OPEN

Textbook Outcome as a Quality Metric in Liver Transplantation

Austin D. Schenk, MD, PhD, FACS,[1] Jing L. Han, MD,[1] April J. Logan, MS,[1] Jeffrey M. Sneddon, BS,[1] Guy N. Brock, PhD,[1] Timothy M. Pawlik, MD, MPH, PhD,[1] and William K. Washburn, MD[1]

Background. Quality in liver transplantation (LT) is currently measured using 1-y patient and graft survival. Because patient and graft survival rates now exceed 90%, more informative metrics are needed. Textbook outcomes (TOs) describe ideal patient outcomes after surgery. This study critically evaluates TO as a quality metric in LT. **Methods.** United Network for Organ Sharing data for 25 887 adult LT recipients were used to define TO as patient and graft survival >1 y, length of stay ≤10 d, 0 readmissions within 6 mo, absence of rejection, and bilirubin <3 mg/dL between months 2 and 12 post-LT. Univariate analysis identified donor and recipient characteristics associated with TO. Covariates were analyzed using purposeful selection to construct a multivariable model, and impactful variables were incorporated as linear predictors into a nomogram. Five-year conditional survival was tested, and center TO rates were corrected for case complexity to allow for center-level comparisons. **Results.** The national average TO rate is 37.4% (95% confidence interval, 36.8%-38.0%). The hazard ratio for death at 5 y for patients who do not experience TO is 1.22 (95% confidence interval, 1.11-1.34; $P \leq 0.0001$). Our nomogram predicts TO with a C-statistic of 0.68. Center-level comparisons identify 31% of centers as high performing and 21% of centers as below average. High rates of TO correlate only weakly with center volume. **Conclusions.** The composite quality metric of TO after LT incorporates holistic outcome measures and is an important measure of quality in addition to 1-y patient and graft survival.

(*Transplantation Direct* 2022;8: e1322; doi: 10.1097/TXD.0000000000001322).

INTRODUCTION

Textbook outcome (TO) was first described in 2012 as the "perfect hospitalization" after colon cancer resection[1] and has alternatively been reported to characterize an "ideal postoperative course."[2] TO is a composite outcome metric that incorporates quality across multiple domains of performance. TO has been studied in colon,[1] gastric,[2-4] esophageal,[2,3] pancreatic,[5,6] liver,[6,7] bariatric,[8] and vascular surgery[9] and even been used to measure quality associated with endoscopic retrograde cholangiopancreatography and colonoscopy.[10] TO is now being explored in transplantation.

TO is a unique quality metric. Historically, outcome studies have focused on individual binary variables such as survival or readmission. TO attempts to capture multiple critical variables that contribute to an ideal healthcare experience. The composite nature of TO increases the "event rate" surrounding this metric[7] and creates a very "high bar," in which no credit is given for near-perfect care.[11] TO is thought to be patient-centric in part because patients use a similar "high bar" to judge their healthcare experience, in which a single negative event in any domain can taint perception of the entire hospitalization. Because TO is a holistic measure that incorporates patient outcomes and healthcare efficiency, it is

Received 9 November 2021. Revision received 22 February 2022.

Accepted 15 March 2022.

[1] Department of Surgery, The Ohio State University Wexner Medical Center, Columbus, OH.

The funding was received from The Ohio State University Center for Clinical and Translational Science (grant UL1TR00273).

The authors declare no conflicts of interest.

The data reported here have been supplied by the United Network for Organ Sharing as the contractor for the Organ Procurement and Transplantation Network. The interpretation and reporting of these data are the responsibility of the author(s) and in no way should be seen as an official policy of or interpretation by the Organ Procurement and Transplantation Network or the US Government.

A.D.S. and A.J.L. participated in research design, database queries, statistical analyses, article preparation, and article editing.

J.L.H., T.M.P., and W.K.W. participated in research design and article editing. J.M.S. participated in research design, database queries, and article editing. G.N.B. participated in research design, statistical analyses, and article editing.

Supplemental digital content (SDC) is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site (www.transplantationdirect.com).

Correspondence: Austin D. Schenk, MD, PhD, FACS, Department of Surgery, The Ohio State University Wexner Medical Center, Comprehensive Transplant Center, 395 W 12th Ave Suite 168, Columbus, OH 43210. (austin.schenk@osumc.edu).

ISSN: 2373-8731

DOI: 10.1097/TXD.0000000000001322

of interest to multiple stakeholders, including patients, providers, payers, and regulatory agencies.

Transplantation is, by necessity, a highly regulated field. Quality metrics in transplant have enormous implications for individual patients, transplant programs, and hospitals. The value of current metrics is fiercely debated,[12,13] and a joint taskforce appointed by the American Society of Transplantation and American Society of Transplant Surgeons has stated that "current publicly reported short-term outcome metrics are no longer reflective of the quality of care delivered and have unintended consequences, including limiting organ utilization and excluding high-risk patients who would benefit from transplantation."[14] Recently, the Organ Procurement and Transplantation Network hosted national webinars, providing the first glimpse of proposed changes in quality monitoring. Concurrent with this national committee work, we have developed the quality metric of TO in liver transplantation (LT). We define TO after LT using a large contemporary cohort and suggest that TO is a novel method for measuring and improving quality in LT.

MATERIALS AND METHODS

Healthcare professionals from the liver transplant team at our institution met on 4 separate occasions to establish a consensus definition of TO after LT. Our team consists of 5 liver transplant surgeons, 6 transplant hepatologists, 6 advanced transplant providers, and 3 transplant fellows, and an additional 5 health service researchers provided input. In defining the ideal patient experience after LT, consideration was given to relevant clinical outcomes, perceived patient satisfaction, complications, and efficient resource utilization. Questions used to frame the discussion are listed in Table S1 (SDC, http://links.lww.com/TXD/A419). This ideal list was then narrowed to data available in the United Network for Organ Sharing Standard Transplant Analysis and Research (UNOS STAR) files. Once TO criteria were defined, it was necessary to define threshold values for TO variables. Within our data set, the median length of stay was 10 d, and we selected ≤10 d as our threshold. Similar methodology was applied in choosing threshold values for the other criteria. Our final definition of TO after LT is provided in Table 1 (with complete details of the derivation provided in Table S2 (SDC http://links.lww.com/TXD/A419).

STAR file data for LT recipients transplanted between 2013 and 2017 were queried (based on Organ Procurement and Transplantation Network data as of June 12, 2020). Exclusion criteria included age <18 y at the time of transplant, living donor transplant, retransplant within 1 y of the initial transplant, and simultaneous liver-kidney transplant. Covariates were chosen (Table 2; complete details provided inTable S3, SDC, http://links.lww.com/TXD/A419), and univariate analysis of donor and recipient characteristics was performed using the Welch *t* testing for continuous variables and chi-square testing for categorical variables. Table S4 (SDC, http://links.lww.com/TXD/A419) lists the percentage of data missing for each covariate. In total, 7.3% of recipients had a missing value in at least 1 of the listed covariates and were excluded from the study. Sensitivity analyses indicated that multiple imputation would not alter results.

To identify covariates with the greatest impact on the model, we performed logistic regression using purposeful selection[15] as the multivariable selection technique. Before fitting the regression model, model for end-stage liver disease (MELD) scores were capped at 40, and cold ischemia times were capped at 16 h. Recipient creatinine, international normalized ratio, and bilirubin at transplant were omitted from model fitting, as these are redundant with MELD. All covariates with a univariate *P* value of <0.0001 were initially included in model fitting but systematically dropped if the maximum *P* value of the variable failed to meet the retention criteria ($P < 0.0001$) and failed to change at least 1 parameter estimate by ≥50%. Dropped variables were retested for inclusion. Categorical variables were separated into component parts, and if any part met the criteria for inclusion in the purposeful selection, the entire categorical variable was included in the model. The final cutoffs chosen ($P < 0.0001$ and 50% change in parameter estimate) were selected by iteratively rerunning purposeful selection with intent to sensibly reduce the model. Least absolute shrinkage and selection operator[16] technique was then used to assess for a further possible reduction of the model using cross-validation to find the optimal shrinkage value λ, repeatedly fitting the model with nine tenths of the data and withholding one tenth of the data for the evaluation of λ for 10 iterations (Figure S1A, SDC, http://links.lww.com/TXD/A419).

To rank covariates identified in purposeful selection with respect to their impact on the model, odds ratios (ORs) with 95% confidence intervals (CIs) were calculated from the regression. Standardized estimates, calculated as the product of regression coefficient and SD, were then used to compare categorical and continuous variables, with binary indicators for categorical variables. Finally, absolute values of standardized estimates were used to rank order variables with respect to their impact on the model. Impactful variables were incorporated as linear predictors into a nomogram for predicting TO after LT. The nomogram was validated using bootstrapping technique with resampling performed 1000 times. Unselected samples were used to generate an averaged receiver operating curve and mean C-statistic reflecting predictive accuracy of the nomogram (Figure S1B, SDC, http://links.lww.com/TXD/A419).

Conditional survival analysis was performed using standard Kaplan-Meier and Cox proportional hazards techniques with TO set as the conditional variable. Seventy-two percent of the patient cohort had 5-y survival data available and were included in the analysis.

Center-level quality was calculated as the ratio of observed (O) to expected (E) TO rate, and centers performing <20 liver

TABLE 1.

Textbook outcome after liver transplantation

Criteria	No. and % of patients meeting criteria	Median (IQR)
1. Patient and graft survival ≥1 y	23 395 (90.4)	–
2. Absence of rejection during first year posttransplant	19 983 (77.2)	–
3. Length of stay ≤10 d	14 206 (54.9)	10 (7–16)
4. No readmissions during first 6 mo posttransplant	19 371 (74.7)	–
5. Total bilirubin not >3 during month 2–12 of follow-up	23 677 (91.5)	0.6 (0.4–0.9)

IQR, interquartile range.

TABLE 2.

Donor and recipient characteristics in transplants with and without textbook outcome

	Textbook outcome, n (%) or mean [SD]		
	Yes N = 9691 (37.4%)	No N = 16 196 (62.6%)	*P*
Recipient characteristics			
Age (y)	57 [10]	55 [11]	<0.0001
Gender			<0.0001
Male	6920 (71.4)	10 332 (63.8)	
Female	2771 (28.6)	5864 (36.2)	
BMI (kg/m^2)	28.7 [5.4]	29.0 [6.1]	0.0011
Ethnicity			<0.0001
White	7110 (73.4)	11 382 (70.3)	
Black	784 (8.1)	1529 (9.4)	
Hispanic	1190 (12.3)	2367 (14.6)	
Other	607 (6.3)	918 (5.7)	
ABO			0.1591
A	3546 (36.6)	5982 (36.6)	
AB	520 (5.4)	769 (4.8)	
B	1310 (13.5)	2186 (13.5)	
O	4315 (44.5)	7313 (45.2)	
Primary diagnosis			<0.0001
NASH	1069 (11.0)	1953 (12.1)	
Alcoholic liver disease	1572 (16.2)	3371 (20.8)	
Hepatitis B or C	1767 (18.2)	2759 (17.0)	
Cholestatic liver disease	660 (6.8)	1324 (8.2)	
Autoimmune liver disease	171 (1.8)	484 (3.0)	
Metabolic liver disease	229 (2.4)	450 (2.8)	
Malignant liver disease	3437 (35.5)	3934 (24.3)	
Graft failure	37 (0.4)	161 (1.0)	
Acute liver failure	216 (2.2)	674 (4.2)	
Idiopathic and other	533 (5.5)	1086 (6.7)	
Diabetic			0.8140
Yes	2653 (27.4)	4412 (27.2)	
No	7038 (72.6)	11 784 (72.8)	
Hx of malignancy			<0.0001
Yes	2784 (28.7)	3398 (21.0)	
No	6907 (71.3)	12 798 (79.0)	
Hx upper abdominal surgery			0.3045
Yes	5121 (52.8)	8665 (53.5)	
No	4570 (47.2)	7531 (46.5)	
Hx of prior liver transplant			<0.0001
Yes	138 (1.4)	413 (2.6)	
No	9553 (98.6)	15 783 (97.5)	
Portal vein thrombosis			0.0004
Yes	1272 (13.1)	2381 (14.7)	
No	8419 (86.9)	13 815 (85.3)	
TIPS			0.0002
Yes	872 (9.0)	1691 (10.4)	
No	8819 (91.0)	14 505 (89.6)	
Dialysis before transplant			<0.0001
Yes	325 (3.4)	2441 (15.1)	
No	9366 (96.7)	13 755 (84.9)	
Waitlist time (d)	274 [413]	242 [420]	<0.0001
Karnofsky fxn status at Tx			<0.0001
0%–40%	2783 (28.7)	8312 (51.3)	
50%–70%	4684 (48.3)	5920 (36.6)	
80%–100%	2224 (23.0)	1964 (12.1)	
Pre-Tx medical condition			<0.0001
In ICU	536 (5.5)	3193 (19.7)	
Hospitalized, not ICU	1205 (12.4)	3777 (23.3)	
Not hospitalized	7950 (82.0)	9226 (57.0)	
Intubated before transplant			<0.0001
Yes	102 (1.1)	1151 (7.1)	
No	9589 (99.0)	15 045 (92.9)	
Ascites			<0.0001
Absent	3225 (33.3)	3833 (23.7)	
Slight	4186 (43.2)	6787 (41.9)	
Moderate-severe	2280 (23.5)	5576 (34.4)	
Encephalopathy (grade)			<0.0001
None	4366 (45.1)	5371 (33.2)	
1–2	4712 (48.6)	8389 (51.8)	
3–4	613 (6.3)	2436 (15.0)	
MELD at transplant	19 [10]	25 [12]	<0.0001
Total bilirubin at Tx (mg/dL)	6.1 [8.7]	11.0 [12.5]	<0.0001
Serum Cr at Tx (mg/dL)	1.2 [0.8]	1.5 [1.1]	<0.0001
INR at Tx	1.8 [1.0]	2.1 [1.4]	<0.0001
Exception cases			<0.0001
HCC	3330 (34.4)	3018 (18.6)	
Non-HCC	6361 (65.6)	13 178 (81.4)	
Serum albumin	3.2 [0.7]	3.1 [0.7]	<0.0001
Donor characteristics			
Age (y)	42 [16]	42 [16]	0.0773
BMI (kg/m^2)	28.1 [6.7]	28.0 [6.6]	0.1728
Cold ischemia time (h)	6.0 [2.4]	6.3 [2.4]	<0.0001
DCD			0.0050
Yes	696 (7.2)	1018 (6.3)	
No	8995 (92.8)	15 178 (93.7)	
Allocation type			<0.0001
Local	6900 (71.2)	10 216 (63.1)	
Regional	2410 (24.9)	5256 (32.5)	
National	381 (3.9)	724 (4.5)	
Increased risk donor			0.0721
Yes	2340 (24.2)	3752 (23.2)	
No	7351 (75.9)	12 444 (76.8)	
Donor w/ clinical infection			<0.0001
Yes	6968 (71.9)	12 005 (74.1)	
No	2723 (28.1)	4191 (25.9)	

BMI, body mass index; Cr, creatinine; DCD, donation after circulatory death; fxn, function; HCC, hepatocellular carcinoma; Hx, history; ICU, intensive care unit; INR, international normalized ratio; MELD, model for end-stage liver disease; NASH, nonalcoholic steatohepatitis; SD, standard deviation; TIPS, transjugular intrahepatic portosystemic shunt; Tx, transplant; /w, with.

transplants per year were excluded from analysis. Expected log odds of TO for each center were calculated on the basis of averaging the patient-level expected values from the purposeful selection logistic regression model. Bootstrapping was used to calculate 95% CIs for the O:E ratio, using 1000 bootstrap samples stratified by center. Margins of error for each center were based on half differences between the 2.5th and 97.5th bootstrap percentiles. SAS (9.4 TS1M3, Cary, NC) and R software (R 3.6.1, Vienna, Austria) were used for analyses.

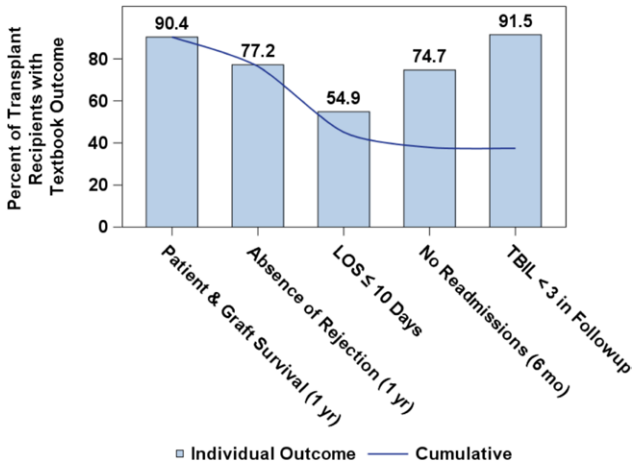


FIGURE 1. National TO rate by individual quality domain is shown. TO after liver transplantation in a national cohort of 25 877 recipients is shown. Bars represent percentage of recipients achieving the threshold for TO in each quality domain. Line represents cumulative percentage TO achieved across all quality domains. LOS, length of stay; TO, textbook outcome.

Patient-level identifiers were not used in the analyses, and this study was exempt from institutional review board approval.

RESULTS

A total of 25 887 adults who received liver transplants between 2013 and 2017 and had complete UNOS STAR file data were included in the study. TO after LT (Table 1) was defined as (1) patient and graft survival ≥1 y, (2) absence of rejection episodes during the first year posttransplant, (3) length of stay for the transplant admission ≤10 d, (4) no readmissions during the first 6 mo posttransplant, and (5) total bilirubin not >3 mg/dL during months 2 to 12 of follow-up. TO was achieved in 37.4% (95% CI, 36.8%-38.0%) of LT recipients. Excessive length of stay and early readmission accounted for the vast majority of non-TOs (Figure 1). Patient death, graft loss, and significant hyperbilirubinemia accounted for the minority of non-TOs, whereas treatment for rejection was identified in just >20% of LT recipients.

Univariate analyses of both categorical and continuous variables are presented in Table 2. LT recipients with TO were slightly older (57 versus 55), more commonly male (71% versus 29%), and more frequently associated with White or "other" ethnicity. Recipients with malignant liver disease or hepatitis (B or C) had a higher likelihood of achieving TO. History of prior upper abdominal surgery did not change probability of TO; however, history of transjugular intrahepatic portosystemic shunt, portal vein thrombosis, or prior liver transplant all negatively impacted likelihood of TO. Factors associated with critical illness, including intensive care unit (ICU) status, high MELD, pretransplant dialysis, intubated state, marked ascites, severe encephalopathy, or reduced Karnofsky functional status at the time of transplant, reduce probability of TO. Donor characteristics associated with TO include short cold ischemia time, local allocation, absence of donor infection, and brain-dead donor type.

The top 9 covariates associated with TO in multivariable modeling are shown in Table 3. ICU status (OR, 1.71; 95% CI, 1.50-1.96; $P<0.0001$), high MELD (OR, 0.98; 95% CI, 0.97-0.98; $P<0.0001$), pretransplant dialysis (OR, 1.73; 95% CI, 1.50-1.98; $P<0.0001$), and reduced Karnofsky functional score (OR, 1.58; 95% CI, 1.44-1.73; $P<0.0001$) were strongly correlated with decreased odds of TO. Gender was the only demographic recipient characteristic not associated with disease severity that factored prominently in the model. Long cold ischemia time and national allocation were the donor variables most strongly associated with reduced odds of TO. These 9 covariates were then integrated into the nomogram shown in Figure 2. Internal validation with bootstrapping yielded a C-statistic of 0.68 (Figure S1, SDC, http://links.lww.com/TXD/A419, displays least absolute shrinkage and selection operator reduction analysis (A) and averaged receiver operating curve curve (B)). As shown in Figure 2C, critically ill patients with low nomogram point scores and high MELD are least likely to experience a TO.

Conditional survival analysis with TO as the conditional variable was used to validate TO as a relevant clinical quality metric. LT recipients who achieved TO had significantly increased survival 5 y posttransplant (Figure 3); the hazard ratio (HR) for death at 5 y for patients who do not experience TO is 1.22 (95% CI, 1.11-1.34; $P\leq0.0001$). Component parts of TO were also analyzed individually. The HR for death was 1.18 (95% CI, 1.04-1.34) for patients experiencing rejection, 1.25 (95% CI, 1.14-1.37) for patients with a length of stay exceeding 10 d, 1.11 (95% CI, 0.99-1.24) for patients with readmission during the first 6 mo posttransplant, and 2.22 (95% CI, 1.81-2.72) for patients with bilirubin ≥3 mg/dL during months 2 to 12 posttransplant.

The logistic regression model informing the nomogram provides an E TO rate at each level of patient complexity, in which low nomogram point scores reflect the complexity of sicker higher MELD patients. To enable center-level comparisons

TABLE 3.

Multivariable model of TO after liver transplantation

Variable	Comparison	OR for TO (95% CI)	Absolute value of standardized estimate
Pre-Tx medical condition	Not hospitalized vs hospitalized in ICU	1.711 (1.496-1.957)	0.1399
MELD	1 pt increase in MELD score	0.978 (0.974-0.981)	0.1329
Dialysis	Not on dialysis vs on dialysis	1.726 (1.504-1.982)	0.0930
Karnofsky fxn status at Tx	80%–100% vs 0%–40%	1.578 (1.442-1.726)	0.0926
Intubated	Not intubated vs intubated	2.087 (1.660-2.623)	0.0871
Encephalopathy	Absent vs grade 3–4	1.301 (1.165-1.452)	0.0702
Cold ischemia time	1 additional hour	0.956 (0.944-0.968)	0.0551
Gender	Female vs male	0.817 (0.771-0.865)	0.0525
Allocation type	National vs local	0.738 (0.646-0.843)	0.0338

CI, confidence interval; fxn, function; ICU, intensive care unit; MELD, model for end-stage liver disease; OR, odds ratio; pt, patient; TO, textbook outcome; Tx, transplant.

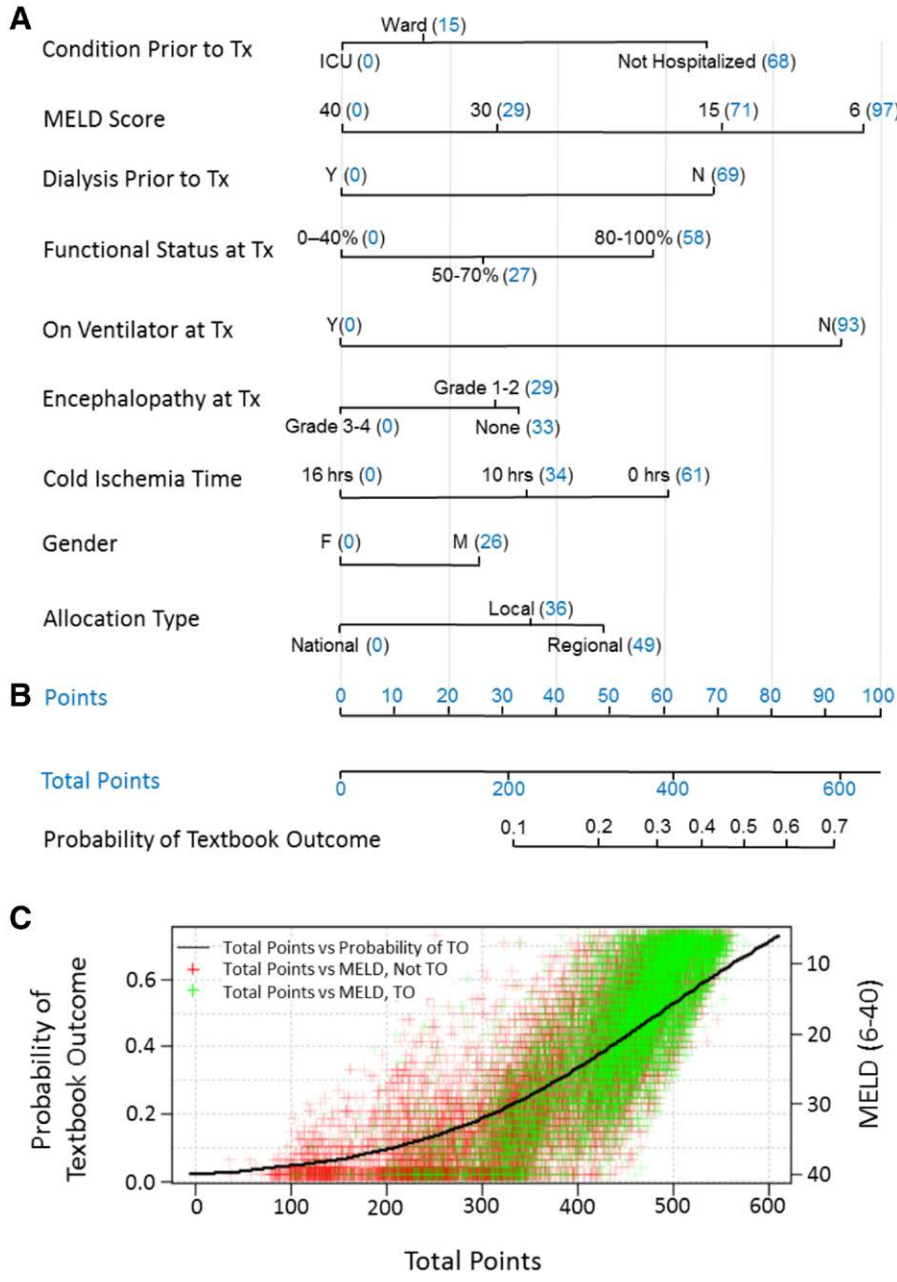
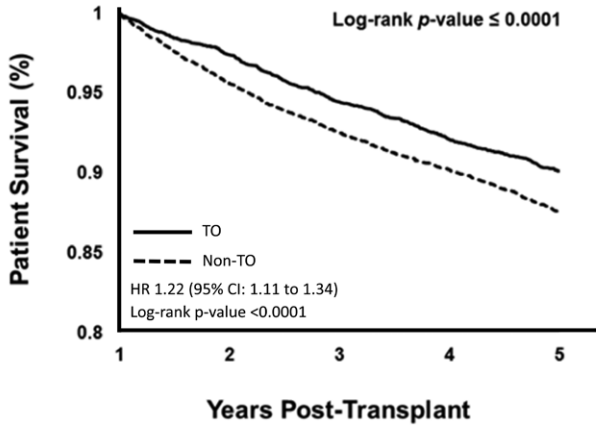


FIGURE 2. Nomogram for predicting probability of TO. A, Point assignments for nomogram predicting probability of TO. B, Scale for conversion of point score to probability (%). C, Probability of TO plotted by total nomogram point score (left y-axis) with individual clinical outcomes (TO [green] or non-TO [red]) by MELD overlaid (right y-axis). ICU, intensive care unit; MELD, model for end-stage liver disease; TO, textbook outcome; Tx, transplant.

and account for the average-case complexity at each center, expected odds of TO were calculated for the patient cohort associated with each center and compared with the O TO rate for that center (Figure 4A). Thirty-one percent of centers are identified as overperforming (O:E TO rate >1 with ≥95% confidence), and 21% of centers are identified as underperforming. Only weak correlation is observed between transplant center volume and center TO rate (Figure 4B), indicating that both low- and high-volume transplant centers are capable of offering high-quality transplant care.

DISCUSSION

When applying TO to a new surgical procedure, it is first necessary to define the ideal course after that operation. We convened a multidisciplinary group of medical and surgical LT providers at our institution to brainstorm relevant outcomes. The initial list generated bore a striking resemblance to benchmarks previously reported by Muller et al[17] and previously studied by others,[18-20] including mortality, graft loss, length of ICU and hospital stay, complications including biliary tract complications and hepatic artery thrombosis, volume of intraoperative blood transfusion, and need for renal replacement therapy. This ideal list was reshaped by necessity, as highly granular data (eg, volume of intraoperative blood transfusion) are not present in the national UNOS data set, highlighting a well-recognized need for better data collection in transplantation.[21-23] Readmission turned out to be the most reliable catch-all for surgical complications, and likewise, bilirubin >3mg/dL between months 2 and 12 posttransplant was



No. at Risk					
Non TO	10220	9827	9536	9292	9025
TO	6916	6736	6548	6386	6246

FIGURE 3. Five-year conditional survival analysis for liver transplant recipients who have and have not achieved a TO. CI, confidence interval; HR, hazard ratio; TO, textbook outcome.

the most reliable method to detect significant biliary complications. In keeping with the patient-centric intention of TO, we felt it necessary to include rejection, as this can cause psychological distress to recipients.[24]

When we applied our definition of TO after LT to a large contemporary national cohort of LT recipients, 37.4% of patients experienced a TO. In multivariable analysis, top determinants of TO (ICU status, MELD, pretransplant dialysis) were all associated with disease severity, suggesting, as expected, that TO is most difficult to achieve in critically ill recipients. It is critically important to note that many non-TOs are still of enormous benefit to the LT recipient. In contrast to surgical oncology, where an incomplete resection with positive margins may negate the benefit of the operation, failure to achieve TO in LT because of a single readmission or rejection episode does not negate the value of lifesaving surgery.

Recently, Moris et al[25] studied TO in 231 liver transplant recipients at their institution. Like ours, their definition of TO included patient and graft survival, length of stay, readmission, and rejection and also included numerous elements not measurable in the UNOS data set, including ICU readmission, intraoperative complications, early allograft dysfunction, and transfusion requirements. They report an overall TO rate of 31%, and in their small single-institution cohort, they were unable to demonstrate differences in overall or rejection-free survival. In our 5-y national cohort, we are able to demonstrate significantly increased 5-y survival for LT recipients who achieve TO (HR for death = 1.22 in non-TO recipients), and these data are consistent with reports of diminished overall survival after esophagectomy (HR, 2.38) and gastrectomy (HR, 2.58)[2] for patients who do not experience TO. The opportunity for transplant centers to use a composite quality metric predictive of survival, rather than survival alone, offers important opportunities for quality-directed healthcare delivery.

A link between TO and cost is established in hepatopancreatic surgery, in which Medicare payments among patients who achieved TO were markedly lower than among patients who did not, indicating that TO may also be a surrogate for value.[6] Similarly, Moris et al[25] report approximately $60 000 less in total hospital charges for LT recipients who experienced a TO. We have not yet analyzed costs in our national data set.

Great care must be taken when using the patient-level metric of TO as a center-level comparator. As Figure 2C illustrates, both low- and high-MELD patients achieve TO; however, low MELD strongly favors TO. These data clearly indicate that the application of TO as a center-level quality metric requires adjustment for case complexity. Aggressive transplant centers using donation after circulatory death and nonideal donors and routinely caring for high-acuity recipients have lower unadjusted center TO rates. Our goal is to measure quality across the spectrum of liver disease, enable equitable comparisons between centers, and do nothing that might discourage

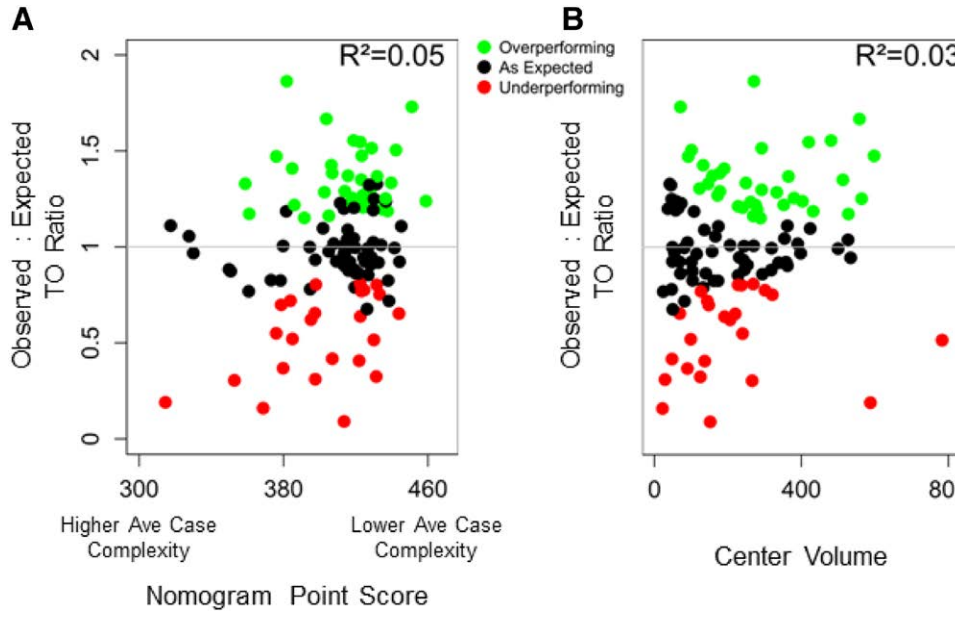


FIGURE 4. O:E TO rates by center as a function of (A) center-level average-case complexity and (B) center volume. Overperforming and underperforming centers have O:E ratios >1 or <1 with 95% confidence. O:E, observed:expected; TO, textbook outcome.

transplantation of critically ill patients. Acknowledging the limitations of MELD as a single variable,[26,27] we incorporate 9 variables predictive of TO in our risk adjustment. We find that quality, as measured by O:E center-level TO rates, is delivered to both low- and high-acuity patients by many transplant centers independent of case volume. These results are consistent with our goal of associating quality with all patients who experience a successful, efficient, and uncomplicated liver transplant. The O:E methods that we use are analogous to those used by the Centers for Medicare and Medicaid Services to compare quality across hospitals.[28]

By whom and how should TO be used? First and foremost, TO offers an opportunity for transplant programs to compare themselves, on a risk-adjusted basis, with peer institutions and identify target areas, in which they can improve patient experience. Even high-performing institutions will find domains within the composite outcome in which there is room for performance improvement. Many centers seeking to improve internal quality using our definition of TO would focus on length of stay and readmission.

For potential LT recipients, TO offers opportunity to use their own pretransplant characteristics to calculate the probability of having both a "smooth" perioperative course and enhanced 5-y survival. It is critical to note that not every quality metric should be co-opted as a regulatory metric. In kidney transplantation, it is well established that regulatory flagging decreases kidney offer acceptance and may actually decrease access to transplantation.[29] Many authors note the precarious balance wherein regulatory monitoring can inadvertently discourage transplantation of higher-risk candidates.[30,31] Similar concerns apply if payers choose to co-opt TO. Almost all transplants, regardless of TO status, are both lifesaving and cost-saving.[32-34]

Our work has several limitations. First, definitions of TO are admittedly arbitrary, and reaching professional consensus may be an iterative process. Our definition was formulated at our institution, and a national provider survey may be necessary to refine this definition. Second, in its truest form, TO is a patient-centric metric. We are actively working to correlate TO with patient-reported outcome data; however, we have not yet established this link. Finally, current liver transplant metrics publicly reported by the Scientific Registry of Transplant Recipients include a 5-tier quality score calculated in the domains of waitlist survival, transplant rate, and 1-y liver survival.[35] Within the transplant community, there is now widespread recognition that access to transplantation,[36] quality of care for waitlisted patients,[37] and time to transplant[38] are as important as the peritransplant and posttransplant experience, and these factors undoubtedly contribute to patients' perception of an ideal transplant experience. In some fields, "front end" variables, such as time from initial consult to colonoscopy,[10] have been incorporated into the definition of TO. At present, it is necessary to omit waitlist survival and transplant rate from center-level definitions of TO because of inadequate data addressing the burden of liver disease in different geographic regions,[39] lack of standardized listing practices across centers,[40-42] inequity in access to care,[43,44] and variability in organ procurement organization practice.[45,46] Nonetheless, our definition of TO in LT will improve when we are able to incorporate comprehensive data reflective of the entire transplant experience.

In summary, as patient and graft survival rates after LT continuously rise, quality improvement initiatives demand more meaningful measures of quality. We propose a novel definition of TO after LT, report a national TO rate of 37.4%, demonstrate that TO is predictive of long-term survival, and suggest that TO can be equitably applied as a center-level comparator.

ACKNOWLEDGMENTS

The authors gratefully acknowledge funding from The Ohio State University Center for Clinical and Translational Science (grant UL1TR00273). The authors thank Daria Faulkner for her assistance in article preparation.

REFERENCES

1. Kolfschoten NE, Kievit J, Gooiker GA, et al. Focusing on desired outcomes of care after colon cancer resections; hospital variations in 'textbook outcome.' *Eur J Surg Oncol.* 2013;39:156–163.
2. van der Kaaij RT, de Rooij MV, van Coevorden F, et al. Using textbook outcome as a measure of quality of care in oesophagogastric cancer surgery. *Br J Surg.* 2018;105:561–569.
3. Busweiler LA, Schouwenburg MG, van Berge Henegouwen MI, et al; Dutch Upper Gastrointestinal Cancer Audit (DUCA) group. Textbook outcome as a composite measure in oesophagogastric cancer surgery. *Br J Surg.* 2017;104:742–750.
4. Priego P, Cuadrado M, Ballestero A, et al. Comparison of laparoscopic versus open gastrectomy for treatment of gastric cancer: analysis of a textbook outcome. *J Laparoendosc Adv Surg Tech A.* 2019;29:458–464.
5. van Roessel S, Mackay TM, van Dieren S, et al; Dutch Pancreatic Cancer Group. Textbook outcome: nationwide analysis of a novel quality measure in pancreatic surgery. *Ann Surg.* 2020;271:155–162.
6. Merath K, Chen Q, Bagante F, et al. Textbook outcomes among Medicare patients undergoing hepatopancreatic surgery. *Ann Surg.* 2020;271:1116–1123.
7. Merath K, Chen Q, Bagante F, et al. A multi-institutional international analysis of textbook outcomes among patients undergoing curative-intent resection of intrahepatic cholangiocarcinoma. *JAMA Surg.* 2019;154:e190571.
8. Poelemeijer YQM, Marang-van de Mheen PJ, Wouters MWJM, et al. Textbook outcome: an ordered composite measure for quality of bariatric surgery. *Obes Surg.* 2019;29:1287–1294.
9. Karthaus EG, Lijftogt N, Busweiler LAD, et al; Dutch Society of Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit, the Dutch Institute for Clinical Auditing. Textbook outcome: a composite measure for quality of elective aneurysm surgery. *Ann Surg.* 2017;266:898–904.
10. Salet N, Bremmer RH, Verhagen MAMT, et al. Is textbook outcome a valuable composite measure for short-term outcomes of gastrointestinal treatments in the Netherlands using hospital information system data? A retrospective cohort study. *BMJ Open.* 2018;8:e019405.
11. Nolan T, Berwick DM. All-or-none measurement raises the bar on performance. *JAMA.* 2006;295:1168–1170.
12. Wey A, Salkowski N, Kasiske BL, et al. Five-tier utility: a start on the path to better reporting, in response to Schold and Buccini. *Am J Transplant.* 2019;19:608–609.
13. Schold JD, Buccini LD. Five-tier futility: this should end any remaining debate. *Am J Transplant.* 2019;19:607.
14. Chandraker A, Andreoni KA, Gaston RS, et al; AST/ASTS Transplant Metrics Taskforce. Time for reform in transplant program-specific reporting: AST/ASTS transplant metrics taskforce. *Am J Transplant.* 2019;19:1888–1895.
15. Bursac Z, Gauss CH, Williams DK, et al. Purposeful selection of variables in logistic regression. *Source Code Biol Med.* 2008;3:17.
16. Tibshirani R. The LASSO method for variable selection in the Cox model. *Stat Med.* 1997;16:385–395.
17. Muller X, Marcon F, Sapisochin G, et al. Defining benchmarks in liver transplantation: a multicenter outcome analysis determining best achievable results. *Ann Surg.* 2018;267:419–425.

18. Burroughs AK, Sabin CA, Rolles K, et al; European Liver Transplant Association. 3-month and 12-month mortality after first liver transplant in adults in Europe: predictive models for outcome. *Lancet.* 2006;367:225–232.
19. Jay CL, Lyuksemburg V, Ladner DP, et al. Ischemic cholangiopathy after controlled donation after cardiac death liver transplantation: a meta-analysis. *Ann Surg.* 2011;253:259–264.
20. Axelrod DA, Dzebisashvilli N, Lentine KL, et al. National assessment of early biliary complications after liver transplantation: economic implications. *Transplantation.* 2014;98:1226–1235.
21. Schold JD, Patzer RE, Pruett TL, et al. Quality metrics in kidney transplantation: current landscape, trials and tribulations, lessons learned, and a call for reform. *Am J Kidney Dis.* 2019;74:382–389.
22. Cross NB, Dittmer I. Quantity over quality: metrics in solid organ transplantation. *Transplantation.* 2018;102:1031–1032.
23. Brett KE, Ritchie LJ, Ertel E, et al. Quality metrics in solid organ transplantation: a systematic review. *Transplantation.* 2018;102:e308–e330.
24. Heinrich TW, Marcangelo M. Psychiatric issues in solid organ transplantation. *Harv Rev Psychiatry.* 2009;17:398–406.
25. Moris D, Shaw BI, Gloria J, et al. Textbook outcomes in liver transplantation. World J Surg. 2021;45:1258–1257.
26. Gotthardt D, Weiss KH, Baumgärtner M, et al. Limitations of the MELD score in predicting mortality or need for removal from waiting list in patients awaiting liver transplantation. *BMC Gastroenterol.* 2009;9:72.
27. Bittermann T, Makar G, Goldberg DS. Early post-transplant survival: interaction of MELD score and hospitalization status. *J Hepatol.* 2015;63:601–608.
28. Ash AS, Louis TA, Normand S-LT, et al. *Statistical Issues in Assessing Hospital Performance*. Population and Quantitative Health Sciences Publications; 2012.
29. Bowring MG, Massie AB, Craig-Schapiro R, et al. Kidney offer acceptance at programs undergoing a systems improvement agreement. *Am J Transplant.* 2018;18:2182–2188.
30. Jay C, Schold JD. Measuring transplant center performance: the goals are not controversial but the methods and consequences can be. *Curr Transplant Rep.* 2017;4:52–58.
31. Mikolajczyk AE, Rao VL, Diaz GC, et al. Can reporting more lead to less? The role of metrics in assessing liver transplant program performance. *Clin Transplant.* 2019;33:e13385.
32. Rana A, Gruessner A, Agopian VG, et al. Survival benefit of solid-organ transplant in the United States. *JAMA Surg.* 2015;150:252–259.
33. Pullen LC. The economy of transplant: saving money and saving lives. *Am J Transplant.* 2021;21:1353–1354.
34. Whiting JF, Kiberd B, Kalo Z, et al. Cost-effectiveness of organ donation: evaluating investment into donor action and other donor initiatives. *Am J Transplant.* 2004;4:569–573.
35. Wey A, Salkowski N, Kasiske BL, et al. A five-tier system for improving the categorization of transplant program performance. *Health Serv Res.* 2018;53:1979–1991.
36. Mathur AK, Ashby VB, Fuller DS, et al. Variation in access to the liver transplant waiting list in the United States. *Transplantation.* 2014;98:94–99.
37. Samuel D, Coilly A. Management of patients with liver diseases on the waiting list for transplantation: a major impact to the success of liver transplantation. *BMC Med.* 2018;16:113.
38. Axelrod D, Yeh H. Liver transplantation equity: supply, demand, and access. *Am J Transplant.* 2017;17:2759–2760.
39. Desai AP, Mohan P, Roubal AM, et al. Geographic variability in liver disease-related mortality rates in the United States. *Am J Med.* 2018;131:728–734.
40. Kling CE, Perkins JD, Biggins SW, et al. Listing practices and graft utilization of hepatitis C-positive deceased donors in liver and kidney transplant. *Surgery.* 2019;166:102–108.
41. Volk ML, Biggins SW, Huang MA, et al. Decision making in liver transplant selection committees: a multicenter study. *Ann Intern Med.* 2011;155:503–508.
42. Luo X, Massie AB, Haugen CE, et al. Baseline and center-level variation in simultaneous liver-kidney listing in the United States. *Transplantation.* 2018;102:609–615.
43. Ross-Driscoll K, Kramer M, Lynch R, et al. Variation in racial disparities in liver transplant outcomes across transplant centers in the United States. *Liver Transpl.* 2021;27:558–567.
44. Lynch RJ, Magliocca JF, Hundley JC, et al. Moving past "think local, act global": a perspective on geographic disparity. *Am J Transplant.* 2019;19:1907–1911.
45. Cannon RM, Jones CM, Davis EG, et al. Patterns of geographic variability in mortality and eligible deaths between organ procurement organizations. *Am J Transplant.* 2019;19:2756–2763.
46. Adler JT, Yeh H, Markmann JF, et al. Is donor service area market competition associated with organ procurement organization performance? *Transplantation.* 2016;100:1349–1355.