## ORIGINAL ARTICLE

OPEN

# Validation of risk scores for allograft failure after liver transplantation in Germany: a retrospective cohort analysis

**Simon Moosburner**[1,2] | **Leke Wiering**[3] | **Nathalie N. Roschke**[1] | **Axel Winter**[1] | **Münevver Demir**[3] | **Joseph M.G.V. Gaßner**[1,2] | **Maximilian Zimmer**[1] | **Paul Ritschl**[1,2] | **Brigitta Globke**[1,2] | **Georg Lurje**[1] | **Frank Tacke**[3] | **Wenzel Schöning**[1] | **Johann Pratschke**[1] | **Robert Öllinger**[1] | **Igor M. Sauer**[1] | **Nathanael Raschzok**[1,2]

[1]Department of Surgery, Experimental Surgery, Campus Charité Mitte|Campus Virchow-Klinikum, Charité—Universitätsmedizin Berlin, Universität zu Berlin and Berlin Institute of Health

[2]BIH Charité Clinician Scientist Program, Berlin Institute of Health (BIH)

[3]Department of Hepatology and Gastroenterology, Campus Charité Mitte|Campus Virchow-Klinikum, Charité—Universitätsmedizin Berlin, Universität zu Berlin and Berlin Institute of Health, Berlin, Germany

**Correspondence**
Simon Moosburner, Department of Surgery, Campus Charité Mitte|Campus Virchow-Klinikum, Charité—Universitätsmedizin Berlin, Augustenburger Platz 1, Berlin 13353, Germany.
Email: simon.moosburner@charite.de

## Abstract

A growing number of clinical risk scores have been proposed to predict allograft failure after liver transplantation. However, validation of currently available scores in the Eurotransplant region is still lacking. We aimed to analyze all clinically relevant donor and recipient risk scores on a large German liver transplantation data set and performed a retrospective cohort analysis of liver transplantations performed at the Charité—Universitätsmedizin Berlin from January 2007 until December 2021 with organs from donation after brain death. We analyzed 9 previously published scores in 906 liver transplantations [Eurotransplant donor risk index (ET-DRI/DRI), donor age and model for end-stage liver disease (D-MELD), balance of risk (BAR), early allograft dysfunction (EAD), model for early allograft function (MEAF), liver graft assessment following transplantation (L-GrAFT$_7$), early allograft failure simplified estimation (EASE), and a score by Rhu and colleagues). The EASE score had the best predictive value for 3-month, 6-month, and 12-month graft survival with a *c*-statistic of 0.8, 0.77,

---

**Abbreviations**: AST/ALT, aspartate- and alanine-aminotransferase; BAR, balance of risk score; DBD, donation after brain death; DCD, donation after circulatory death; D-MELD, donor age and model for end-stage liver disease; DRI, donor risk index; EAD, early allograft dysfunction; EASE, early allograft failure simplified estimation; ET-DRI, Eurotransplant donor risk index; ECD, extended criteria donors; GGT, gamma-glutamyltransferase; ICU, intensive care unit; L-GrAFT$_7$, liver graft assessment following transplantation; MEAF, model for early allograft function; MELD, model for end-stage liver disease; score by Rhu, Rhu et al.

S.M. and L.W. contributed equally to this work.

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

and 0.78, respectively. In subgroup analyses, the EASE score was suited best for male recipients with a high-MELD ($> 25$) and an EAD organ. Scores only based on pretransplant data performed worse compared to scores including postoperative data (eg, ET-DRI vs. EAD, $p < 0.001$ at 3-month graft survival). Out of these, the BAR score performed best with a $c$-statistic of 0.6. This a comprehensive comparison of the clinical utility of risk scores after liver transplantation. The EASE score sufficiently predicted 12-month graft and patient survival. Despite a relatively complex calculation, the EASE score provides significant prognostic value for patients and health care professionals in the Eurotransplant region.

## INTRODUCTION

Liver transplantation is the only curative treatment option for patients with acute liver failure, chronic liver disease due to other reasons, and distinct stages of hepatocellular carcinoma.[1,2] Nevertheless, the demand for transplantation by far exceeds the number of available liver grafts. Strategies to overcome this disparity include increased usage of split livers, living donor grafts as well as usage of grafts from extended criteria donors (ECD). ECDs are donors of older age, donors with elevated liver enzymes, risk factors such as smoking, obesity, and diabetes as well as steatosis hepatis.[3] All these criteria have been shown to be associated with an increased risk for ischemia reperfusion injury and consequently higher rates of primary nonfunction and early allograft dysfunction (EAD).[4,5]

Recently, efforts have been made to adequately predict the risk for inferior graft performance based either on donor risk factors, recipient risk factors, or both.[6,7] Since the establishment of the concept of EAD by Olthoff et al,[8] there has been a steady trend aimed at improving sensitivity and specificity of predictive scores for liver transplantation.[9] Most scores are powered for short-term graft survival, that is, 3–12 months, but are often also useful in predicting patient survival.

Currently available scores for risk prediction in liver transplantation can be grouped according to 4 categories:

i. Scores only factoring in donor risk factors such as the donor risk index (DRI)[6] and Eurotransplant donor risk index (ET-DRI).[10]
ii. Scores with a combination of donor and recipient risk factors such as the product of donor age and model for end-stage liver disease (D-MELD)[7] and balance of risk score (BAR).[11]
iii. Conventional recipient outcome scores such as EAD by Olthoff et al(EAD),[8] model for early allograft function (MEAF),[9] a recently published easy to use score by Rhu et al[12] and finally
iv. Scores factoring in postoperative laboratory dynamic, such as the liver graft assessment following transplantation (L-GrAFT$_7$)[13] score and the early allograft failure simplified estimation (EASE)[14] (Table 1).

The most impressive $c$-statistic has so far been achieved by the EASE score with 0.85.[15] Both the EASE and the liver graft assessment following transplantation (L-GrAFT$_7$) differ from previously available scores by adjusting for the dynamics of aspartate- and alanine-aminotransferase (AST/ALT) values as well as platelet numbers, making them more difficult to calculate, even if smartphone-based solutions have become available.[16] Nevertheless, the scores propose to allow valuable predictive insight based on laboratory values in the first 10 postoperative days.

Previous studies have analyzed a selection of available prediction scores,[13,17] but a complete overview concerning available donor and recipient risk scores is still lacking. Furthermore, despite these scores being well internally validated, it has been shown for previous scores that validation in different countries and transplantation networks is necessary.[10] An external validation in a Eurotransplant data set for L-GrAFT$_7$ and EASE is missing so far. This is of special interest since donation after circulatory death (DCD) is only available in few Eurotransplant countries and not an option in Germany. In addition, the proportion of donation after brain death (DBD) donors fulfilling extended criteria for transplantation is quite high, with 69% of graft fulfilling more than 1 ECD criterion.[18,19] Overall and at least in

Germany, there has been a rapid decline of overall available donors of up to 30% in the last decade.[20] We therefore wanted to validate all currently available liver transplantation outcome scores in a region with high ECD proportion, without DCD, and with an overwhelming organ scarcity.

## METHODS

### Study design

This is a retrospective, single-center analysis of all patients undergoing liver transplantation at the Charité—Universitätsmedizin Berlin from January 2007 to December 2021. Exclusion criteria were pediatric patients and retransplantation. Patients were routinely transferred to a specialized intensive care unit (ICU) immediately after transplantation. After discharge from the normal ward, patients presented for routine follow-up exams at our in-house outpatient clinic. Patient data were extracted from the electronic health records and anonymized; data on organ donors were retrieved from Eurotransplant. Data collection was performed from December 2021 until February 2022. The study was approved by the institutional ethics board (Ethik-kommission der Charité—Universitätsmedizin Berlin, EA1/369/16). The need for informed consent was waived by the ethics committee of the Charité—Universitätsmedizin Berlin due to the retrospective nature of this study. All research in this study was performed in accordance with both the Declarations of Helsinki and Istanbul, and the relevant guidelines and regulations of the ethics committee.

### Outcome scores

The objective of this study was to validate available outcome scores for liver transplantation based on donor data, recipient data, or data of both. Included in the analysis were the DRI,[6] ET-DRI,[10] D-MELD,[7] BAR,[11] EAD,[8] MEAF,[9] L-GrAFT$_7$ score,[13] EASE score,[14] and Rhu using AST, total bilirubin and coagulation factor with a superior *c*-statistic compared to EAD and MEAF.[12] Scores were calculated as described in their respective original publications and included donor characteristics such as age, gamma-glutamyltransferase (GGT), height, cause of death, type of graft (whole organ or "split or partial"), cold ischemia time and rescue allocation (Table 1). All subjects were set "Caucasian" for all calculations as the origin is not generally recorded in Germany. Recipient parameters included age, model for end-stage liver disease (MELD), retransplantation, life support prior to transplantation, and amount of intraoperative packed red blood cells. Postoperative

laboratory values included blood levels of bilirubin, AST, ALT, international normalized ratio, and platelet count until the 10th postoperative day. In addition, if an arterial or portal vein thrombosis occurred, this was documented. No values were missing for the calculation of DRI, D-MELD, BAR, and EAD. However, the data gap for score calculation was 31.6% for the ET-DRI, 9.6% for the MEAF, 43.2% for the L-GrAFT$_7$ score, 60.2% for the EASE score, and 4.2 for the Rhu score. Data was retrieved from electronic patient health records and supplemented by data from Eurotransplant, if necessary.

### Survival analyses

For survival analysis, cutoff values were chosen from previously published literature. The cutoffs were $\geq 2$ for DRI and ET-DRI, $> 1628$ for the D-MELD, $\leq 18$ for the BAR, 1 for EAD, $\leq 8$ for MEAF, $> 0$ for EASE and L-GrAFT$_7$ and $\geq 2$ for the score by Rhu. Graft survival was calculated based on outcome score cutoffs, plotted as Kaplan-Meier curves and compared using the log-rank method. In addition cox proportional hazard models were created for all scores for overall graft and patient survival. Results are reported as HR and 95% CI.

### Statistical analysis

Statistical analysis was performed using *R* (version 4.1.2) and *R* Studio (version 2021.09.0) for macOS (R Foundation for Statistical Computing, Vienna, Austria).[21] Additional required packages for graph plotting and analysis were tidyverse, survminer, receiver operating curve (ROCit), pROC, survival, and gtsummary. Descriptive data are reported as median and interquartile range (IQR). Missing values were not imputed. Overall, a *p* value <0.05 was considered statistically significant.

Area under the receiver operating curve are reported for 3-month, 6-month, and 12-month graft and patient survival. Graft loss was defined as the need for retransplantation or patient death due to graft failure. Curves were compared using DeLong test for 2 ROC curves.[22]

## RESULTS

### Patient and donor demographics

Between January 1, 2007 and December 31, 2021, a total number of 1185 liver transplantations were performed at our center. Out of these, 60 were living donor liver transplantations (5.1%), 93 pediatric liver transplantations (7.8%), and 194 liver retransplantations

**TABLE 1**  Score data

| Characteristic | DRI | ET-DRI | D-MELD | BAR | EAD | MEAF | L-GrAFT$_7$ | EASE | Rhu |
|---|---|---|---|---|---|---|---|---|---|
| Object of score | Donor quality | Donor quality | Donor-recipient match | Donor-recipient match | Graft quality | Graft recovery | Graft recovery | Graft recovery | Graft quality |
| Endpoint | Graft failure | Graft failure | Graft failure/patient death | Patient death | Graft dysfunction | Graft failure | Graft failure | Graft failure | Graft failure |
| Cutoff | ≥2 | ≥2 | >1628 | ≤18 | 1 | ≤8 | >0 | >0 | ≥2 |
| Days of evaluation to LT | Intraoperative | Intraoperative | -1 | Intraoperative | 7 | 3 | 7 | 10 | 7 |
| Donor | | | | | | | | | |
| Age | X | X | X | X | | | | | |
| GGT | | X | | | | | | | |
| Race[a] | X | | | | | | | | |
| Height | X | | | | | | | | |
| Cause of death | X | X | | | | | | | |
| DCD[b] | X | X | | | | | | | |
| Partial or split | X | X | | | | | | | |
| Location | X | X | | X | | | | | |
| Cold ischemia time | X | X | | | | | | | |
| Rescue allocation | | X | | | | | | | |
| Recipient | | | | | | | | | |
| Age | | | | X | | | | | |
| MELD | | | X | X | | | | X | |
| Retransplantation | | | | X | | | | | |
| Life support | | | | X | | | | | |
| Packed red blood cells | | | | | | | | X | |
| High-volume center | | | | | | | | X | |
| After transplant | | | | | | | | | |
| INR >1.6 POD 7 | | | | | X | | | | |
| Bilirubin >10 mg/dL POD 7 | | | | | X | | | | |
| ALT/AST >2000 U/L POD 7 | | | | | X | | | | |
| ALT max. POD 1–3 | | | | | | X | | | |
| INR max. POD 1–3 | | | | | | X | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Score bilirubin POD 3** | | | | | X | | | | |
| AST POD 1–7 | | | | | | | X | | |
| Bilirubin POD 1–7 | | | | | | | X | | |
| Platelets POD 1–7 | | | | | | | X | | |
| INR max. POD 1–7 | | | | | | | X | | |
| AST POD 1, 2, 3, 7, and 10 | | | | | | | | X | |
| Platelets POD 1, 3, 7, and 10 | | | | | | | | X | |
| Bilirubin POD 1, 3, 7, and 10 | | | | | | | | X | |
| Vascular thrombosis POD 1–10 | | | | | | | | X | |
| ALT max. POD 1–7 | | | | | | | | | X |
| Bilirubin max. POD 3–7 | | | | | | | | | X |
| INR max. POD 3–7 | | | | | | | | | X |
| No. of variables | 8 | 8 | 2 | 6 | 3 | 3 | 3 | 7 | 3 |
| Discrimination ability at 90 d | | | | | | | | | |
| In the derivation set | NA | 0.63 | 0.70 and 0.64 | NA | 0.72 | NA | 0.83 | 0.85 | 0.74 |
| In the validation set | 0.57 | 0.58 | 0.72 and 0.64 | 0.7 | 0.63 | 0.73 | 0.78 | 0.78 | 0.73 |

[a]Race was "Caucasian" for all donors.

[b]Not applicable in Germany.

Abbreviations: ALT indicates alanine aminotransferase; AST, aspartate aminotransferase; DCD, donation after circulatory death; GGT, gamma-glumatyltransferase; INR, international normalized ratio; LT, liver transplantation; MELD, model for end-stage liver disease; POD, postoperative day.

Modified from Avolio et al.[15] Adaptations are themselves works protected by copyright. So in order to publish this adaptation, authorization must be obtained both from the owner of the copyright in the original work and from the owner of copyright in the translation or adaptation.
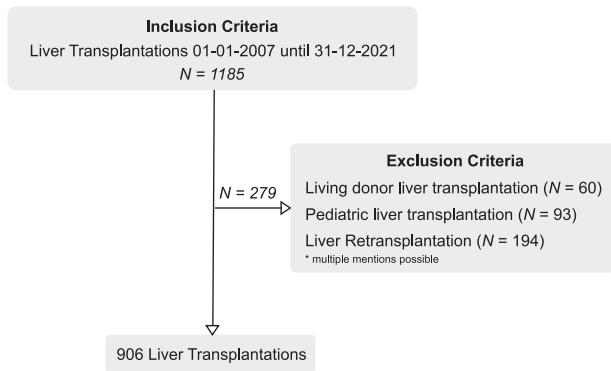
**Inclusion Criteria**
Liver Transplantations 01-01-2007 until 31-12-2021
*N = 1185*

*N = 279*

**Exclusion Criteria**
Living donor liver transplantation (*N* = 60)
Pediatric liver transplantation (*N* = 93)
Liver Retransplantation (*N* = 194)
* multiple mentions possible

906 Liver Transplantations

**FIGURE 1** Patient flow diagram with inclusion and exclusion criteria. Overall, 906 patients were analyzed in the study.

(16.4%). For downstream analysis, a subset of all adult recipient with a primary liver transplantation from a brain-dead donor (n = 906) was used (Figure 1).

Recipients had a median age of 56.6 years (IQR: 12.42 y) and were transplanted most frequently due to liver cirrhosis (42.7%) or hepatic malignancies (26%). Less recipients were female (33%). The median MELD at transplantation was 17 points (IQR: 16). The median

**TABLE 2** Recipients

| Characteristic | N = 906 |
|---|---|
| Age (y) | 56.6 (12.42) |
| Sex (female), n (%) | 295 (33) |
| BMI (kg/m$^2$) | 26.1 (6.1) |
| Blood group, n (%) | |
| A | 390 (43) |
| AB | 76 (8.4) |
| B | 139 (15) |
| O | 301 (33) |
| Indication, n (%) | |
| Acute hepatic failure | 64 (7.1) |
| Benign liver tumors | 23 (2.5) |
| Cancers | 236 (26) |
| Cholestatic disease | 88 (9.7) |
| Cirrhosis | 387 (42.7) |
| Congenital biliary disease | 7 (0.7) |
| Metabolic disease | 28 (3.1) |
| Vascular | 5 (0.6) |
| Viral | 6 (0.7) |
| Other liver diseases | 62 (6.8) |
| LabMELD | 17 (16) |
| Transfused RBCs | 6 (7) |
| Intensive care unit (d) | 9 (14) |
| Length of hospital stay (d) | 32 (31) |

Data are represented as median (IQR) unless it is mentioned.
IQR indicates interquartile range; RBC, red blood cell; MELD, model for end-stage liver disease.

**TABLE 3** Donors

| Characteristic | N = 906 |
|---|---|
| Cause of death, n (%) | |
| Anoxia | 16 (1.8) |
| Cerebrovascular accident | 832 (92) |
| Trauma | 37 (4.1) |
| Other | 21 (2.3) |
| Age | 57 (24) |
| Blood group, n (%) | |
| A | 391 (43) |
| AB | 62 (6.8) |
| B | 120 (13) |
| O | 333 (37) |
| Sex (female) | 444 (49) |
| BMI (kg/m$^2$) | 26.0 (5.0) |
| Sodium (mmol/l) | 140 (7) |
| AST (U/L) | 39 (56) |
| ALT (U/L) | 30 (42) |
| GGT (U/L) | 33 (47) |
| Intensive care unit (d) | 3.0 (6) |
| Cold ischemia time (h) | 9.23 (2.93) |
| DRI | 1.56 (0.5) |
| ET-DRI | 1.76 (0.4) |

Data represented as median (IQR) unless it is mentioned.
Abbreviations: ALT indicates alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; DRI, donor risk index; ET-DRI, Eurotransplant donor risk index; GGT, gamma-glutamyltransferase; IQR, interquartile range.

cold ischemia time was 9.23 hours (IQR: 2.93 h) and a median of 6 units packed red blood cells were used intraoperatively. The median postoperative ICU stay was 9 days, and the overall length of hospital stay was 32 days (Table 2).

Donors had a median age of 57 years (IQR: 24 y) and the most frequent cause of death was a cerebrovascular accident (92%). Laboratory parameters of liver damage were within the reference range, median ICU stay before donation was 3 days (IQR: 6 d). The median DRI and ET-DRI were 1.56 and 1.76, respectively (Table 3).

## Predictive quality of outcome scores

The highest c-statistic was achieved in all scores for the prediction of 3-month graft survival, with the L-GrAFT$_7$ and the EASE score both achieving 0.8 (Figure 2 and Supplementary Table 1, http://links.lww.com/HC9/A34). Worst overall outcome prediction was at 12-month patient survival. With only the EASE score achieving a c-statistic of 0.69 (95% CI: 0.6–0.78), the L-GrAFT$_7$ and the EASE score were not significantly different at any analyzed time point.
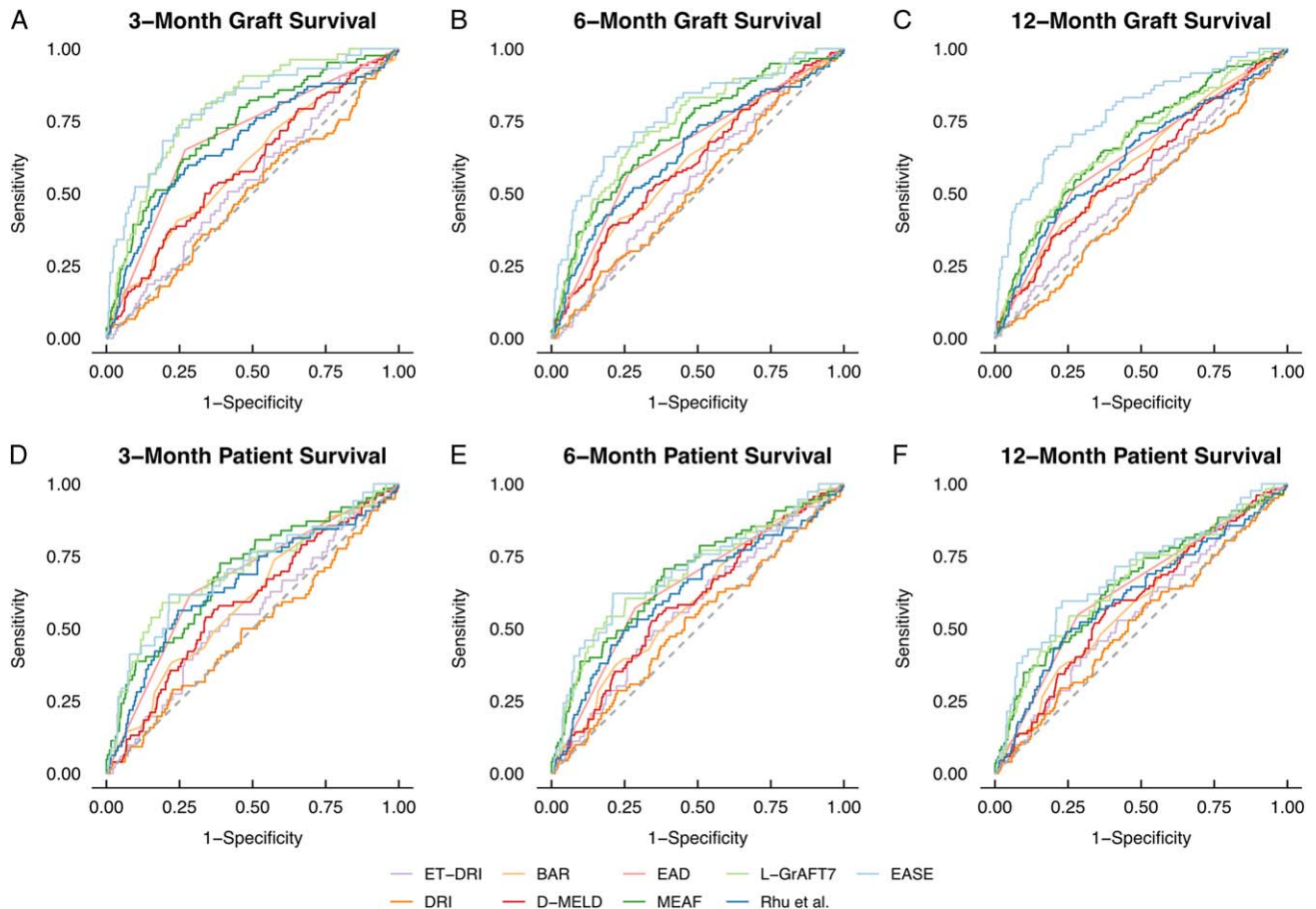
**FIGURE 2** Receiver operatic characteristic curves for relevant liver transplantation outcome scores. Abbreviations: BAR indicates balance of risk score; DRI, donor risk index; EAD, early allograft dysfunction; ET-DRI, Eurotransplant donor risk index; EASE: early allograft failure simplified estimation score; L-GrAFT$_7$: liver graft assessment following transplantation; MEAF, model for early allograft function, Rhu and colleagues. Score proposed by Rhu et al,[12] D-MELD: product of donor age and model for end-stage liver disease score.

The DRI had the least predictive power for 3-month graft survival, with a *c*-statistic of 0.5 (95% CI: 0.44–0.55). Throughout the board, scores only factoring in donor risk factors, that is, DRI and ET-DRI, had inferior predictive qualities compared to scores factoring in at least 1 recipient characteristic, that is, D-MELD and BAR (*c*-statistic DRI 0.5 vs. 0.6 for bar, *p* = 0.02 for 3-month graft survival). Out of scores with data available preoperatively, the BAR score performed best with a *c*-statistic of 0.6 for 3-month graft survival. In classic scores of outcome prediction such as EAD, MEAF and the Rhu and colleagues score, MEAF had the best *c*-statistic at 3 months graft survival with 0.72 (95% CI: 0.67–0.78).

## Subgroup analysis for sex, MELD, and EADs

Of special interest was, if previously published outcome scores were robust for prediction in subgroups with regards to MELD and proportion of ECD organs. The *c*-statistic for all nine scores did not differ between high-MELD and low-MELD recipients (> 25). This was

regardless of outcome for patient or graft survival in the first 12 months after transplantation. For ECD organs, the EASE score performed nominally better for prediction of 3-month graft survival (0.83 vs. 0.54, *p* = 0.06) than the MEAF for 12-month graft survival (0.69 vs. 0.47, *p* = 0.04) (Supplementary Tables 2 and 3, http://links.lww.com/HC9/A34). MEAF had a higher *c*-static in male recipients (0.78 vs. 0.62, *p* = 0.03) for 3-month, 6-month, and 12-month graft survival. The EASE score was equally superior for 6-month and 12-month graft survival only in male recipients (0.83 vs. 0.64, *p* = 0.046; 6-month graft survival) (Supplementary Table 4, http://links.lww.com/HC9/A34). Analyzing score input variables by recipient sex, revealed a higher labMELD in female patients (21.9, IQR: 11.2 vs. 18.2, IQR: 9.53 in male patients; *p* < 0.001).

## Risk classification for graft survival

Liver transplantation outcome scores were stratified into risk groups based on previously published cut-off values. One-year graft survival from donors with an ET-DRI or DRI ≥ 2 was not significantly worse than
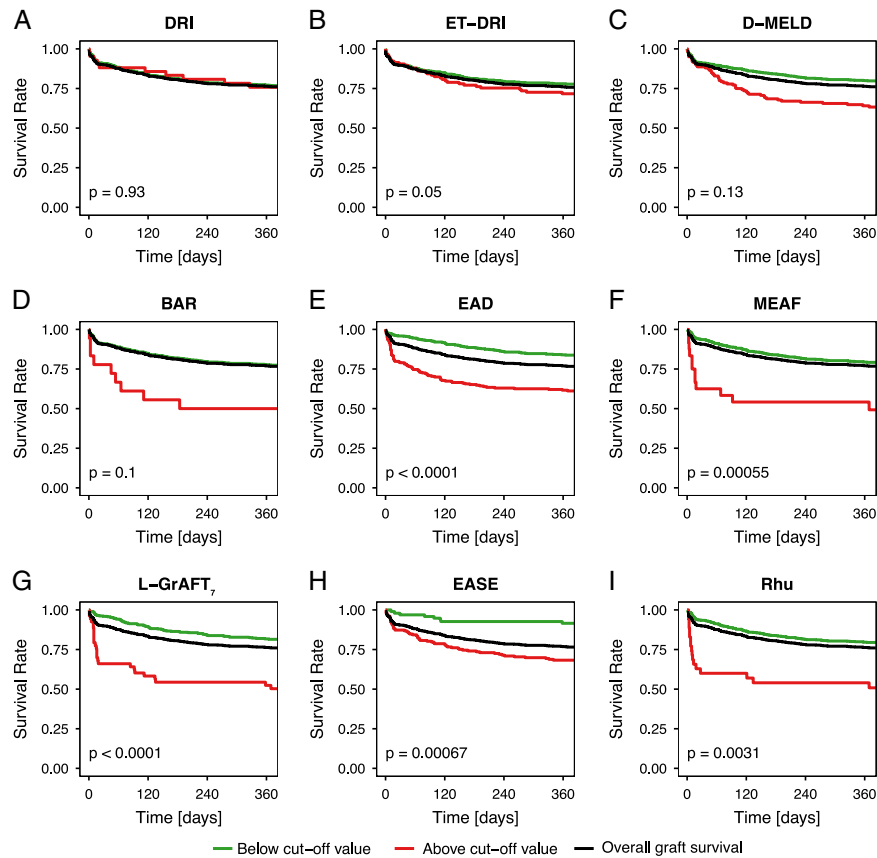
**FIGURE 3** One-year graft survival based on published cutoff values. DRI $\geq 2$, ET-DRI $\geq 2$, D-MELD $> 1628$, BAR $\leq 18$, EAD $> 0$, MEAF $\leq 8$, L-GrAFT$_7$ $> 0$, EASE $> 0$, Rhu et al $\geq 2$. Abbreviations: BAR indicates balance of risk score; DRI, donor risk index; EAD, early allograft dysfunction; ET-DRI, Eurotransplant donor risk index; EASE: early allograft failure simplified estimation score; L-GrAFT$_7$: liver graft assessment following transplantation; MEAF, model for early allograft function, Rhu and colleagues. Score proposed by Rhu et al,[12] D-MELD: product of donor age and model for end-stage liver disease score.

after transplantation of grafts from lower risk donors ($p = 0.05$ and $0.93$, respectively; Figure 3). This was equally true for D-MELD $> 1628$ and BAR $\leq 18$ points ($p = 0.13$ and $0.1$). Patients with EAD had a significantly reduced 1-year graft survival (61.6%) compared to patients without EAD (83.8%, $p < 0.001$). The 1-year graft survival was 79% for patients with a MEAF $\leq 8$ and 54.2% for patients with a MEAD $> 8$ ($p < 0.001$). For L-GrAFT$_7$ $> 0$, 1-year graft survival was 81.4% compared to 52.4% ($p < 0.001$). The best 1-year graft survival was seen in patients with an EASE score $> 0$ (92% compared to 68.3%, $p < 0.001$). The score by Rhu and colleagues was stratified at $\geq 2$. Above this score, the 1-year graft survival was 54% and 79.5% below that cut-off ($p = 0.003$). For overall graft and patient survival, cox proportional hazard models were created (Table 2). Again, DRI and ET-DRI proved insufficient in predicting outcome after liver transplantation. All other scores had a significant effect (Supplementary Table 5, http://links.lww.com/HC9/A34). Most frequent cause for graft failure was primary nonfunction (1.5%) followed by arterial thrombosis (1.3%) and infection (1.2%) (Supplementary Table 6, http://links.lww.com/

HC9/A34). Analyzing causes for graft failure revealed no direct link between a higher prevalence of vascular anastomosis complications, biliary complications, or rejection between risk classification groups. Primary nonfunction was the only cause for graft failure classified well by all scores except for the D-MELD.

## DISCUSSION

Liver transplantation outcome scores provide valuable insight in predicting graft or patient outcome, which is of special interest in times of broader use of EADs and potentially reduced graft quality.[3] Timely identification of patients at increased risk can help monitor them more closely, improve survival through intervention if necessary, and enable further research in this area. However, a comprehensive analysis on how these different risk prediction-scoring systems perform in "real-life conditions" at large liver transplant centers was lacking up to now. We here show that the L-GrAFT$_7$ and EASE score are superior compared to the remaining scores in predicting graft outcome in our cohort from Germany, with a $c$-statistic of 0.8 for 3-month graft survival.

This study is the first external validation in the Eurotransplant area of the L-GrAFT$_7$ and EASE scores, which were developed in 2018 and 2020. Their good performance proves their applicability in liver transplantation programs with relevant local differences in allocation policy and graft quality. In Germany these are namely no DCD organs, a decline in overall available grafts and a large proportion of ECD organs, especially from patients of older age or grafts with steatosis hepatis. Our findings are therefore relevant for other countries with low donor rates as well as for countries with higher donor rates, such as the United States, as a change in the donor population is also observed there.

Previous studies have shown the validity in Italian, American, British, and Chinese liver transplantation centers.[13,15,17,23] Our own c-statistic of 0.8 for the EASE score for 3-month graft survival was, however, inferior to the original publication of 0.85.[15] We assume that this is due to our retrospective study design, the percentage of missing values, and the smaller data set, in addition to longer cold ischemia times and higher recipient MELD compared to the originally used data sets.

From an organ acceptance point-of-view and to help with risk mitigation, we performed additional subgroup analyses, grouped according to recipient MELD scores, organ quality (ECD or non ECD), and analyzed the effect of sex on the applicability of the available scores. Overall worse c-statistics for graft survival in female patients must be regarded with caution; higher MELD scores for our female patients are most likely due to less MELD exceptions. As patients with hepatocellular carcinoma and lower MELD can be prioritized in Eurotransplant, a disease disproportionality affecting men due to gender-specific lifestyles, exposure to risk factors such as alcohol in addition to hepatitis B and C infection.[24,25] Nevertheless, this underlines the necessity to add continued focus on gender imbalances in liver transplantation. In the United States, women are 2 times less likely to receive a liver transplant and are of higher risk of death or becoming too sick for liver transplantation on the waitlist.[26] Future validation and score development should include gender-based approaches to solve this imbalance.

Regarding graft quality and recipient MELD, we found no significant differences in time-dependent AUC for graft or patient survival in subgroup analysis. This is not necessarily unexpected as many risk factors, that is, donor or recipient age are included in several scores and can be adjusted. However, even in the MEAF or L-GrAFT$_7$ scores without any of the named characteristics, we saw no differences. This highlights the relative robustness of the analyzed scores. Our results are comparable to results achieved in the validation data sets of each study and are only in part inferior. c-Statistics for DRI and ET-DRI were not as promising as expected, however with a c-statistic of 0.58 in the validation data set for the ET-DRI for

3-month graft survival, the predictive power was already limited.

Nevertheless, the value of scores with data available pretransplant should not be underestimated in assisting clinical decision making. Inferior predictive quality on the one hand compared to having a tool for real-time decision making for organ acceptance on the other hand is a dilemma. At least for the BAR score we found a c-statistic of 0.6 for 3-month graft survival. Combining available pretransplant donor and recipient data seems to be the best alternative for clinical decision making.

Missing and needed for future analysis is a score for outcome prediction after liver machine perfusion. In times of widespread trials and adoption of machine perfusion prior to liver transplantation, a future score should implement machine perfusion as an additional variable.[27,28] Furthermore, the "holy grail" of outcome prediction would still be a score, based on preoperatively available data on donor and recipient, potentially augmented by machine perfusion parameters and fueled by artificial intelligence, that would allow for more accurate assessment of perioperative risk.[29] With this, involved caretakers could be even more aware of the risks at play and potentially positively influence the outcome—if possible.

## Limitations

We are limited by the single-center design and the high rates of missing values for GrAFT$_7$ and EASE scores. Nevertheless, both scores prove to be of superior predictive value in our data set. The original limitations of the L-GrAFT$_7$ and EASE score remain the same; they are complicated to calculate, and retrospective sampling of patient data is cumbersome for that many parameters, despite online calculators and smartphone apps have been recently proposed by Avolio and colleagues. Nonetheless, the retrospective single-center design carries the advantage of a relatively homogeneous, standardized clinical management from recipient evaluation, donor organ selection, surgical procedures, and post-transplant care. The limitations of our analyses call for prospective validation of clinical outcome scores for liver transplantation in large multicenter trials.

## CONCLUSION

This is the most comprehensive comparison of outcome scores in a liver transplantation data set. The EASE score had the highest overall c-statistic and was the only score sufficiently predicting 12-month graft and patient survival. Out of scores with preoperatively available data, the BAR score performed best with a c-statistic of 0.6 for 3-month graft survival. Despite a relatively complex calculation, the EASE score provides

significant prognostic value in an Eurotransplant region center for patients and caretakers.

## CONFLICT OF INTEREST
A.W. received grants from GBA of the federal government of Germany. The remaining authors declare that they have no financial conflict of interest.

## ORCID
*Simon Moosburner* http://orcid.org/0000–0003–1879–4788
*Leke Wiering* http://orcid.org/0000–0003–4144–3652
*Axel Winter* http://orcid.org/0000–0001–7435–4993
*Münevver Demir* http://orcid.org/0000–0002–7050–797X
*Joseph M.G.V. Gaßner* http://orcid.org/0000–0002–6984–2738
*Brigitta Globke* http://orcid.org/0000–0002–1482–6255
*Georg Lurje* http://orcid.org/0000–0001–9674–0756
*Frank Tacke* http://orcid.org/0000–0001–6206–0226
*Wenzel Schöning* http://orcid.org/0000–0002–0446–5126
*Johann Pratschke* http://orcid.org/0000–0001–9839–1369
*Robert Öllinger* http://orcid.org/0000–0002–4499–1673
*Igor M. Sauer* http://orcid.org/0000–0001–9355–937X
*Nathanael Raschzok* http://orcid.org/0000–0001–5074–7063

## REFERENCES
1. Stepanova M, Kabbara K, Mohess D, Verma M, Roche-Green A, AlQahtani S, et al. Nonalcoholic steatohepatitis is the most common indication for liver transplantation among the elderly: data from the United States Scientific Registry of Transplant Recipients. Hepatol Commun. 2022;6:1506–15.
2. Ivanics T, Shwaartz C, Claasen M, Patel MS, Yoon P, Raschzok N, et al. Trends in indications and outcomes of liver transplantation in Canada: a multicenter retrospective study. Transpl Int. 2021;34:1444–54.
3. Orman ES, Mayorga ME, Wheeler SB, Townsley RM, Toro-Diaz HH, Hayashi PH, et al. Declining liver graft quality threatens the future of liver transplantation in the United States. Liver Transpl. 2015;21:1040–50.
4. Spitzer AL, Lao OB, Dick AA, Bakthavatsalam R, Halldorson JB, Yeh MM, et al. The biopsied donor liver: incorporating macrosteatosis into high-risk donor assessment. Liver Transpl. 2010;16:874–4.
5. Kan C, Ungelenk L, Lupp A, Dirsch O, Dahmen U Ischemia-reperfusion injury in aged livers—the energy metabolism, inflammatory response, and autophagy. 2018;102:368–77.
6. Feng S, Goodrich NP, Bragg-Gresham JL, Dykstra DM, Punch JD, DebRoy MA, et al. Characteristics associated with liver graft failure: the concept of a donor risk index. Am J Transplant. 2006;6:783–90.
7. Halldorson JB, Bakthavatsalam R, Fix O, Reyes JD, Perkins JD. D-MELD, a simple predictor of post liver transplant mortality for optimization of donor/recipient matching. Am J Transplant. 2009;9:318–26.
8. 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.
9. 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.
10. Braat AE, Blok JJ, Putter H, Adam R, Burroughs AK, Rahmel AO, et al. The Eurotransplant donor risk index in liver transplantation: ET-DRI. Am J Transplant. 2012;12:2789–96.
11. Dutkowski P, Oberkofler CE, Slankamenac K, Puhan MA, Schadde E, Mullhaupt B, et al. Are there better guidelines for allocation in liver transplantation? A novel score targeting justice and utility in the model for end-stage liver disease era. Ann Surg. 2011;254:745–53; discussion 53.
12. Rhu J, Kim JM, Kim K, Yoo H, Choi G-S, Joh J-W. Prediction model for early graft failure after liver transplantation using aspartate aminotransferase, total bilirubin and coagulation factor. Sci Rep. 2021;11:12909 .
13. Agopian VG, Harlander-Locke MP, Markovic D, Dumronggittigule W, Xia V, Kaldas FM, et al. Evaluation of early allograft function using the liver graft assessment following transplantation risk score model. JAMA Surg. 2018;153:436–44.
14. Agopian VG, Markovic D, Klintmalm GB, Saracino G, Chapman WC, Vachharajani N, et al. Multicenter validation of the liver graft assessment following transplantation (L-GrAFT) score for assessment of early allograft dysfunction. J Hepatol. 2021;74:881–92.
15. Avolio AW, Franco A, Schlegel A, Lai Q, Meli S, Burra P, et al. Development and validation of a comprehensive model to estimate early allograft failure among patients requiring early liver retransplant. JAMA Surg. 2020;155:e204095.
16. Moosburner S, Gassner J, Pratschke J, Sauer IM, Raschzok N. Smartphone Apps to stratify the risk of early allograft failure are just the beginning for next-generation outcome prediction in transplantation medicine. Hepatol Commun. 2021;6:249.
17. Chen S, Wang T, Luo T, He S, Huang C, Jia Z, et al. Prediction of graft survival post-liver transplantation by L-GrAFT risk score model, EASE score, MEAF scoring, and EAD. Front Surg. 2021;8:753056 .
18. Moosburner S, Raschzok N, Schleicher C, Bosebeck D, Gassner J, Ritschl PV, et al. Declined liver grafts—analysis of the german donor population from 2010 to 2018. Z Gastroenterol. 2020;58:945–54.
19. Moosburner S, Sauer IM, Gassner JMGV, Schleicher C, Bösebeck D, Rahmel A, et al. Macrosteatosis is a huge problem in liver transplantation—however, not the only one we face. Am J Transplant. 2019;19:2661–2.
20. Schulte K, Borzikowsky C, Rahmel A, Kolibay F, Polze N, Fränkel P, et al. Decline in organ donation in Germany. Dtsch Aerztebl Int. 2018;19:2661–2.
21. Team RC. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44:837–45.
23. Avolio AW, Lai Q, Cillo U, Romagnoli R, De Simone P. L-GrAFT and EASE scores in liver transplantation. Need for a reciprocal external

validation and comparison with other scores. J Hepatol. 2020;75: 729–31.

24. Kotsiliti E. Understanding HCC sex disparities. Nat Rev Gastroenterol Hepatol. 2022;19:147.

25. Ji F, Zhang J, Liu N, Gu Y, Zhang Y, Huang P, et al. Blocking hepatocarcinogenesis by a cytochrome P450 family member with female-preferential expression. Gut. 2022;71:2313–24.

26. Sarkar M, Watt KD, Terrault N, Berenguer M. Outcomes in liver transplantation: does sex matter? J Hepatol. 2015;62: 946–55.

27. Bruinsma BG, Sridharan GV, Weeder PD, Avruch JH, Saeidi N, Ozer S, et al. Metabolic profiling during ex vivo machine perfusion of the human liver. Sci Rep. 2016;6:22415.

28. Nasralla D, Coussios CC, Mergental H, Akhtar MZ, Butler AJ, Ceresa CDL, et al. A randomized trial of normothermic preservation in liver transplantation. Nature. 2018;557: 50–6.

29. Pettit RW, Corr S, Havelka J, Rana AA. Using artificial intelligence to improve post-transplant survival predictions. J Am Coll Surg. 2020;231:S302–3.