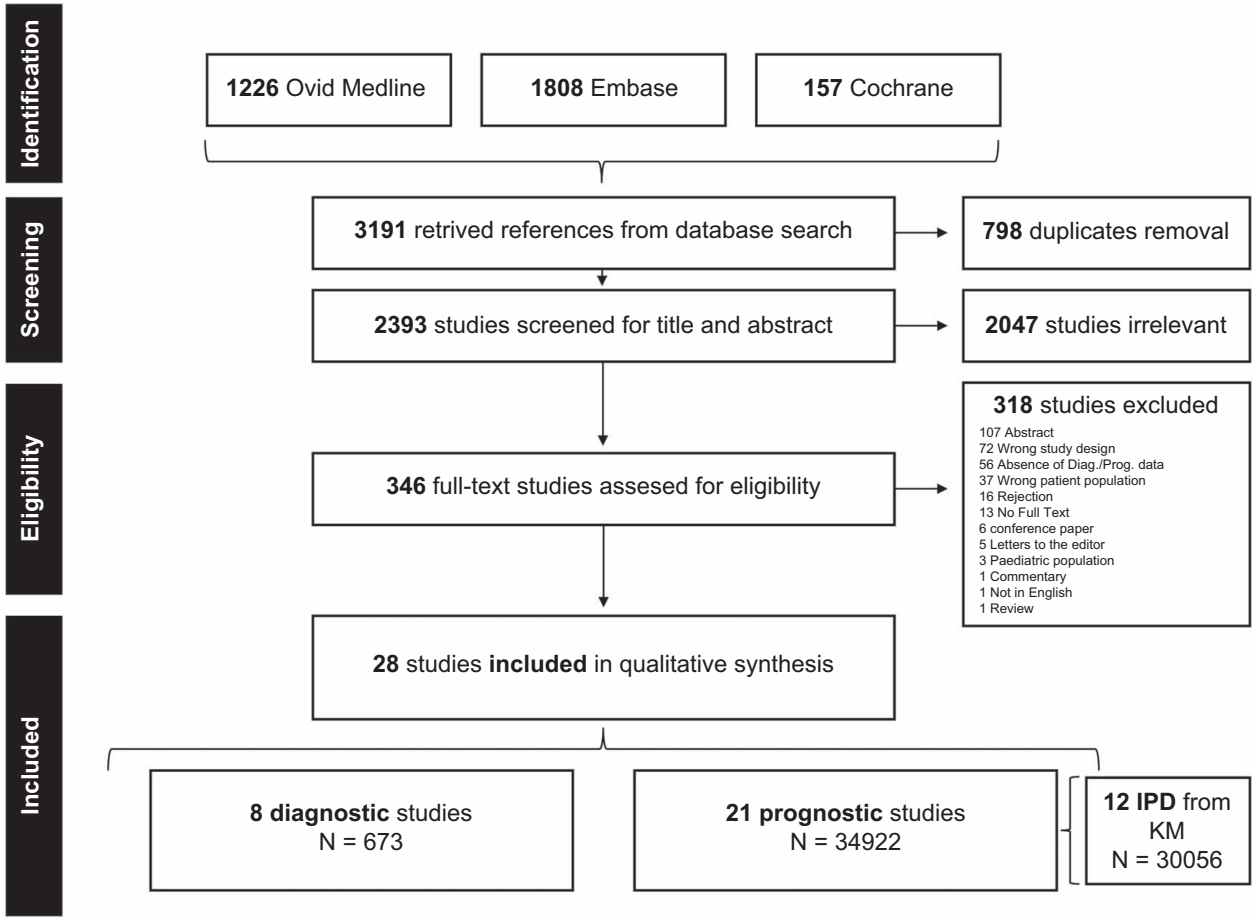

FIGURE 1 PRISMA flow diagram of study selection. Abbreviations: IPD, individual patient data; KM, Kaplan–Meier; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

(Figure 1). Two independent reviewers (Andrea Camera and Gustavo Jose Justo Silva) screened titles and abstracts in a blinded manner, followed by full-text screening according to the eligibility criteria. Disagreements were solved by a third independent reviewer (Tawhidul Islam). In relation to the eligibility criteria, studies were included if the following criteria were met: (1) case–control or cohort study, retrospective or prospective, cross-sectional study; (2) LT; (3) study with at least one biomarker for prognosis after LT, and (4) study should report measures of effect: non-adjusted or adjusted odds ratio (aOR), non-adjusted or adjusted hazard ratio (aHR) for outcomes of graft dysfunction, graft failure, retransplantation, or mortality. Studies were excluded if at least one of the following criteria was met: (1) pediatric population; (2) split LT; (3) multi-organ transplantation; (4) presence of allograft rejection; (5) the population is limited to a single non-graft-related complication (infections, acute and chronic renal injury, and malignancies); (6) abstracts, reviews, systematic reviews, commentaries, letters to the editor, book chapters, and conference papers; and (7) manuscript not in English.

Diagnostic and prognostic risk of bias assessment

The risk of bias was assessed using the Quality in Prognostic factor Studies (QUIPS) tool based on 6 domains: Study Participation, Study Attrition, Prognostic Factor Measurement, Outcome Measurement, Study Confounding, and Statistical Analysis and Reporting.[17] For assessment of biased estimation of diagnostic accuracy, we used the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool, encompassing 4 domains: Patient Selection, Index Test, Reference Standard, and Flow and Timing.[18] QUIPS and QUADAS-2 domains were evaluated respectively within each included prognostic or diagnostic study (Gustavo Jose Justo Silva).

Data extraction

Data extraction was performed by one reviewer (Andrea Camera) using the online systematic review tool Covidence (Melbourne, Australia). The following information was extracted from the included studies: general and demographic information (first author's name, publication

year, country, age, gender, study design) and specific items which are related to the reported prognostic biomarker associated with its measurements of effect. In particular: biomarker name, method of detection, sample type, number of patients, time of collection, cutoff, effect size (aHR or aOR), and corresponding 95% CI, *P*, adjustment factors, follow-up time, and outcome definition. Two of the reviewers (Andrea Camera and Gustavo Jose Justo Silva) assessed the authenticity of the data extracted. Any disagreements were resolved by a third researcher (Reza Parvan).

Outcome definition and data harmonization

All extracted outcomes were revised by two reviewers (Andrea Camera and Gustavo Jose Justo Silva), and terms with equivalent meaning were harmonized under the same name. Disagreements were resolved by a third researcher (Reza Parvan). After achieving agreement, the following outcomes are reported: EAD, primary non-function, initial poor graft function, graft failure, graft survival, Retransplantation, mortality, and survival. Early allograft dysfunction (EAD): reversible dysfunction of the liver graft post-LT, previously associated with adverse outcomes of graft failure and patient death.[6,8] Primary non-function: irreversible failure of a liver graft within the first postoperative week, which results in either retransplantation or death.[6] Initial poor graft function: older definition of impaired graft function, defined by AST and/or ALT >1500 U/L within the first 72 hours, whose clinical features overlap with early allograft dysfunction.[5,19]

Meta-analysis

Aggregated meta-analysis

Statistical analysis was conducted in Stata version 18, inputting pooled aHR and aOR. To estimate effect sizes, we applied either a fixed-effect or a random-effects model using inverse variance or DerSimonian–Laird tests, respectively. Inter-study heterogeneity was evaluated using the Cochran *Q* test and the I^2 statistic.[20] Galbraith plots, leave-one-out plots, and funnel plots were used to evaluate, respectively, the study's heterogeneity, the contribution of each study to the overall data, and the risk of publication bias. Custom Forest plots were generated using the R package (https://github.com/BioAlgorhythm/ForestplotR).

KM-based IPD analysis and risk stratification

Reconstructed individual patient data (IPD) from reported KM curves were analyzed using an established approach.[21] Coordinates from KM curves were extracted using WebPlotDigitizer v5 (https://automeris.io), and reconstruction of IPD was executed using the IPDfromKM tool. The accuracy of data reconstruction was assessed using several metrics: root mean square error (RMSE, ≤ 0.05), mean absolute error (≤ 0.02), maximum absolute error (≤ 0.05), and Kolmogorov–Smirnov test. Each metric obtained from the reconstruction process of each individual KM curve was systematically assessed following the aforementioned thresholds recommended by Liu et al.[21] Reconstructed KM curves that did not meet the thresholds or exhibited discrepancies between the observed and estimated curves were re-extracted.

The risk stratification analysis was conducted by pooling the reconstructed IPD time-to-event information for each biomarker from multiple studies, segregated by the thresholds defined by the original studies. We categorized patients into low-risk or high-risk groups for graft-related, survival, or composite outcomes. Low-risk or high-risk groups were formed with patient data of a single biomarker or a combination of them. We then performed a log-rank test (significant p value <0.05) between low-risk and high-risk groups using GraphPad Prism version 10.

RESULTS

Study selection and risk of bias assessment

As depicted in Figure 1, our search strategy retrieved records from 3 different databases: Ovid Medline ($n=1226$), Embase ($n=1808$), and Cochrane ($n=157$). These 2393 records were assessed based on their Title and Abstract in accordance with inclusion/exclusion criteria, and 346 were selected for full-text assessment. Twenty-one studies with prognostic (Supplemental Table S1, http://links.lww.com/LVT/A944) and 8 studies with diagnostic (Supplemental Table S2, http://links.lww.com/LVT/A944) data were considered eligible for qualitative assessment based on our inclusion/exclusion criteria. The 21 studies with prognostic data were included in an aggregate meta-analysis for 34922 patients. Among the prognostic studies, 12 reported KM curves that were available for KM-based IPD Analysis. Data extracted from the 8 diagnostic studies could not be used for diagnostic aggregate analysis, due to the lack of data affecting all the biomarkers available.

For diagnostic studies, the results from the QUADAS-2 questionnaire (Supplemental Figure S1A, http://links.lww.com/LVT/A944) showed that 62.5% of the included references presented a moderate risk of bias for the Patient Selection. Concerning the Reference Standard, Patient Flow and Timing, and Interpretation of the Index Test domains, all studies have a low risk of bias (Supplemental Table S3, http://links.lww.com/LVT/A944). The 21 included prognostic studies were

assessed for the risk of bias (Supplemental Figure S1B, http://links.lww.com/LVT/A944) using the QUIPS questionnaire. Regarding the risk of bias for Study Confounding, 52.4% of the studies presented a moderate or high risk, and 47.6% for Prognostic Factor Measurement domains. Supplemental Table S4, http://links.lww.com/LVT/A944 describes the QUIPS results per domain for each included study.

Aggregated meta-analysis

The data extracted from all 21 prognostic studies were stratified according to the time of sampling [early ($\leq$15 days post-LT) or late ($>$15 days post-LT) biomarkers] (Supplemental Table S5, http://links.lww.com/LVT/A944). This cutoff is meant to be at a reasonable distance from the defined time intervals for the onset of early graft dysfunction post-LT. We performed subanalysis for the reported outcomes (ie, graft-related, hard endpoints, or composite outcomes). Due to the lack of consistency in the reported effect size, we pooled aHR and aOR for the analysis of the early biomarkers. In this way, we were able to appraise the largest and most comprehensive number of patients and biomarkers reported in the included studies.

Early biomarkers

Single biomarkers

Eight out of the 21 included studies reported prognostic data (2245 patients) on single biomarkers sampled within the first 15 days (or $\leq$15 days) following LT and are addressed here (Figure 2 and Supplemental Figure S2, http://links.lww.com/LVT/A944). Concerning graft-related outcomes, the overall HR/OR was 0.95 (0.94–0.97), heterogeneity $I^2 = 97.8\%$, $H^2 = 44.9$. When we explore the single components, coagulation factor V and creatinine, they were significantly associated with early allograft dysfunction, with respectively HR/OR 0.95 (0.94–0.97) and HR/OR 3.16 (1.70–5.88); lactate showed significant association with initial poor graft function (HR/OR 169 [52.49–544.13]). Concerning mortality, overall HR/OR was 1.00 (0.97–1.04), heterogeneity $I^2 = 79.8\%$, $H^2 = 4.9$. Although the overall HR/OR was not significant, ALT HR/OR 3.16 (1.26–7.94), AST HR/OR 3.45 (1.50–7.92), and coagulation factor V HR/OR 2.60 (1.22–5.52) were associated with mortality after transplantation. Concerning composite outcomes (graft failure, transplantation, or mortality), overall HR/OR was 1.02 (0.98–1.07), heterogeneity $I^2 = 91.7\%$, $H^2 = 12.1$. Here, GlycoTransplantTest, a serum glycemic signature,

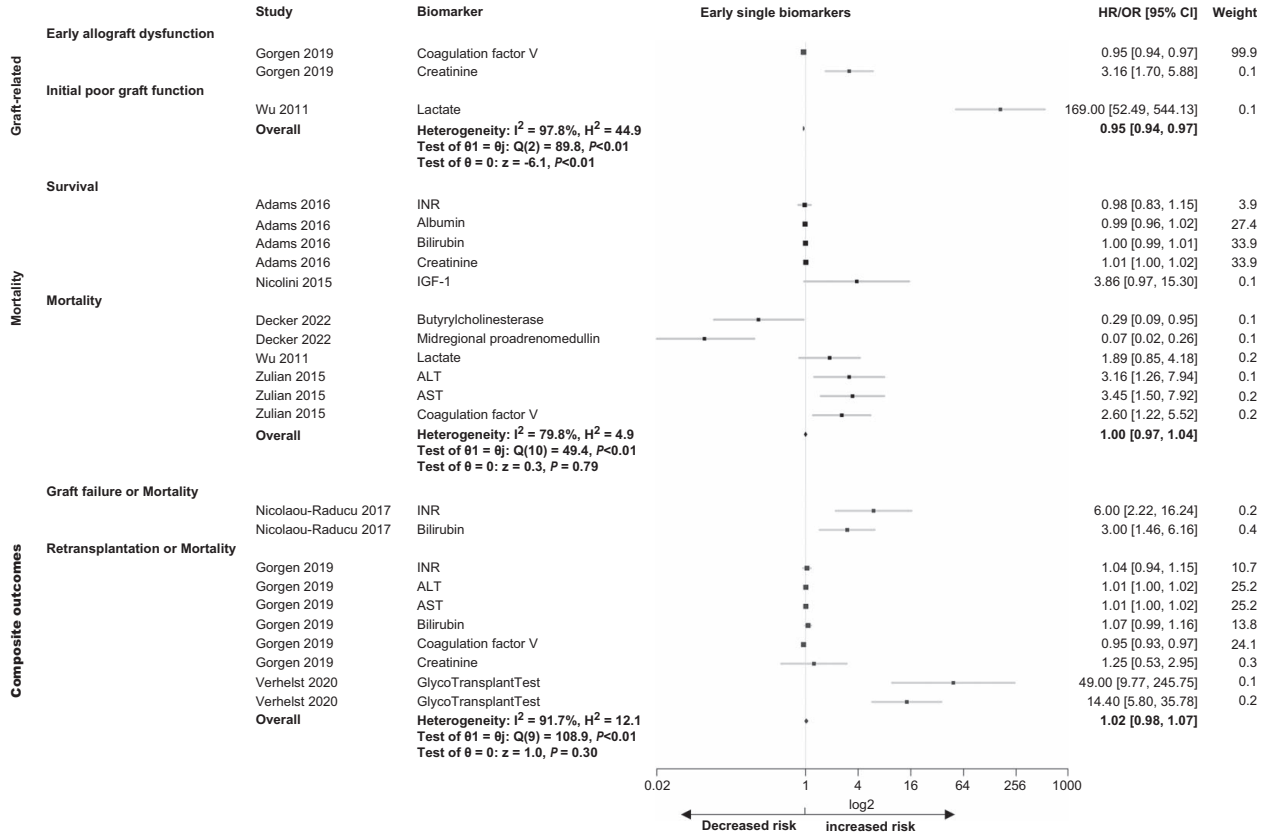

FIGURE 2 Forest plot of prognostic single biomarkers from 8 studies (2245 patients) collected within 15 days from LT assessing graft-related outcomes, mortality, and composite outcomes. The vertical line centered at one expresses no association between biomarker and outcome; values <1 indicate a reduction in risk, >1 an increase in risk.

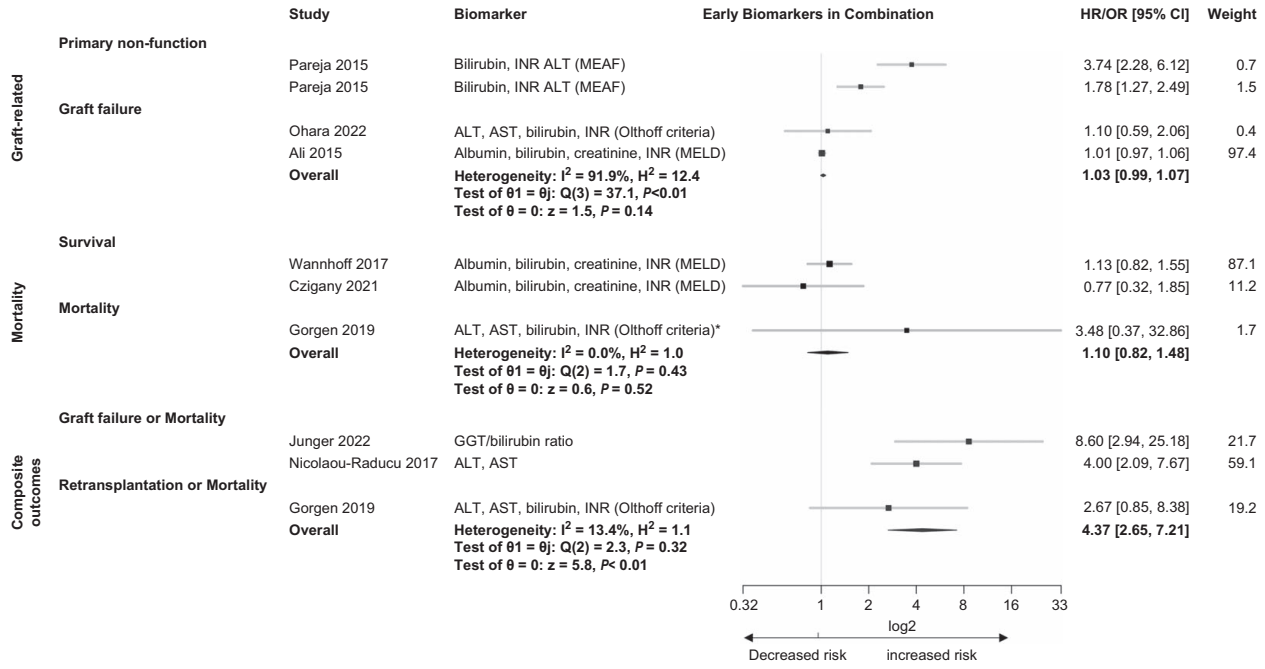

FIGURE 3 Forest plot of prognostic biomarkers in combination from 9 studies (3815 patients) collected within 15 days of LT assessing graft-related outcomes, mortality, and composite outcomes. The vertical line centered at one expresses no association between biomarker and outcome, values <1 a reduction in risk, and >1 an increase in risk.

was associated with risk of composite outcomes (HR/OR 22.00 [7.02–68.93]).

Biomarkers in combination

Nine out of 21 studies assessed different combinations of biomarkers retrieved within the first 15 days (or $\leq$15 days) following LT, accounting for 3815 patients (Figure 3 and Supplemental Figure S3, http://links.lww.com/LVT/A944). Concerning graft-related outcomes, overall aHR/aOR was 1.03 (0.99–1.07), heterogeneity $I^2 = 91.9\%$, $H^2 = 12.4$. In this category, MEAF showed a significant association in the development and validation cohorts, with aHR/aOR 3.74 (2.28–6.12) and 1.78 (1.27–2.49) for primary non-function. For mortality, none of the combinations was significant, with an overall aHR/aOR 1.10 (0.82–1.48), heterogeneity $I^2 = 0.0\%$, $H^2 = 1.0$. Instead, prognostication of composite (graft failure or mortality) outcomes had an overall aHR/aOR 4.37 (2.65–7.21), with a very low heterogeneity $I^2 = 13.4\%$, $H^2 = 1.2$. This was feasible using in combination AST and ALT, revealing aHR/aOR 4.00 (2.09–7.67) or GGT/bilirubin ratio, revealing aHR/aOR 8.60 (2.94–25.18) for composite outcomes.

Late biomarkers

Five out of the 21 studies retrieved biomarkers after the first 15 days post-LT (excluding day 15), accounting for 914 patients. While early single biomarkers and early biomarkers in combination were analyzed separately, all the late biomarkers were assessed together according to the reported outcomes (ie, mortality, survival, graft failure or mortality; retransplantation or mortality) (Figure 4 and Supplemental Figure S4, http://links.lww.com/LVT/A944). Concerning mortality, the overall aHR/aOR was 1.02 (1.00–1.04), heterogeneity $I^2 = 83.0\%$ and $H^2 = 5.9$. Here, Creatinine Excretion Rate aHR/aOR 2.92 (1.47–5.81) and aHR/aOR 1.93 (1.07–3.49), Elastography Grade aHR/aOR 1.11 (1.47–3.60), and NAFLD (glucose, AST, ALT, platelets, and albumin) aHR/aOR 1.90 (1.27–2.84) were significantly associated with risk of death after transplantation (Figure 4). For composite outcomes, retransplantation or mortality and graft failure or mortality were pooled together, with an overall aHR/aOR 1.00 (0.99–1.01), heterogeneity $I^2 = 73.3\%$, $H^2 = 3.7$. Here, APRI (AST and platelets) aHR/aOR 3.10 (1.71–5.63), FIB-4 (AST, ALT, and platelets) aHR/aOR 2.40 (1.30–4.42), and hemoglobin aHR/aOR 3.84 (1.11–13.22) were associated with risk of composite outcomes.

Analysis of a panel of biomarkers with prognostic value for graft and patient survival using KM-based IPD

To further assess the prognostic value of the biomarkers, we here applied computationally reconstructed IPD

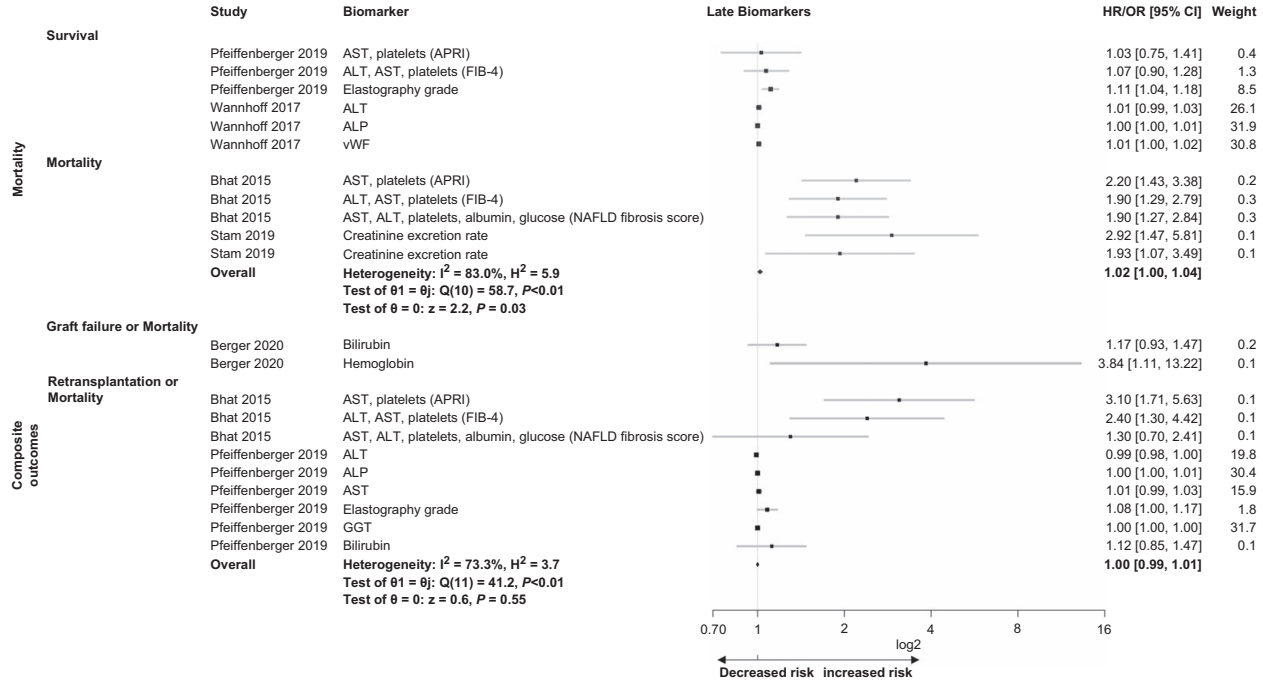

FIGURE 4 Forest plot of prognostic biomarkers in combination from 5 studies (914 patients) collected after 15 days from LT assessing mortality and composite outcomes. The vertical line centered at one expresses no association between biomarker and outcome; values <1 indicate a reduction in risk, >1 an increase in risk.

analysis from the reported KM curves. Six out of the 12 prognostic studies reported KM curves on early biomarkers for 2 different outcomes: graft survival or patient survival. Reconstructed patient data (time-to-event) from different studies assessing the same outcome or composite outcomes were then pooled together as low-risk or high-risk groups. We then generated a matrix of IPD data for graft survival (EAD, ALT, coagulation factor V, HLA-DR13) and patient survival (EAD, coagulation factor V, IGF-1), and assessed all the possible permutations of 4, 3, 2, or 1 biomarker at the time (Figure 5A).

Graft survival

When pooling all 4 biomarkers (EAD+ALT+coagulation factor V+HLA-DR13), we were able to predict graft failure, HR 1.06 [1.00–1.11], log-rank $p = 0.03$. Furthermore, the 3-biomarker panel composed of ALT+coagulation factor V+HLA-DR13 showed association with graft failure, HR 1.098 (1.04–1.16), log-rank $p < 0.001$. EAD+coagulation factor V+HLA-DR13 was still mildly associated with graft failure HR 1.05 (1.00–1.11), log-rank $p = 0.04$. ALT+HLA-DR13 showed a significant association, HR 1.08 (1.03–1.14), log-rank $p < 0.01$. Coagulation factor V+HLA-DR13 was associated with graft failure, HR 1.10 (1.04–1.16), log-rank $p < 0.001$. Remarkably, the combination of coagulation factor V +ALT was significantly associated with graft failure HR 2.12 (1.44–3.12), with a log-rank $p < 0.001$ (Figure 5B). Of note, coagulation factor V alone displayed an increased risk of graft failure HR 3.02 (1.75–5.20), log-rank $p < 0.001$ (Supplemental Figure S5A, http://links.lww.com/LVT/A944). Liver transplanted patients that presented the HLA-DR13 haplotype presented an increased risk of graft failure, HR 1.08 (1.02–1.14) with a log-rank $p < 0.01$ (Supplemental Figure S5D, http://links.lww.com/LVT/A944). Individually, the other biomarkers (eg, ALT and EAD) were unable to show association with graft failure (Supplemental Figures S5B, C, http://links.lww.com/LVT/A944).

Patient survival

In contrast with graft failure, the 3-biomarker panel composed of EAD+coagulation factor V+IGF-1 did not predict mortality (HR 1.12 [0.84–1.50], log-rank $p = 0.44$) (Figure 5A). Interestingly, the combination of coagulation factor V and IGF-1 presented a significant association with mortality (HR 2.97 [1.79–4.91], log-rank $p < 0.001$) (Figures 5A–C). Coagulation factor V alone was already associated with survival (HR 2.89 [1.68–4.98], log-rank $p < 0.001$) (Supplemental Figure S5E, http://links.lww.com/LVT/A944), but not IGF-1 (HR 3.86 [0.97–15.30], the log-rank $p = 0.04$) (Supplemental Figure S5F, http://links.lww.com/LVT/A944). EAD alone did not show an association with mortality (aHR 0.97 [0.65–1.45], log-rank $p = 0.02$) (Supplemental Figure S5G, http://links.lww.com/LVT/A944).

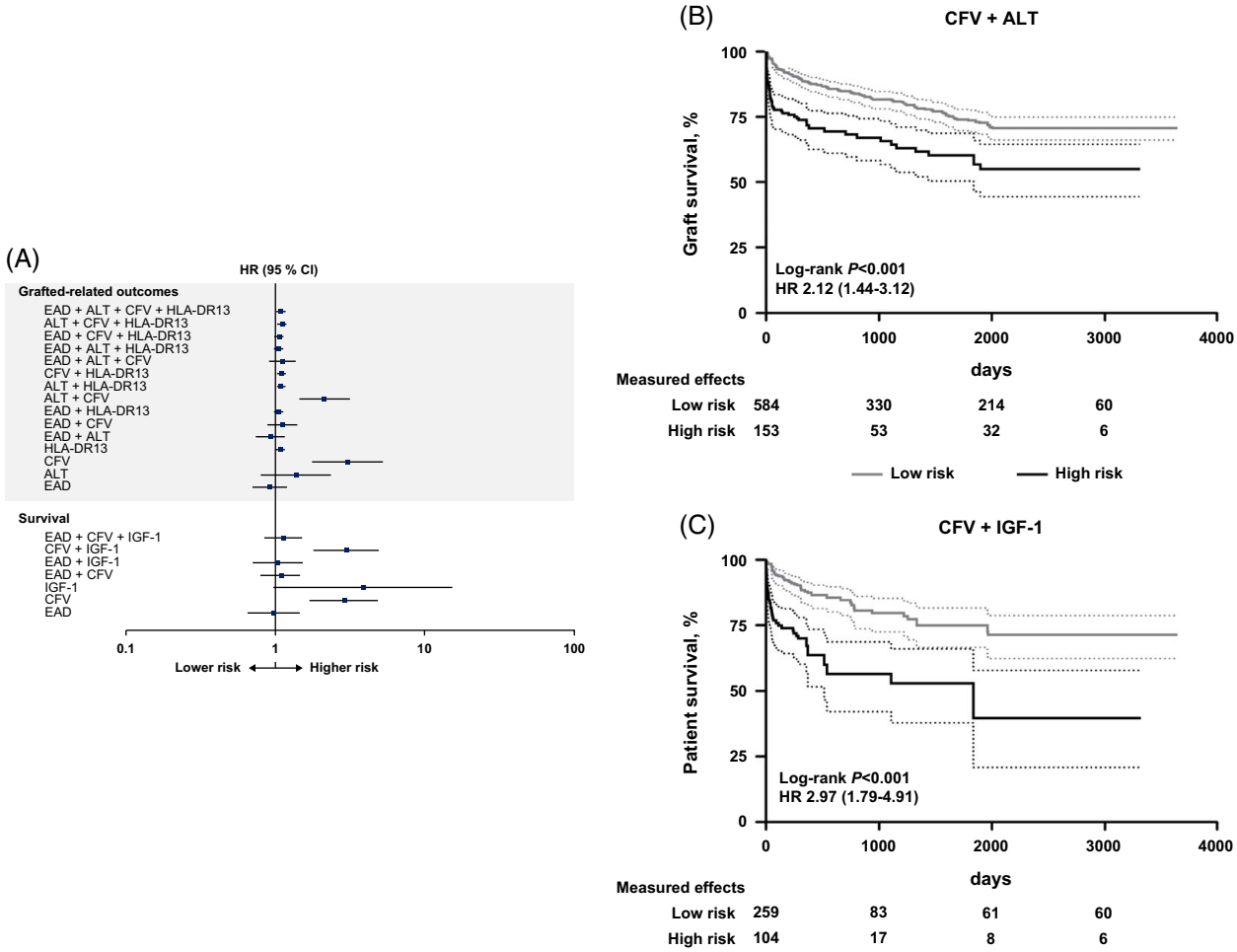

FIGURE 5 Prognostic potential of identified panels of biomarkers in liver transplanted patients. Forest plot for HR (95% CI) (A) for different combinations of biomarkers associated with graft-related or survival outcomes. Kaplan–Meier curves the combination of coagulation factor V+ALT for graft survival (B) and coagulation factor V+IGF-1 (C) for patient survival.

DISCUSSION

This meta-analysis evaluated the prognostic performance of biomarkers to predict EAD and hard outcomes (ie, graft failure, retransplantation, or mortality) in LT patients. The aggregated meta-analysis of 21 studies (34922 patients) showed that single biomarkers sampled ≤ 15 days post-LT were significantly associated with graft-related outcomes but did not predict mortality or composite outcomes (graft failure, retransplantation, or mortality). The assessment of biomarkers in combination (GGT/bilirubin ratio, ALT+AST, or ALT+AST+bilirubin+INR) predicted the risk of developing composite outcomes (graft failure or mortality). Biomarkers sampled at late stage (>15 days post-LT) did not show association with mortality or composite outcomes. Furthermore, we performed a computational reconstruction of individual patient data from KM curves on 12 of the studies (30056 patients). Serving as a complement to the aggregated results, the IPD-based KM analysis corroborates the significance of coagulation factor V as a prognostic marker for both graft and patient survival when evaluated individually, as well as in combination with ALT for graft survival and IGF-1 for patient survival. Late biomarkers measured from day 15 onward after LT showed association with mortality and graft failure for single biomarkers such as hemoglobin, elastography grade (evaluated by ultrasound), and creatinine excretion. Serum biomarker combination meant to assess liver fibrosis (NFALD, FIB-4, and APRI) showed an association with retransplantation or mortality.

Risk scores to predict outcomes in liver transplantation

The application of scoring systems based on multiple biomarkers has become a hallmark of modern outcome prediction, particularly for estimating disease progression, response to treatment, or mortality rates.[16,22,23] In the field of LT, various tools have been developed for graft assessment, including the MELD, MEAF, Olthoff criteria, and others, significantly transforming clinical practices. These tools are not only relevant for diagnosing and predicting graft-

related outcomes but have also demonstrated the ability to predict hard endpoints such as retransplantation and mortality.[12] However, challenges remain in accurately performing risk stratification with precision in LT patients.

The MELD score, in use since 2002, evaluates the severity of chronic liver disease severity by considering total bilirubin, creatinine and sodium, and INR. It is primarily employed to prioritize waiting lists for LT, and has shown predictive value for short-term mortality.[12] Over the years, the MELD score has undergone several modifications to enhance accuracy and reduce bias[24] Other scoring systems focus on assessing liver dysfunction, a condition most often defined by the Olthoff criteria of EAD. EAD is a major complication of LT and serves as a surrogate marker for post-LT complications.[6,7,25] In this condition, the injured, dysfunctional liver maintains a certain level of function that is sufficient for life.[6] Although no uniform definition has been established, the most accepted description of this nosocomial entity is proposed by Olthoff et al[2] as AST or ALT >2000 U/L within the first week after LT, total bilirubin >10 mg/dL, or INR >1.6 on post-LT day 7.[26]

In contrast, the more recent Model for Early Allograft Function (MEAF) criteria, unlike the Olthoff criteria, define EAD on a continuous scale calculated on ALT, INR, and bilirubin levels.[6] Whereas multiple studies suggest EAD is associated with adverse outcomes after transplantation, within our data, there is no difference in graft and patient survival between the patients with and without EAD when graft dysfunction was diagnosed with the Olthoff criteria.[2,7] Among the scoring systems evaluated in this meta-analysis, which includes MELD and Olthoff criteria, MEAF was the only model capable of effectively stratifying the risk of developing graft-related outcomes such as primary non-function and graft failure. In assessing hard clinical endpoints, the Olthoff criteria did not outperform the MELD score in terms of predicting mortality following LT. This finding is not entirely unexpected, given the numerous factors that influence LT mortality in the postoperative setting. Indeed, Jochmans et al[27] directly compared the predictive performance of MEAF and Olthoff criteria, finding that the Harrell C-index statistical measure for evaluating the predictive accuracy of prognostic models fell below the necessary threshold required for efficient and reliable mortality in LT. Furthermore, the authors recommended that these scoring tools should be used to complement clinical decision-making rather than dictate it. Interestingly, our aggregated meta-analysis data indicated that the Olthoff criteria, along with GGT/bilirubin ratio, and ALT and AST levels, performed well in stratifying the risk of composite outcomes such as graft failure or retransplantation, or mortality.

IPD as a further step toward the discovery of new biomarker combinations

We have also appraised the clinical prognostic evidence for multiple biomarkers available in the liver transplantation literature, many of which have been repurposed from their original intended use in hepatology.[6,25,26,28–34] Our approach involved a combination of different analytical tools, spanning from aggregated data meta-analysis to assessment of various biomarkers—both individually or in combination—across a large cohort of LT patients. In addition, we conducted a meta-analysis on time-to-event data at the IPD level.

We employed a novel strategy to reconstruct IPD from KM curves, which enables the extraction of time-to-event data at the patient level for each biomarker and evaluates the overall pooled IPD prognostic performance, and even proposes the combinations of high-performance biomarkers. Notably, some combinations of biomarkers were able to predict outcomes after transplantation: in particular, pooled data for coagulation factor V+ALT predict graft survival with a 95% CI that is narrower than that of coagulation factor V alone. The combination of coagulation factor V+IGF-1 predicts patient survival. These findings suggest novel combinations of biomarkers for outcome prediction, different from the graft dysfunction diagnostic tools (eg, Olthoff criteria or MEAF) previously investigated in the included studies.

Coagulation factor V

Coagulation factor V is a protein synthesized by the liver involved in the coagulation cascade. With its short half-life, coagulation factor V reflects liver function directly, making it a potential candidate for early detection of hepatic dysfunction and associated long-term outcomes, regardless of the presence of clinically manifest EAD.[26,35,36] Another meta-analysis evaluated coagulation factor V as a potential diagnostic biomarker for EAD; however, in that study, the association with long-term outcomes such as graft or patient survival was not explored. Our combined aggregated and IPD results suggest that this biomarker is significant for the determination of graft and patient survival, which is also supported by recent findings.[37] Coagulation factor V should be further investigated in larger clinical studies on graft and patient survival.

Insulin-like growth factor 1

Serum IGF-1, synthesized by the liver after growth hormone stimulation, mediates complex endocrine regulation. The impairment of its biological pathways is linked to the pathophysiology of diverse conditions

such as cardiovascular diseases and cancer.[9,38–40] Despite the endocrine interactions that influence the synthesis, fluctuations in the hormone level can be associated with dysfunction in cirrhotic livers and, for this reason, the relationship with outcomes post-LT was investigated.[9,10] However, both our aggregate meta-analysis and KM-based IPD results did not show a significant association with mortality; however, this could be due to the small sample size. In our IPD analysis, the combination of coagulation factor V and IGF-1 was significantly associated with patient survival; however, larger clinical studies are required to further investigate the prognostic significance of this biomarker.

At a later stage, our aggregated data meta-analysis revealed that the assessment of biomarkers modestly predicted mortality following LT (HR/OR 1.02 [1.00–1.04]). In contrast, scoring systems designed to estimate liver fibrosis (APRI, FIB-4, NAFLD, and elastography),[29,33] along with urinary creatinine excretion rate (used as an indicator of muscle wasting during the postoperative period)[41] showed promising performance in stratifying risk of mortality after LT.

Fibrosis scores

Non-alcoholic fatty liver disease fibrosis score (NAFLD), fibrosis score 4 (FIB-4), and aspartate aminotransferase-to-platelet ratio index (APRI) are used as noninvasive combined markers to stage liver fibrosis, a common complication of LT. Each of these fibrosis scores combines a specific set of serum biomarkers in an equation to yield a risk value: AST, ALT, platelet count, glucose, and albumin in NAFLD, AST, ALT, and platelet count in FIB-4, and AST and platelet count in APRI.[29] We report 2 studies, Bhat et al[29] and Pfeiffenberger et al,[33] assessing whether biomarkers normally employed diagnostically for liver-fibrosis staging could prognosticate hard endpoints such as mortality or retransplantation. Additionally, it has recently emerged that the use of noninvasive imaging-based methods, such as liver elastography, to assess liver stiffness as a surrogate quantitative marker for fibrosis.[42] Performed by using ultrasound or MRI, these methods have proven to be convenient and reproducible in the diagnosis of chronic liver diseases, but also in predicting clinical outcomes. Here, we showed that even when assessed on a later stage after LT, liver elastography grade predicts mortality alone and combined outcomes (eg, retransplantation or mortality).[33] In general, liver fibrosis has a considerable impact on outcomes after transplantation, particularly increasing the risk of mortality and graft failure. Recipients with fibrotic grafts exhibited a higher occurrence of specific causes of death, including cancer and cardiovascular disease, highlighting the need for additional clinical studies to refine donor selection criteria. Future investigations should focus on understanding the underlying mechanisms of fibrosis associated with transplantation and developing approaches to mitigate its adverse effects and, ultimately, improving long-term survival rates and graft function for individuals undergoing liver transplantation.[43]

Study strengths and limitations

The strengths of this study include a comprehensive systematic review of the literature across diverse databases, the criterion-based inclusion of relevant studies, and thorough quality assessment using different questionnaires. This approach allowed us to perform a meta-analysis in more studies and patients. We successfully assessed multiple biomarkers, alone or in combination, collected at different stages of post-LT follow-up. We focused our analysis on the recipient patients, excluding studies that investigated markers from the donors, as well as rejection-related biomarkers. Furthermore, we analyzed the prognostic and diagnostic capacity of the included biomarkers for different outcomes relevant for LT (eg, graft-related outcomes, retransplantation, and mortality). Complementary, we employed a method to extract time-to-event data from individual patients from KM curves, allowing us to assess the prognostic performance of a specific panel of biomarkers. The integration of all those approaches confers one of the strengths of our meta-analysis.

However, our investigation has some limitations that reflect the status of the literature. First, most of the included studies are retrospective ($n = 15$, 53.6%), single-center cohorts ($n = 25$, 89.3%), performed on a small number of patients, mostly coming from Europe and North America (75.0%), conferring limited generalizability. Additionally, we observed substantial heterogeneity associated with the overall aggregated data meta-analysis. The main cause is the lack of consistency when reporting the measure of effect (aHR or aOR), which forced us to analyze together a limited percentage of the data available, hence reducing the sample and study size. Last, the interpretation of the results on the identified panels derived from the KM-based IPD analysis warrants caution and should be the subject of further clinical validation.

CONCLUSIONS

We showcase the functionality of a pipeline of analytical tools developed to assess large, aggregate, and/or individual patient datasets of biomarkers, aiming to develop a prognostic panel of multi-modal markers with

clinical relevance for risk stratification of early and late outcomes in LT patients. We were able to identify biomarkers that could be used for the prediction of graft-related outcomes as well as mortality after LT. Here, not only serum biomarkers could be included in a potential outcome prediction algorithm, but also imaging data such as ultrasound-based elastography grade, genetic markers such as the HLA-DR13 haplotype, as well as frailty and sarcopenia metrics.[33,44,45] Nevertheless, to report results with low risk of bias, we stress the importance of standardization of biomarker cutoffs and units for the new clinical studies in the field. Also, we deem it necessary that new clinical studies will provide complete information on missing data and participant dropouts as well as well as the criteria used to identify the eligible population, which is critical when addressing the differences in outcomes between circulatory and brain death donors.

AUTHOR CONTRIBUTIONS

Andrea Camera, Reza Parvan, Gustavo Jose Justo Silva, and Kåre-Olav Stensløkken participated in the study conception and design. Andrea Camera and Gustavo Jose Justo Silva collected the data. Reza Parvan performed statistical analyses. Andrea Camera, Reza Parvan, Tawhidul Islam, Søren Erik Pischke, Gustavo Jose Justo Silva, and Kåre-Olav Stensløkken interpreted the data. Andrea Camera, Gustavo Jose Justo Silva, and Kåre-Olav Stensløkken drafted the manuscript. Andrea Camera, Reza Parvan, Tawhidul Islam, Søren Erik Pischke, Gustavo Jose Justo Silva, and Kåre-Olav Stensløkken participated in the critical revision. Gustavo Jose Justo Silva and Kåre-Olav Stensløkken supervised the project. All authors critically reviewed and revised the manuscript draft and approved the final version for submission.

FUNDING INFORMATION

Research reported in this publication was supported by the UiO:Life Science Convergence Environments to Kåre-Olav Stensløkken and by the Norwegian Association for Public Health to Gustavo Jose Justo Silva (Grant No. 19560), Marie Skłodowska-Curie Actions (Scientia Fellows) to Tawhidul Islam (Grant No. 801133).

ACKNOWLEDGMENTS

We would like to express our gratitude to the librarian Marie Susanna Isachsen (University of Oslo) for the invaluable assistance in developing the search strategy. We also extend our thanks to Yousef Moradi (Social Determinants of Health Research Center, Research Institute for Health Development, Kurdistan University of Medical Sciences, Iran) for the critical support and expertise in the data analysis.

CONFLICTS OF INTEREST

The authors have no conflicts to report.

REFERENCES

1. Adam R, Karam V, Cailliez V, O Grady JG, Mirza D, Cherqui D, et al. 2018 Annual Report of the European Liver Transplant Registry (ELTR)—50-year evolution of liver transplantation. Transpl Int. 2018;31:1293–317.
2. Olthoff KM, Kulik L, Samstein B, Kaminski M, Abecassis M, Emond J, et al. Validation of a current definition of early allograft dysfunction in liver transplant recipients and analysis of risk factors. Liver Transpl. 2010;16:943–9.
3. Organización Nacional de Trasplantes W. International Report on Organ Donation and Transplantation Activities 2023; 2024.
4. Ziogas IA, Kakos CD, Esagian SM, Skarentzos K, Alexopoulos SP, Shingina A, et al. Liver transplant after donation from controlled circulatory death versus brain death: A UNOS database analysis and publication bias adjusted meta-analysis. Clin Transplant. 2022;36:e14521.
5. Masior L, Grat M. Primary nonfunction and early allograft dysfunction after liver transplantation. Dig Dis. 2022;40:766–76.
6. Pareja E, Cortes M, Hervas D, Mir J, Valdivieso A, Castell JV, et al. A score model for the continuous grading of early allograft dysfunction severity. Liver Transpl. 2015;21:38–46.
7. Pollok JM, Tinguely P, Berenguer M, Niemann CU, Raptis DA, Spiro M; ERAS4OLT collaborative. Enhanced recovery for liver transplantation: Recommendations from the 2022 International Liver Transplantation Society consensus conference. Lancet Gastroenterol Hepatol. 2023;8:81–94.
8. Lee DD, Croome KP, Shalev JA, Musto KR, Sharma M, Keaveny AP, et al. Early allograft dysfunction after liver transplantation: An intermediate outcome measure for targeted improvements. Ann Hepatol. 2016;15:53–60.
9. Gariani K, Toso C, Philippe J, Orci LA. Effects of liver transplantation on endocrine function: A systematic review. Liver Int. 2016;36:1401–11.
10. Nicolini D, Mocchegiani F, Palmonella G, Coletta M, Brugia M, Montalti R, et al. Postoperative insulin-like growth factor 1 levels reflect the graft's function and predict survival after liver transplantation. PLoS One. 2015;10:e0133153.
11. Nagai S, Chau LC, Schilke RE, Safwan M, Rizzari M, Collins K, et al. Effects of allocating livers for transplantation based on model for end-stage liver disease-sodium scores on patient outcomes. Gastroenterology. 2018;155:1451–62.e3.
12. Lucey MR, Furuya KN, Foley DP. Liver transplantation. N Engl J Med. 2023;389:1888–900.
13. Attard JA, Dunn WB, Mergental H, Mirza DF, Afford SC, Perera M. Systematic review: Clinical metabolomics to forecast outcomes in liver transplantation surgery. OMICS. 2019;23:463–76.
14. Koch PF, Ludwig K, Krenzien F, Hillebrandt KH, Schoning W, Pratschke J, et al. miRNA as potential biomarkers after liver transplantation: A systematic review. Transplant Rev (Orlando). 2024;38:100831.
15. Lopez-Lopez V, Perez-Sanz F, de Torre-Minguela C, Marco-Abenza J, Robles-Campos R, Sanchez-Bueno F, et al. Proteomics in liver transplantation: A systematic review. Front Immunol. 2021;12:672829.
16. Song W, Xiong X, Ge W, Zhu H. Prognostic value of protein biomarkers in liver transplantation: A systematic review. Proteomics Clin Appl. 2022;16:e2100038.
17. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013;158:280–6.
18. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155:529–36.
19. Nanashima A, Pillay P, Verran DJ, Painter D, Nakasuji M, Crawford M, et al. Analysis of initial poor graft function after orthotopic liver transplantation: Experience of an Australian

single liver transplantation center. Transplant Proc. 2002;34:1231–5.
20. DerSimonian R, Laird N. Meta-analysis in clinical trials revisited. Contemp Clin Trials. 2015;45(Pt A):139–45.
21. Liu N, Zhou Y, Lee JJ. IPDfromKM: reconstruct individual patient data from published Kaplan–Meier survival curves. BMC Med Res Methodol. 2021;21:111.
22. Liu D, Han Y, Liu A. Marginal, conditional, and pseudo likelihood ratio approaches for biomarker combination to predict a binary disease outcome. Stat Med. 2022;41:2574–85.
23. Norman JS, Li PJ, Kotwani P, Shui AM, Yao F, Mehta N. AFP-L3 and DCP strongly predict early hepatocellular carcinoma recurrence after liver transplantation. J Hepatol. 2023;79:1469–77.
24. Kim WR, Mannalithara A, Heimbach JK, Kamath PS, Asrani SK, Biggins SW, et al. MELD 3.0: The Model for End-Stage Liver Disease Updated for the Modern Era. Gastroenterology. 2021;161:1887–95 e4.
25. Ohara S, Macdonough E, Egbert L, Brooks A, Lizaola-Mayo B, Mathur AK, et al. Decreasing significance of early allograft dysfunction with rising use of nonconventional donors. Medicina. 2022;58:821.
26. Gorgen A, Prediger C, Prediger JE, Chedid MF, Backes AN, De Araujo A, et al. Serum Factor V is a continuous biomarker of graft dysfunction and a predictor of graft loss after liver transplantation. Transplantation. 2019;103:944–51.
27. Jochmans I, Fieuws S, Monbaliu D, Pirenne J. "Model for early allograft function" outperforms "early allograft dysfunction" as a predictor of transplant survival. Transplantation. 2017;101:e258–64.
28. Ali JM, Davies SE, Brais RJ, Randle LV, Klinck JR, Allison ME, et al. Analysis of ischemia/reperfusion injury in time-zero biopsies predicts liver allograft outcomes. Liver Transpl. 2015;21:487–99.
29. Bhat M, Ghali P, Rollet-Kurhajec KC, Bhat A, Wong P, Deschenes M, et al. Serum fibrosis biomarkers predict death and graft loss in liver transplantation recipients. Liver Transpl. 2015;21:1383–94.
30. Czigany Z, Kramp W, Lurje I, Miller H, Bednarsch J, Lang SA, et al. The role of recipient myosteatosis in graft and patient survival after deceased donor liver transplantation. J Cachexia, Sarcopenia Muscle. 2021;12:358–67.
31. Junger H, Muhlbauer M, Brennfleck FW, Schurr LA, Goetz M, Eggenhofer E, et al. Early gammaGT and bilirubin levels as biomarkers for regeneration and outcomes in damaged bile ducts after liver transplantation. Clin Transplant. 2023;37:e14880.
32. Nicolau-Raducu R, Cohen AJ, Bokhari A, Bohorquez H, Bruce D, Carmody I, et al. Predictive model and risk factors associated with a revised definition of early allograft dysfunction in liver transplant recipients. Clin Transplant. 2017;31:1–8.
33. Pfeiffenberger J, Hornuss D, Houben P, Wehling C, Von Haken R, Lozanovski, et al. Routine liver elastography could predict actuarial survival after liver transplantation. J Gastrointestin Liver Dis. 2019;28:271–7.
34. Wannhoff A, Rauber C, Friedrich K, Rupp C, Stremmel W, Weiss KH, et al. Von Willebrand factor and alkaline phosphatase predict re-transplantation-free survival after the first liver transplantation. United European Gastroenterol J. 2017;5:86–93.
35. Sanchez-Gonzalez C, Fernandez Aguilar JL, Sanchez Perez B, Munoz MAS, Daga JAP, Reyes MP, et al. Primary graft dysfunction: Factor V's value for its early diagnosis. Transplant Proc. 2022;54:2531–4.
36. Zulian MC, Chedid MF, Chedid AD, Filho TJMG, Leipnitz I, De Araujo A, et al. Low serum factor V level: Early predictor of allograft failure and death following liver transplantation. Langenbecks Arch Surg. 2015;400:589–97.
37. Gravely AK, Claasen M, Ivanics T, Winter E, Peralta P, Selzner M, et al. Factor V Serves as an Early Biomarker for Graft Loss After Liver Transplant: A Prospective Evaluation. Clin Transplant. 2025;39:e70086.
38. Rahmani J, Montesanto A, Giovannucci E, Zand H, Barati M, Kopchick JJ, et al. Association between IGF-1 levels ranges and all-cause mortality: A meta-analysis. Aging Cell. 2022;21:e13540.
39. Delafontaine P, Song YH, Li Y. Expression, regulation, and function of IGF-1, IGF-1R, and IGF-1 binding proteins in blood vessels. Arterioscler Thromb Vasc Biol. 2004;24:435–44.
40. Ungvari Z, Csiszar A. The emerging role of IGF-1 deficiency in cardiovascular aging: recent advances. J Gerontol A Biol Sci Med Sci. 2012;67:599–610.
41. Stam SP, Oste MCJ, Eisenga MF, Blokzijl H, van den Berg AP, Bakker SJL, et al. Posttransplant muscle mass measured by urinary creatinine excretion rate predicts long-term outcomes after liver transplantation. Am J Transplant. 2019;19:540–50.
42. Safi K, Pawlicka AJ, Pradhan B, Sobieraj J, Zhylko A, Struga M, et al. Perspectives and tools in liver graft assessment: A transformative era in liver transplantation. Biomedicines. 2025;13:494.
43. Bian C, Luo J, Tang R, Huang H, Zeng Z. The impact of donor liver fibrosis on the outcomes of patients who undergo liver transplant: A cohort study from UNOS database. Front Med (Lausanne). 2025;12:1547200.
44. Futagawa Y, Waki K, Cai J. The association of HLA-DR13 with lower graft survival rates in hepatitis B and primary sclerosing cholangitis Caucasian patients receiving a liver transplant. Liver Transpl. 2006;12:600–4.
45. Park B, Bhat S, Xia W, Barazanchi AWH, Frampton C, Hill AG, et al. Consensus-defined sarcopenia predicts adverse outcomes after elective abdominal surgery: Meta-analysis. BJS Open. 2023;7:zrad065.

How to cite this article: Camera A, Islam T, Parvan R, Pischke SE, Silva GJJ, Stensløkken K. Prognostic and diagnostic biomarkers in liver transplantation: A systematic review and meta-analysis. Liver Transpl. 2025;31:1499–1510. https://doi.org/10.1097/LVT.0000000000000666