Potential Relationships Between NAFLD Fibrosis Score and Graft Status in Liver Transplant Patients

By Haaris Kadri[1], Raja R. Narayan MD MPH[2], Sunnie Y. Wong MD PhD[2], and Marc L. Melcher MD PhD[2]

[1] Department of Molecular and Cell Biology, University of California, Berkeley
[2] Department of Surgery, Stanford University

ABSTRACT

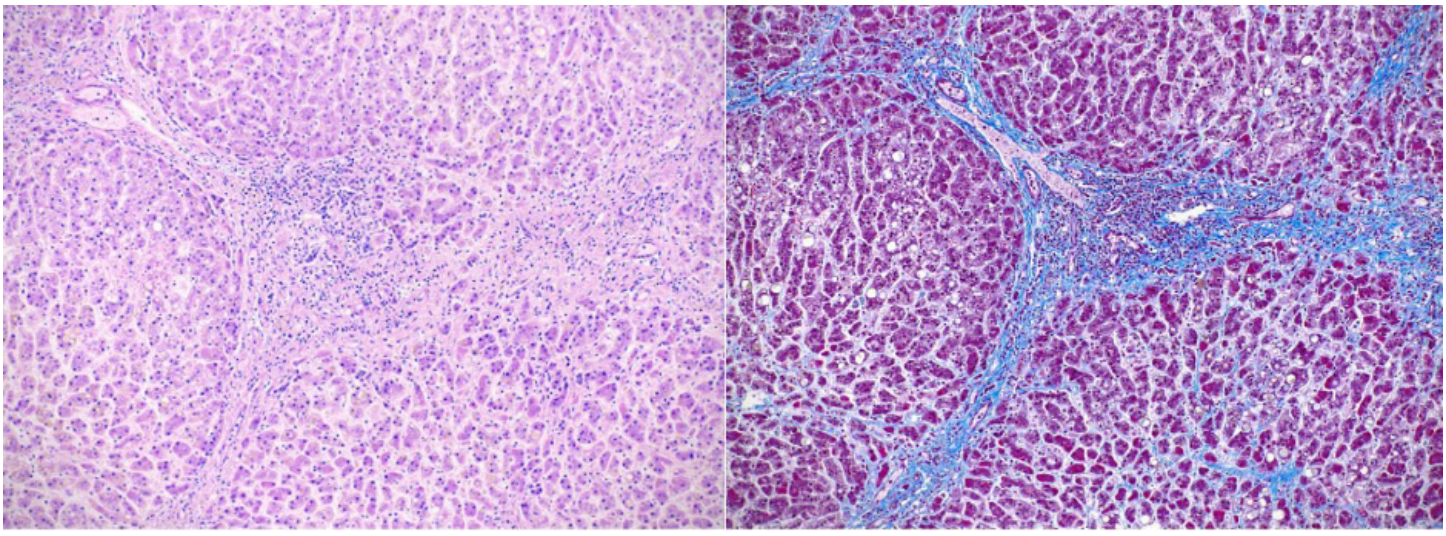
Non-alcoholic fatty liver disease is projected to be the most common cause of liver failure in the coming decade and is a very common reason for liver transplantation. One measure of its severity is the level of hepatic fibrosis, traditionally assessed by a liver biopsy. The non-alcoholic fatty liver disease fibrosis score was developed to non-invasively predict the degree of fibrosis using patient characteristics and laboratory values. We hypothesized that this score could also be used to assess the quality of donated livers, since many donors are obese and thus have a higher risk of fatty liver disease. Using data from the United Network for Organ Sharing over two decades, this study tests whether graft failure is associated with the donor liver's non-alcoholic fatty liver disease fibrosis score. Statistical analysis yielded that the relationship between the score and time till graft failure is insignificant: A chi-square test of independence between the two gives a *p*-value of .1311, and a Kaplan-Meier survival analysis yielded a *p*-value of .2, neither of which were under the significance level of .05. Though the results were not statistically significant, future studies on non-invasive assessments and their use may illuminate possibilities for clinical applications.

INTRODUCTION

Liver transplantation is the only definitive treatment for patients with end-stage liver disease.[1] Increase in the incidence of non-alcoholic fatty liver disease (NAFLD)-related liver failure has caused it to become one of the most common reasons for liver transplantation. The severity of the disease itself can range from hepatic steatosis (lipid retention in the liver) to steatohepatitis (lipid retention coupled with inflammation of liver tissue), and it can lead to advanced cirrhosis and hepatocellular carcinoma, among other ill sequelae.[1] The decision to use a liver from a deceased donor for transplantation into a recipient depends on histopathological, physical, and biochemical factors. Therefore, the increase in the prevalence of NAFLD, which damages the liver at the cellular level, reduces the quality of the liver donor pool and can be problematic when procuring organs for transplantation.[2]

Currently, the gold standard method to assess a donor liver's quality is the histologic review of a biopsy obtained at procurement time. This review screens for the degree of hepatic steatosis, hepatocellular injury and disease, inflammation, and fibrosis; all of which can preclude transplantation.[3] However, there are significant limitations to this method. For example, the biopsy only assesses a small portion of the liver parenchyma, the functional tissue of the organ, and might not represent the global histopathology of

Figure 1: Different levels of fibrosis in liver parenchyma. *This image shows the five accepted stages of fibrosis common in pathological analysis of liver fibrosis scoring, F0-F4. F0: no fibrosis, F1: portal fibrosis without septa, F2: few septa, F3: numerous septa without cirrhosis, F4: cirrhosis.*



Donor	n = 366	
Age (years)	Median Age	47
Sex (%)	% Male	40.7104
	% Female	59.2896
BMI (kg/m²)	Median BMI	28.3008
COD (%)	Anoxia	38.2514
	CVA/Stroke	36.3388
	Head Trauma	24.3170
	CNS Tumor	0
	Other	1.0929
CIT (hours)	Median CIT	7.3

Recipient	n = 366	
Age (years)	Median Age	58
Sex (%)	% Male	32.5137
	% Female	67.3967
BMI (kg/m²)	Median BMI	26.7760
Graft Status (%)	Survived	80.3279
	Failed	19.6721
Time to Graft Failure (days) (n = 361)	Median Time	1101

Table 1: Demographics Data of all Included Donor and Recipient Pairs. *Abbreviations: BMI: Body Mass Index, COD: Cause of Death, CVA: Cardiovascular Accident, CNS: Central Nervous System, CIT: Cold Ischemia Time.*

the liver. Histological assessment can also vary from pathologist to pathologist, which is problematic when it comes to clearly and effectively communicating the severity of a biopsy reading and can have severe consequences when the liver is being vetted for transplantation.[4,5]

Given the aforementioned limitations of the liver biopsy, there is an interest in developing alternative tools to avoid biopsies when possible.6 One such tool, the NAFLD fibrosis score (NFS), predicts the level of hepatic fibrosis from patient age, BMI, and diabetes status, as well as several laboratory values including aspartate aminotransferase (AST) and alanine aminotransferase (ALT), platelet, and albumin levels. The NFS is used to predict postoperative complications and mortality and was developed specifically for use in NAFLD patients.[7,8] It can be calculated using readily-available patient data to provide results almost instantly, giving patients a lower-cost alternative to biopsy.[9] A high NFS value indicates a high probability for advanced fibrosis and a low NFS value indicates the absence of advanced fibrosis.[7] Notably, the NFS does not predict low fibrosis risk, but merely the presence or absence of high-level fibrosis. In particular, metrics like the NFS are known for their high negative predictive value when it comes to eliminating the potential for high-grade and advanced fibrosis.[10] Negative predictive value (NPV), in this case, is the probability that a subject does not have some high-grade fibrosis given that their NFS does not indicate that they do. Due to its high NPV, the NFS most accurately detects advanced fibrosis in afflicted livers when used alongside other non-invasive indexes, like FIB-4 and BARD scores.[11]

Graft failure is a postoperative risk for transplantation patients and can lead to death. Within one year of transplantation, 10% of livers fail.[12] Also, in the United States, as of 2013, approximately 3000 people on the transplant waitlist either die or become too sick to undergo transplant per year, and in 2017, 742 livers were thrown out post-procurement for a variety of reasons, including histological anomalies,[13, 14] leading to not only a loss of life, but

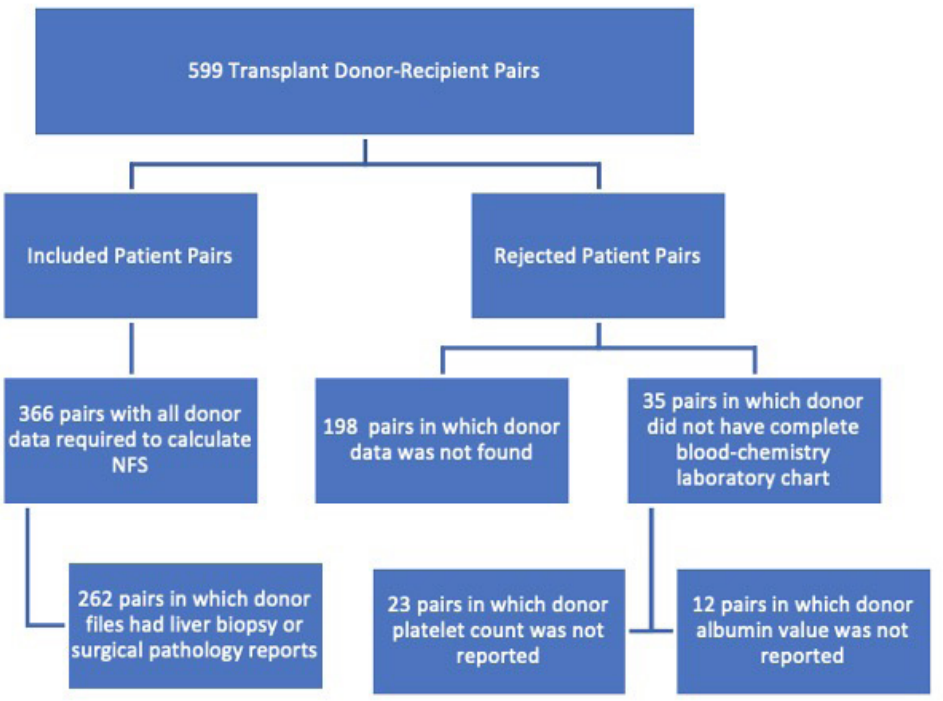


Figure 2: Flowchart of Patient Inclusion. *This chart shows how many patients started off in the initial database and how the number was whittled down upon enforcement of the inclusion criteria.*

NAFLD Fibrosis Score	Correlated Fibrosis Severity
- 1.455 > Score (Low)	F0 - F1 (Low severity)
-1.455 < Score < 0.675 (Medium)	F2 (Moderate severity)
Score > 0.675 (High)	F3 - F4 (High severity)

Table 2: NAFLD Fibrosis Scores and Their Assigned Fibrosis Severity. *The above NAFLD Fibrosis score ranges are best associated with their paired pathologist scores, which can be seen as ranked on a scale from 0 to 4. F0: no fibrosis, F1: mild fibrosis, F2: moderate fibrosis, F3: severe fibrosis, F4: cirrhosis.*

also a loss of healthcare dollars. Thus, it is important to develop stronger predictors for liver success and outcomes post-transplant. Even though the NFS was not created to assess liver graft failure, we hypothesize that it would be indicative, to some degree, of postoperative graft failure and time to graft failure.

METHODS

Patient Selection

Patients were selected from a prospectively maintained United Network for Organ Sharing (UNOS) database of liver donations and transplants performed at a single center between September 1999 and May 2020. The database culls information from the Organ Procurement and Transplantation Network (OPTN) at the national, regional, and individual levels. Liver transplant donor-recipient pairs with the following available data were included in the study: donor NFS, donor liver biopsy data, and recipient graft survival. Relevant donor liver fibrosis data was collected from either available liver biopsy reports or final surgical pathology reports. From a database of 599 transplantation procedures, 366 had the data necessary for inclusion (**Figure 2**). Demographics data for both donors and recipients was collected and tabulated as well (**Table 1**).

Determining Postoperative Outcomes for Donors

Graft survival and time to failure were noted. The distinction between primary and secondary graft failure was also established. As is common convention in transplantation, primary graft failure is defined as complications with the initial engraftment of the liver, and secondary graft failure is defined as complications after engraftment.[16] Due to ambiguity or lack of specification regarding the level of graft failure, this study did not differentiate between the two types when doing analysis, though this distinction is standard practice.

Calculating NAFLD Fibrosis Scores

To calculate NFS for each donor liver, the following formula was utilized.[7] The variable *diabetes_score* was quantified on a binary scale, with a patient diagnosed with diabetes assigned a score of 1 and a patient not diagnosed with diabetes assigned a score of 0.[7]

$$\mathrm{NFS} = -1.675 + [0.037 \cdot \mathrm{age(years)}] + \left[0.094 \cdot \mathrm{BMI}\left(\frac{\mathrm{kg}}{\mathrm{m}^2}\right)\right] + (1.13 \cdot \mathrm{diabetes\_score}) + \left(0.99 \cdot \frac{\mathrm{AST}}{\mathrm{ALT}}\right) - (0.013 \cdot \mathrm{platelet\_count}) - \left[0.66 \cdot \mathrm{albumin}\left(\frac{\mathrm{mg}}{\mathrm{dl}}\right)\right]$$

Analysis of NAFLD Fibrosis Score Categorization and Pathologist Scoring

The donor NFS values were split into three groups with associated risks of fibrosis. All low NAFLD scores (F0 – F1) were assigned 0, medium scores (F2) were assigned 1, and high scores (F3 – F4) were assigned 2. Pathologist scoring for fibrosis is on a scale from 0 to 4, so scores in the range [0, 1) were assigned 0, [1, 3) were assigned 1, and [3, 4] were assigned 2. In the analysis, pathologist scoring is provided in ranges because of discrepancies in how it was reported. For example, some livers were given non-integer fibrosis scores, such as 1.5. In other cases, right-lobe and left-lobe fibrosis readings were inconsistent. In these instances, the average of the two reported values was used.

Statistical Analysis of the Relationship Between NAFLD Fibrosis Score and Graft Failure

A statistical analysis of the patient's NFS and graft failure was done by separating the sample of 366 patients into two groups: one group of patients with graft failure and one group without. The groups were then broken into three strata: those with an NFS less than -1.455, those with a score between -1.455 and 0.675, and those

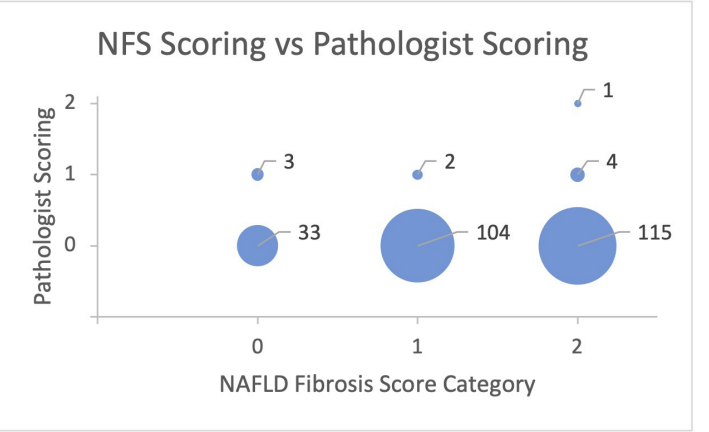


Figure 3: Bubble Plot of NAFLD Score Category and Pathologist Scoring. *This graph plots the intersections between NFS categorization and pathologist scoring in three dimensions. Points with no bubbles indicate no agreement between NFS Score category and pathology scoring.*

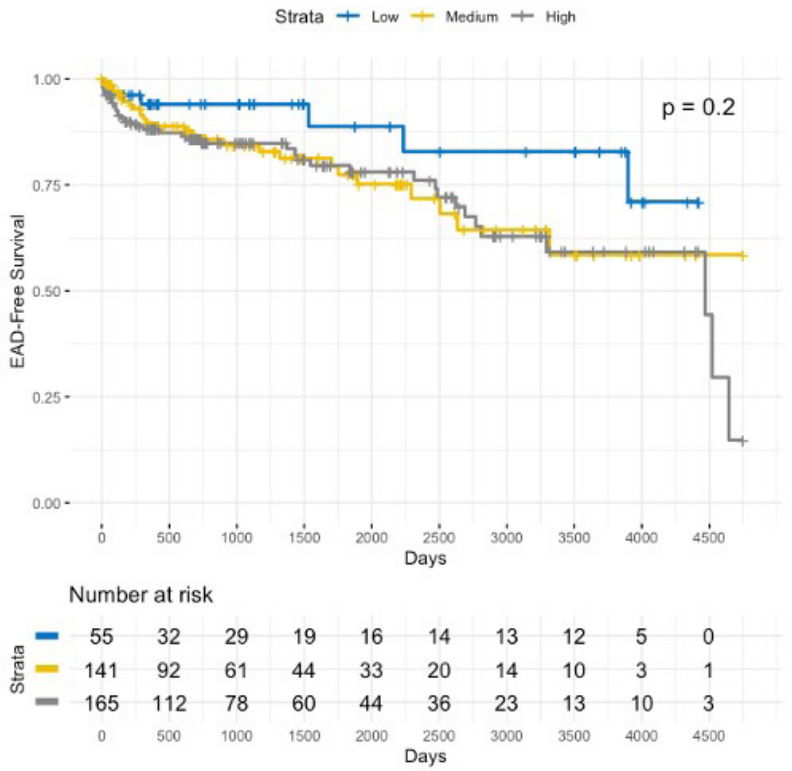


Figure 4: Kaplan-Meier Survival Curve for Graft Survival by Pre-Determined NAFLD Scores. *This set of curves shows the survival of the 266 patients postoperatively as split up into the three NFS categories.*

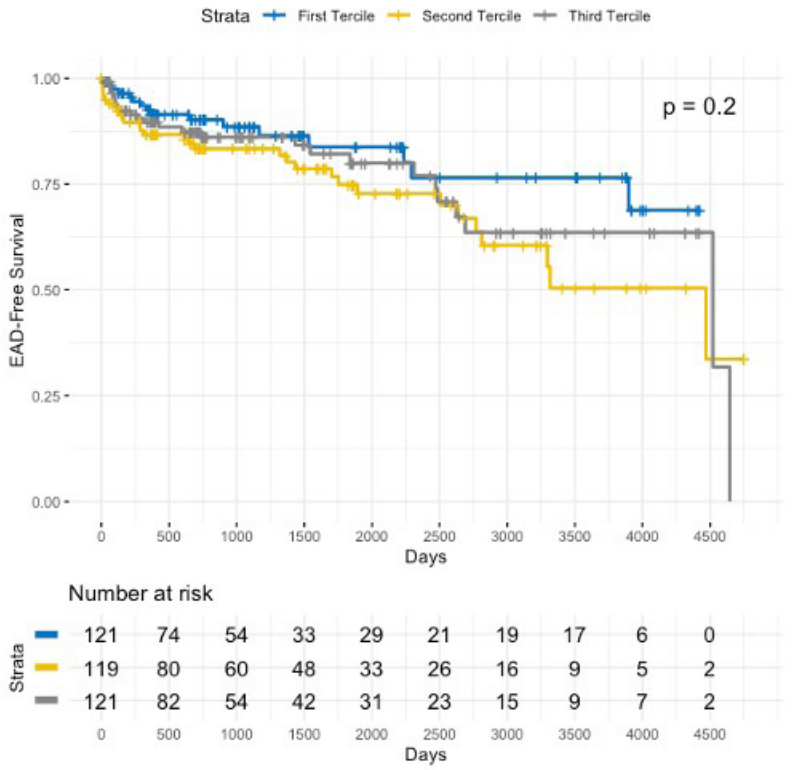


Figure 5: Kaplan-Meier Survival Curve for Graft Survival by NAFLD Score Terciles. *This set of curves shows the survival of the 266 patients postoperatively as split up into three equally sized terciles.*

with a score higher than 0.675 (**Table 2**).[7] The proposed null hypothesis was that graft survival in recipients was independent of NFS scores; this hypothesis was tested using a chi-square test of independence.

A Kaplan-Meier estimation analysis was used to test whether donor NFS scores were associated with time to graft failure. Data was censored when the time to graft failure was not recorded for a patient. The recipients were split into three groups in two different ways, and their survival curves were graphed and tested for any statistical significance. In the first categorization (**Figure 3**), the patients were split into groups based on the pre-determined NFS categories of low, moderate, and high severity fibrosis risk of their transplanted organs. In the second analysis (**Figure 4**), the patients were split into terciles (lower third, middle third, and upper third), with 122 patients in each. The null hypothesis for these tests was that the three groups, in each of the two tests, had similar time-to-failure, or survival times.

RESULTS

Correlation of the NAFLD Fibrosis Score with Pathologist-Reported Fibrosis Scoring

Of the 366 patients included in this study, 262 had pathology data including assessment of fibrosis. The results of the scoring comparisons are compiled in the bubble plot (**Figure 5**), where the *x*-axis is the NAFLD score category, the *y*-axis is the pathologist scoring as per the donor reports, and the size of the bubble itself indicates how many instances there were of that specific intersection. Instances of agreement, shown on the graph as (0,0), (1,1), and (2,2), only account for 36 of the 262 (13.74%) donor livers for which fibrosis data was available.

Relationship between NFS and Graft Failure in Recipients

Out of the 366 recipients included in the study, 72 (19.67%) of them ended up with graft failure. The graft failure incidence at 90 and 365 days was 4% and 10% respectively. The graft outcomes were broken down by NFS severity.[7] While there appears to be a trend towards greater graft loss with higher severity, there was no significant difference in graft failure at 90 (χ^2 (2, N = 366) = 1.9232, p = .3823) and 365 days (χ^2 (2, N = 366) = 1.6022, p = .4488) (**Table 3**).

Kaplan-Meier Survival Analysis for the Determination of Correlation Between NFS and Graft Survival Time

The 366 recipients and their graft survival times were analyzed in two ways through a Kaplan-Meier survival analysis, for which they were split into three groups. Although the survival curve displayed a trend toward better graft survival for the low fibrosis risk group in the first analysis and the first tercile group in the second, this association was not statistically significant (p = .2).

Donor NFS Severity	90 Day Graft Loss (%)	365 Day Graft Loss (%)
Low Severity	3.6	5.5
Medium Severity	2.8	9.8
High Severity	6.0	11.3
Total	4.4	9.8

Table 3: Graft Outcome by Donor NFS Severity. *This table shows the relative percentages of graft loss observed in patients with low, medium, and high NFS severity. The "total" row on the bottom shows a complete percentage of 90- and 365-day graft loss.*

DISCUSSION

Graft failure after liver transplantation can result in death if re-transplantation cannot be performed. Thus, it is important to identify donor livers that are less likely to fail. We hypothesized that a higher donor liver NFS would be associated with a greater risk for graft failure in recipients due to the association between NFS and liver fibrosis, a risk factor in transplantation.

Graft failure was identified in 4.37% of recipients 90-days post-transplant, and 9.84% of recipients one-year post-transplant. Although graft survival was lower at higher NFS, this was not found to be significant. This could either be due to the limited power of this study or the true lack of an association between the NFS and the degree of fibrosis linked to graft failure. More data and collaboration with other centers is being pursued to conduct a higher-powered study to address the former. Regarding the latter, research is being done to find a tool that can diagnose fibrosis associated with graft failure. Employment of artificial intelligence (AI) on this front has been fruitful. The use of AI in histopathological assessment yields accurate, objective results that can be analyzed quickly to provide information regarding associations between histopathological features of the liver tissue and patient outcomes post-transplant.[15]

Very little association was found between donor NFS and biopsy liver fibrosis in pathologist assessments (**Figure 5**). It is not surprising that the pathologist scoring of fibrosis was low in general, as these livers came from a database of livers accepted for donation. Few donor livers selected for transplantation would have high NFS. The NFS may also be inappropriately affected by the donor circumstance of death. As an example, if a patient died of an overdose and was anoxic for a period, their liver may have been impacted leading to increased AST and ALT levels on the lab tests. Elevated AST and ALT increase the NFS, thus, overcalling fibrosis risk.

Donor NFS may also not predict early graft failure well because it only accounts for the quality of the donor organ and not the health of the recipient. Future studies could investigate whether the combination of donor NFS with recipient-specific factors, such as MELD score, would better predict graft-survival post-transplantation. This combination may determine which donor livers are suitable for donation and which potential recipients cannot risk receiving a liver that has histopathological anomalies.

The biopsy remains the gold standard of pathological analysis. It is difficult, however, to compare biopsy results to one another because of the variability in pathologist analysis.[4,5] Inconsistencies in the interpretation of reports between pathologists can lead to more issues with decisions, which are important to avoid in high-risk situations like a transplant. Additionally, although all transplants analyzed in this study were conducted at the same medical center, misinterpretation and lack of clarity between different transplant centers can have adverse impacts.[16] The area of highest agreement between the NFS and pathologist assessment was with low-scoring livers (12.60% of available fibrosis data), which means that there are conditions for which NFS could supplement a liver biopsy to a fair level of accuracy. Further analysis of data and non-invasive tools is important in this regard.

The study was limited by the available data. The initial database accessed had 599 patient files, but 233 (38.90%) of them had to be excluded from the study for being incomplete. Out-of-date files were excluded only if the transplant occurred outside of the study period. Assuming these documents were filled out completely and accurately, the addition of the excluded patients could have led to more insightful statistical results in the end.

Due to the prevalence of liver diseases like NAFLD, liver transplantation is becoming increasingly common. To avoid re-transplantation and recipient mortality, it is important to be certain of the donor liver's viability. Though it could not be concluded that a donor liver's NFS was associated with recipient graft failure, there is the possibility that the use of other tools, or NFS in conjunction with other tools, may prove to be an adequate indicator of patient outcomes, reducing or even eliminating the need for a biopsy in cases, which could provide a more holistic analysis of the potential risks involved in a given transplant procedure, one of which is postoperative graft failure in the recipient. While the biopsy is the "status quo" assessment when assessing livers for transplantation, using noninvasive tools avoids a lot of the subjectivity that comes with pathological assessment, while also being time-efficient and affordable.

ACKNOWLEDGEMENTS

This work was supported in part by Health Resources and Services Administration contract HHSH250-2019-00001C. The content is the responsibility of the authors alone and does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This research was conducted with help from the Department of Surgery at Stanford University.

REFERENCES

1. Pais, R., Barritt, A. S., Calmus, Y., Scatton, O., Runge, T., Lebray, P., Poynard, T., Ratziu, V., & Conti, F. (2016). NAFLD and liver transplantation: Current burden and expected challenges. *Journal of Hepatology*, *65*(6), 1245–1257. https://doi.org/10.1016/j.jhep.2016.07.033
2. Patel, Y. A., Berg, C. L., & Moylan, C. A. (2017). Nonalcoholic fatty liver disease: Key considerations before and after liver transplantation. *Digestive Diseases and Sciences*, *61*(5), 1406–

1416. https://doi.org/10.1007/s10620-016-4035-3
3. Cheah, M. C., McCullough, A. J., & Goh, G. B.-B. (2017). Current modalities of fibrosis assessment in non-alcoholic fatty liver disease. *Journal of Clinical and Translational Hepatology*, *5*(3), 261–271. https://doi.org/10.14218/JCTH.2017.00009
4. Bracamonte, E., Gibson, B. A., Klein, R., Krupinski, E. A., & Weinstein, R. S. (2016). Communicating uncertainty in surgical pathology reports: A survey of staff physicians and residents at an academic medical center. *Academic Pathology*, *3*. https://doi.org/10.1177/2374289516659079
5. Coffin, C. S., Burak, K. W., Hart, J., & Gao, Z. (2006). The impact of pathologist experience on liver transplant biopsy interpretation. *Modern Pathology*, *19*(6), 832–838. https://doi.org/10.1038/modpathol.3800605
6. Dyson, J. K., Anstee, Q. M., & McPherson, S. (2014). Non-alcoholic fatty liver disease: A practical approach to diagnosis and staging. *Frontline Gastroenterology*, *5*(3), 211–218. https://doi.org/10.1136/flgastro-2013-100403
7. Angulo, P., Hui, J. M., Marchesini, G., Bugianesi, E., George, J., Farrell, G. C., Enders, F., Saksena, S., Burt, A. D., Bida, J. P., Lindor, K., Sanderson, S. O., Lenzi, M., Adams, L. A., Kench, J., Therneau, T. M., & Day, C. P. (2007). The NAFLD fibrosis score: A noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology*, *45*(4), 846–854. https://doi.org/10.1002/hep.21496
8. Treeprasertsuk, S., Björnsson, E., Enders, F., Suwanwalaikorn, S., & Lindor, K. D. (2013). NAFLD fibrosis score: A prognostic predictor for mortality and liver complications among NAFLD patients. *World Journal of Gastroenterology*, *19*(8), 1219–1229. https://doi.org/10.3748/wjg.v19.i8.1219
9. McPherson, S., Stewart, S. F., Henderson, E., Burt, A. D., & Day, C. P. (2010). Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. *Gut*, *59*(9), 1265. https://doi.org/10.1136/gut.2010.216077
10. Robinson, A., & Wong, R. J. (2020). Applications and limitations of noninvasive methods for evaluating hepatic fibrosis in patients with nonalcoholic fatty liver disease. *Clinical Liver Disease*, *15*(4). https://doi.org/10.1002/cld.878
11. Xiao, G., Zhu, S., Xiao, X., Yan, L., Yang, J., & Wu, G. (2017). Comparison of laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: A meta-analysis. *Hepatology*, *66*(5), 1486–1501. https://doi.org/10.1002/hep.29302
12. George, L. A., Lominadze, Z., & Kallwitz, E. R. (2018). Con: Patient and graft survival outcome at 1 and 3 years posttransplant do not reflect the value of successful liver transplantation. *Clinical Liver Disease*, *11*(3), 62–65. https://doi.org/10.1002/cld.670
13. Hart, A., Schladt, D. P., Zeglin, J., Pyke, J., Kim, W. R., Lake, J. R., Roberts, J. P., Hirose, R., Mulligan, D. C., Kasiske, B. L., Snyder, J. J., & Israni, A. K. (2016). Predicting outcomes on the liver transplant waiting list in the United States: Accounting for large regional variation in organ availability and priority allocation points. *Transplantation*, *100*(10), 2153–2159. https://doi.org/10.1097/TP.0000000000001384
14. Chu, M. J. J., Dare, A. J., Phillips, A. R. J., & Bartlett, A. S. J. R. (2015). Donor hepatic steatosis and outcome after liver transplantation: A systematic review. *Journal of Gastrointestinal Surgery*, *19*(9), 1713–1724. https://doi.org/10.1007/s11605-015-2832-1
15. Yang, L., Ghosh, R. P., Franklin, J. M., Chen, S., You, C., Narayan, R. R., Melcher, M. L., & Liphardt, J. T. (2020). NuSeT: A deep learning tool for reliably separating and analyzing crowded cells. *PLOS Computational Biology*, *16*(9), Article e1008193. https://doi.org/10.1371/journal.pcbi.1008193
16. Colling, R., Verrill, C., Fryer, E., Wang, L. M., & Fleming, K. A. (2014). Discrepancy rates in liver biopsy reporting. *Journal of Clinical Pathology*, *67*(9), 825–827. https://doi.org/10.1136/jclinpath-2014-202261