

Liver discard rate due to conservative estimations of steatosis: an inference-based approach

Hao Guo BSc<sup>1,2</sup>; Boris L. Gala-Lopez MD, PhD<sup>3</sup>; Ian P.J. Always MD;<sup>4</sup> Kevin C. Hewitt PhD<sup>1</sup>

<sup>1</sup>Department of Physics and Atmospheric Science, Dalhousie University, 6310 Coburg Road, Halifax, NS B3H 4R2, Canada

<sup>2</sup>Department of Medical Physics, Nova Scotia Health Authority, 5820 University Avenue Halifax, NS B3H 1V7, Canada

<sup>3</sup>Department of Surgery, Dalhousie University, 6-300 Victoria Building, 1276 South Park Street, Halifax, NS B3H 2Y9, Canada

<sup>4</sup>Department of Surgery, LUMC Transplant Center, Leiden University Medical Center, 2300 RC Leiden, The Netherlands

Correspondence information:

Kevin C. Hewitt, 6310 Coburg Road, PO Box 15000, Halifax, NS Canada B3H 4R2. E-mail:

[kevin.hewitt@dal.ca](mailto:kevin.hewitt@dal.ca)

Each author's specific contributions to the work:

H.G. conceived of the study and its design; acquisition, analysis, and interpretation of data, drafting and critically revising the manuscript for important intellectual content. K.C.H. provided oversight of the work in their supervisory capacity of the trainee H.G., K.C.H., B.L.G-L. and I.P.J.A. were involved in the critical revision of the manuscript. All authors approved the final version to be published.

Disclosure:

The authors declare no conflicts of interest.

Funding:

H.G. is support by the Vitamin Scholarship, Dalhousie University and the Research Nova Scotia Doctoral Award, Research Nova Scotia.

Abbreviations used in this paper: ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; CAP, controlled attenuation parameter; HS, hepatic steatosis; IQR, interquartile range; IQRC, interquartile range of final controlled attenuation parameter; LT, liver transplantation; M, medium; MaS, macrovesicular steatosis; MiS, microvesicular steatosis; NHANES, National Health and Nutrition Examination Survey; OPTN, Organ Procurement and Transplantation Network; SD, standard deviation; XL, extra-large

## Abstract

**Background:** On-site conservative estimations of steatosis could result in the unnecessary discard of donor livers. This study applied the body mass index (BMI) as an independent statistical indicator to determine the extent of this problem. We explored two hypotheses: (I) that because of varying levels of expertise and protocols (reputational risk for pathologists), biopsies at transplant centers overestimate hepatic steatosis (HS), and (ii) that non-biopsy donor liver assessments are more conservative than biopsy-based evaluations.

**Methods:** The study processed cross-database and intra-database comparisons using data from the National Health and Nutrition Examination Survey (NHANES) and Organ Procurement and Transplantation Network (OPTN) spanning January 2017 to March 2020 in the United States. Post-matching BMI was applied as an independent indicator of statistical risk of HS.

**Results:** *Contrary to our first hypothesis*, biopsies at transplant centers did **not** overestimate HS - biopsy-classified donor livers were found in similar or lower risk categories. *Consistent with our second hypothesis*, absent biopsies, evaluations before and during organ procurement **were** observed to be more conservative, leading to the discard of 11.9% (373) of potential donor livers. .

**Conclusions:** The study concludes that there was a significant (11.9%) disparity caused by on-site non-biopsy assessments of HS, leading to the unnecessary discard of potential donor livers. The findings emphasize the need to develop more accurate intraoperative techniques for assessing HS to optimize donor liver procurement.

**Keywords:** Liver transplantation; Hepatic steatosis; Body mass index; Elastography; Liver biopsy

## 1 Introduction

Hepatic steatosis (HS), especially moderate and severe macrovesicular steatosis (MaS) of donor livers, is associated with an increased risk of graft dysfunction after liver transplantation (LT).<sup>1-4</sup> Although marginal donor livers have been accepted more and more often by transplant centers, HS remains a critical reason for discard.<sup>5</sup> Livers with severe MaS ( $\geq 60\%$ ) are usually considered high risk for graft dysfunction and are often discarded outright.<sup>6</sup> Livers with microvesicular steatosis (MiS) and mild (5%-30%) MaS are usually safe for LT, but transplanting livers with moderate (30-60%) MaS is still controversial.<sup>6-9</sup> Furthermore, there seems to be disagreement between transplant physicians on the exact risk attributed to this steatosis. As a result, surgeons from different centers act differently when offered a steatotic organ.<sup>10</sup> Therefore, accurate classification and quantification of steatotic donor livers are essential.

The current qualification of HS of donor livers relies on semi-subjective and subjective methods.<sup>11</sup> Liver biopsy, though with limitations such as invasiveness, sampling errors, inter-observer variability, weak reproducibility, and occasional low on-site availability, is currently the “gold standard” for assessing HS of donor’s livers.<sup>7,12,13</sup> However, a biopsy is only performed in a portion of donor surgeries. The most widely used method is the surgeon’s visual quality inspection during organ procurement surgeries. An online survey indicated that one-third of surgeon members of the American Society of Transplant Surgeons replied that visual inspection played a critical role in their decision to transplant more than 50% of donor livers.<sup>14</sup> In contrast with visual estimation, only 30% of donors after cardiac death had data from biopsies between 2004 and 2010 in the United States.<sup>3</sup>

Surgeon's visual quality inspection is based on the stiffness and macroscopic appearance of livers during organ procurement surgeries. Experienced surgeons can accurately predict <30% MaS; however, surgeons' prediction of  $\geq 30\%$  MaS was correct only for 52.2% of the cases, according to a prospective, double-blind study of 201 procured donor livers.<sup>15</sup> Compared with the visual quality inspection, the degree of HS in biopsies is assessed more quantitatively; however, it shows a considerable variation among assessments of pathologists.<sup>13</sup> It is also reported that the morphological semiquantitative evaluation of the degree of HS was constantly overestimated in a study on 75 liver biopsies.<sup>16</sup>

The low accuracy and high variability of the degree of HS assessed by the visual inspection and biopsies bring uncertainties to transplanting livers with MaS ranging 20%-60%. A natural hypothesis is that to ensure LT's safety, transplant centers tend to estimate the degree of HS and make the decision of LT using a **conservative** approach, especially in cases without the availability of liver biopsies. Assuming this hypothesis is true, a considerable percentage of donor livers with "actual" MaS<30% are rejected before and/or during organ procurement. Since the "actual" degree of HS cannot be determined using the current gold standard, an alternative method is needed to quantify the discrepancies between the prevalence of HS and the frequency of HS of donor livers based on on-site assessments.

Since the information on the degree of HS is unavailable from donors who did not receive liver biopsies, the body mass index (BMI) can be used as an indirect intermediate parameter to statistically describe "potential" HS in each group of analyses. The risks of HS have been found significantly and independently correlated with increasing BMI in a nonlinear fashion.<sup>17-19</sup> BMI was also found an independent predictor of the extent of fat infiltration.<sup>20</sup> As a predictor for HS,

the receiver operating characteristic (ROC) curve analysis of BMI were reported with an area under the receiver operator characteristic curve (AUROC) ranging from 0.77 to 0.84.<sup>21,22</sup>

In the present study, by adopting BMI as an independent indicator of the statistical HS risk, we applied an inference-based approach to quantify unrecovered steatotic donor livers impacted by on-site conservative estimations of biopsy and non-biopsy assessments. The analysis is supported by data from publicly available national clinical registries.

## **2 Materials and Methods**

We analyzed the National Health and Nutrition Examination Survey (NHANES) 2017-March 2020 pre-pandemic data and the Organ Procurement and Transplantation Network (OPTN) data in the same timeframe in the United States. This research does not require Research Ethics Board review because it relies exclusively on information that is publicly available through a mechanism set out by regulation and that is protected by law, according to the 2018 Tri-Council Policy Statement for the Ethical Conduct for Research Involving Humans.

Designed to assess the health and nutritional status of adults and children in the United States, the NHANES examines a nationally representative sample every year. The National Center for Health Statistics combined the 2019-March 2020 data (uncompleted collection in the 2019-2020 cycle) with the 2017-2018 data using additional weighting procedures to provide nationally representative estimates.<sup>23</sup>

Although NHANES is population-wide, whereas OPTN captures donors, the two registries are comparable. 97% of the donor deaths collected by OPTN in 2009-2020 were caused by anoxia, stroke, head trauma, or central nervous system tumor, the incidence of which are stochastic and independent from BMI or liver diseases.<sup>24</sup> Considering that the online database

developed by OPTN contains data about every organ donation and transplant event in the United States since 1987,<sup>25</sup> the OPTN data can also be regarded as nationally representative.

All statistical data analyses using the NHANES and OPTN data (Table 1) were performed using Stata Version 17.0-Basic Edition (StataCorp LLC, TX, USA). P-values of less than 0.05 were considered statistically significant.

Table 1. Variables of retrieved data from NHANES and OPTN registries.

NHANES Variable	Description	OPTN Variable	Description
SEQN	Respondent sequence number	RECOVERY_DATE_DON	Recovery date (sent to the operating room)
BMXBMI	Body mass index of the participant	BMI_DON_CALC	Donor body mass index
RIAGENDR	Gender of the participant	GENDER_DON	Donor gender
RIDAGEYR	Age in years at screening	AGE_DON	Age in years
LUAXSTAT	Elastography exam status	MACRO_FAT_LI_DON	Macro vesicular fat (%)
LUANMVGP	Count: complete measures from the final wand	MICRO_FAT_LI_DON	Micro vesicular fat (%)
LBDHEM	Hepatitis E IgM (anti-HEV)	LI_BIOPSY_DON	Liver biopsy
LUXCAPM	Median controlled attenuated parameter (CAP)	LI_DISPOSITION	Liver disposition
LUXCPIQR	Controlled attenuated parameter (CAP) interquartile range (IQRc)	LI_REASON_CD	Liver disposition reason
LBXHCR	Hepatitis C RNA	HCV_NAT	Hepatitis C nucleic acid amplification testing
LBDHBG	Hepatitis B surface antigen	HBV_SUR_ANTIGEN_DON	Hepatitis B surface antigen status
ALQ170	Past 30 days # times 4-5 drinks on an occasion	ALCOHOL_HEAVY_DON	Heavy alcohol use (heavy=2+ drinks/day)
LBXSASSI	Aspartate aminotransferase	SGOT_DON	Terminal lab aspartate aminotransferase
LBXSATSI	Alanine aminotransferase	SGPT_DON	Terminal lab alanine aminotransferase
LBDSTBSI	Total bilirubin	TBILL_DON	Terminal lab total bilirubin
LBXSNASI	Sodium	SODIUM170_VAL_DON	Last Serum Sodium
LUAPNME	Exam wand type		

## 2.1. The NHANES Data

Starting in 2017, liver ultrasound transient elastography measured by FibroScan<sup>®</sup> was included in the NHANES. The controlled attenuation parameter (CAP) algorithm assesses ultrasonic attenuation and therefore examines the degree of HS.<sup>26</sup> Following the Liver Ultrasound Transient Elastography Procedures Manuals, the FibroScan<sup>®</sup> model 502 V2 Touch



equipped with a medium (M) or extra-large (XL) probe was used to obtain the elastography measurements and record the CAP measurements.<sup>27,28</sup>

Although many studies have shown good performance of CAP for excluding moderate and severe HS in potential living liver donors, the reported optimal cut-offs of CAP varied and could bring significant uncertainty in grading the degree of HS.<sup>7,29–31</sup> To apply the most reasonable CAP cut-offs for this present cross-database study, we referred to a recent meta-analysis including 13 studies and 2,346 biopsy-controlled CAP datasets collected using the FibroScan® M and XL probes<sup>32</sup>. In this present study, for data collected by the XL probes, the CAP cut-offs of non-HS ( $S_0$ ) versus mild HS ( $S_1$ ),  $S_1$  versus moderate HS ( $S_2$ ), and  $S_2$  versus severe HS ( $S_3$ ) were defined as 297 dB/m, 313 dB/m, and 333 dB/m, respectively.<sup>32</sup> For data collected by the M probes, the CAP cut-offs of  $S_0$  versus  $S_1$ ,  $S_1$  versus  $S_2$ , and  $S_2$  versus  $S_3$  were defined as 290.5 dB/m, 306.5 dB/m, and 326.5 dB/m, respectively, since the XL probes have been shown to negatively affect the CAP scores by 6.5 dB/m.<sup>32</sup>

## 2.2. The OPTN Data

In 2017-March 2020, the OPTN database recorded 36,014 deceased donors from whom at least one solid organ was recovered.

Deceased liver donations were classified into several categories for matching and analyses, as Table 2 describes. The MaS and HS cut-offs of  $S_0$  versus  $S_1$ ,  $S_1$  versus  $S_2$ , and  $S_2$  versus  $S_3$  were defined as 5%, 30%, and 60%, respectively.

Table 2. Classification of liver donations registered in OPTN.

Category	OPTN Variable	Value
Biopsy	LI_BIOPSY_DON	“Y”
Non-biopsy	LI_BIOPSY_DON	“N”
For transplant	LI_DISPOSITION	3 and 4
Not for transplant	LI_DISPOSITION	3 and 4

Non-recovery for reasons not related to steatosis	LI_REASON_CD	201, 202, 203, 204, 206, 207, 208, 209, 210, 211, 213, 214, 220, 221, 222, 223, 224, 225, 226, 295, and 299
Ruled out after evaluation in operating room	LI_REASON_CD	218

### 2.3. Exclusion and Matching of Data

The included NHANES dataset consisted of 3,149 subjects (Figure 1). Elastography measurements were included in the present study if (1) the examination contained at least ten complete measurements, (2) the participant was at least 18 years old, (3) the participant did not claim heavy alcohol use (defined as five or more alcoholic drinks for males or four or more alcoholic drinks for females on at least five days in the past month),<sup>33</sup> (4) the participant did not carry viral hepatitis, and (5) the interquartile range of final CAP (IQRc) measures was non-zero and smaller than 40 dB/m<sup>32,34,35</sup>.

The included OPTN dataset consisted of 14,172 subjects (Figure 1). Deceased liver donations were included in the present study if (1) an authorization of liver donation was obtained, (2) the donor was at least 18 years old, (3) the donor was not found to be a heavy alcohol user (defined as consuming more than two drinks per day),<sup>36</sup> (4) the participant did not carry viral hepatitis, (5) the information of liver disposition was specified, and (6) the donor liver was either recovered, transplanted, or not recovered for reasons related to steatosis.

To eliminate potential interference of risk factors other than HS and BMI on this study, according to the definition of a liver donor marginal by European Association for the Study of the Liver (EASL) Clinical Practice Guidelines,<sup>37</sup> elastography measurements and deceased liver donations were excluded from this study if any of the following criteria was met: age >65 years, serum sodium >165 mmol/L, alanine aminotransferase (ALT) >105 U/L, aspartate aminotransferase (AST) >90 U/L, or serum bilirubin >3 mg/dl.

In our study, we applied propensity score matching to match gender, age, BMI, ALT, AST, total bilirubin, and sodium of elastography participants (NHANES) and donors (OPTN) with biopsy records. As donor liver marginals had been excluded before matching, BMI was the only characteristic being matched for comparing donations (OPTN) with and without biopsy records. Propensity scores were calculated based on multivariable logistic regressions, where the dependent variable was the group identification (0 = group A, 1 = group B), and the independent variables were the characteristics to be matched.

Following the estimation of propensity scores, we used the “**psmatch2**” module in Stata to perform nearest-neighbor one-to-one matching.<sup>38</sup> We assessed the quality of the matches by checking the balance of the covariates across the treated and untreated groups after matching using the “**pstest**” command.

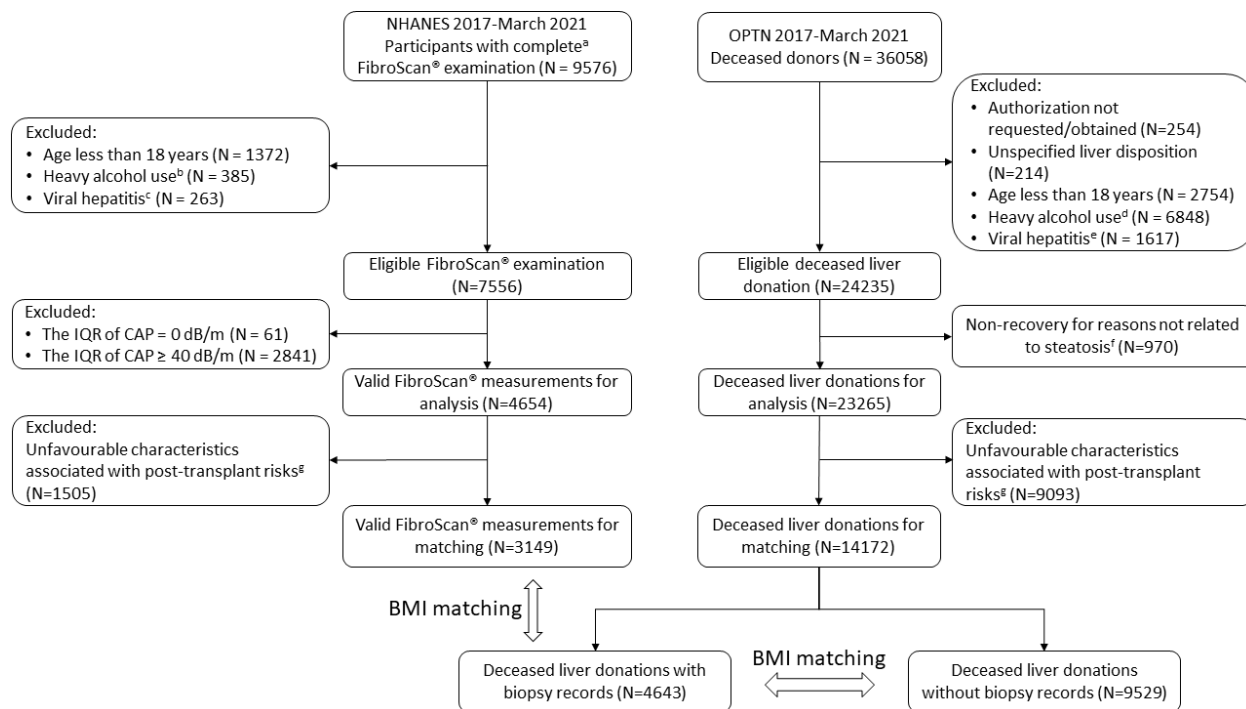


Figure 1. Flowchart of participants in the study of the prevalence of hepatic steatosis. (a) A complete FibroScan® examination contained at least ten complete measurements (the ratio interquartile range/median value of liver stiffness measurement (IQR/M) was not larger than 30%) of a patient fasting at least three hours. (b) Heavy alcohol use is defined as five or more alcoholic drinks for males or four or more alcoholic drinks for females on at least five days in the past month. (c) When the test result of any of hepatitis B surface antigen, hepatitis C RNA, hepatitis D antibody and hepatitis E IgM is positive, the participant was excluded. (d) The OPTN defines heavy alcohol use as consuming more than two drinks per day. (e) When the test result of any of hepatitis B surface antigen and hepatitis C RNA is positive, the donor was excluded. (f) The reasons not related to steatosis include cardiac arrest, positive hepatitis, positive HIV, anatomical abnormalities, vascular damage, no recipient located, donor medical history, donor social history, surgical damage in the operating room, hemodynamically unstable donor, trauma to the organ, time constraints, medical examiner restricted recovery, and other specify. (g) Participants and donors were excluded if at least one of the following criteria was met: age >65 years, serum sodium >165 mmol/L, ALT >105 U/L, AST >90 U/L, or serum bilirubin >3 mg/dl.

### 3 Results

The pre- and post-matching population characteristics of included participants in the (NHANES) elastography study and donors registered in OPTN with biopsy data are shown in Table 3, with the reporting of pre- and post-matching p-values. Records of 1,325 participants and 1,325 donors were matched. The post-matching participants and donors were well-correlated, as there was no statistically significant difference ( $p > 0.05$ ) in the characteristics, including age, BMI, ALT, AST, total bilirubin, and sodium. The post-matching data enabled comparison of prevalence of HS assessed by elastography and biopsy, among participants and donors with statistically same characteristics.

Table 3. Pre- and post-matching population characteristics of participants in the elastography study and donors registered in OPTN with biopsy records.

Characteristics	Pre-matching			Post-matching		
	NHANES	OPTN-Biopsy	P	NHANES	OPTN-Biopsy	P
Gender, N (%)						
Female	1685 (53.5)	2327 (50.1)		637 (46.9)	637 (46.9)	
Male	1464 (46.5)	2316 (49.9)		688 (53.1)	688 (53.1)	
Age, yr, mean ± SD						
Mean ± SD	43.2 ± 14.1	47.2 ± 12.1	<0.001	47.0 ± 13.7	46.8 ± 12.3	0.790
Median (IQR)	44 (31-56)	49 (39-57)		50 (37-59)	49 (38-57)	
BMI, kg/m <sup>2</sup>						
Mean ± SD	29.9 ± 7.2	32.0 ± 8.2	<0.001	30.4 ± 7.3	30.4 ± 7.8	0.911
Median (IQR)	28.8 (25.1-33.5)	30.9 (26.0-36.5)		29.1 (25.3-34.1)	29.2 (24.5-34.8)	
ALT, U/L, median (IQR)						
Female	15 (12-22)	28 (18-46)	<0.001	19 (13-29)	19 (13-29)	0.651
Male	23 (17-32)	29 (19-46)	<0.001	25 (18-35.5)	23 (16-37)	0.831
AST, U/L, median (IQR)						
Female	17 (15-21)	29 (20-45)	<0.001	20 (16-26)	20 (15-27)	0.766
Male	21 (17-26)	31.5 (22-50)	<0.001	22 (19-29)	22.5 (16-32)	0.519
Total Bilirubin, mg/dL, median (IQR)						
Female	0.3 (0.2-0.5)	0.6 (0.4-0.8)	<0.001	0.4 (0.3-0.6)	0.4 (0.3-0.6)	0.845
Male	0.4 (0.3-0.6)	0.7 (0.4-1.0)	<0.001	0.5 (0.4-0.7)	0.5 (0.4-0.8)	0.388
Sodium, mg/dL, median (IQR)						
Female	140 (139-142)	149 (144-154)	<0.001	142 (140-143)	142 (138-146)	0.063
Male	141 (139-142)	149 (143.5-154)	<0.001	142 (140-143)	142 (139-145)	0.052

S<sub>0</sub>, non-steatosis; S<sub>1</sub>, mild steatosis; S<sub>2</sub>, moderate steatosis; S<sub>3</sub>, severe steatosis; SD, standard deviation; BMI, body mass index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; IQR, interquartile range. The median CAP scores classify S<sub>0</sub>-S<sub>3</sub>. Null hypothesis: there is a statistically significant difference in the distribution of a covariate between the treated and control groups after matching.

In Figure 2, a stacked column chart was used to compare the prevalence of HS in the post-matching groups. Considering moderate and severe steatosis as risky for transplantation, elastography classified the most (28%) participants (NHANES) into the “risky steatosis” category. Biopsy classified similar percentage (25%) of donors (OPTN) into the “risky steatosis” category based on the degree of overall HS, showing no overestimation which was due to the hypothesized conservative assessment of donor livers. Applying the degree of MaS (which is more clinically relevant to graft dysfunction) as the classifier, significantly less (13%) donors

were classified into the “risky steatosis” category, indicating that biopsy can lessen unnecessary discarding of steatotic donor livers.

The elastography-based CAP score shows significant disagreement with biopsy in classifying HS and MaS in  $S_0$  and  $S_1$  stages. On condition that the post-matching population characteristics had no significant statistical difference, CAP scored classified the most elastography participants as non-steatotic (61%), whereas biopsies scored classified the most liver donors into mild overall HS (45%) or mild MaS (46%). The disagreement may support that biopsy is more conservative in classifying  $S_0$  and  $S_1$ ; however, as both  $S_0$  and  $S_1$  are considered “non-risky” in LT, this finding did not support that donor HS was assessed in a statistically conservative approach by biopsy.

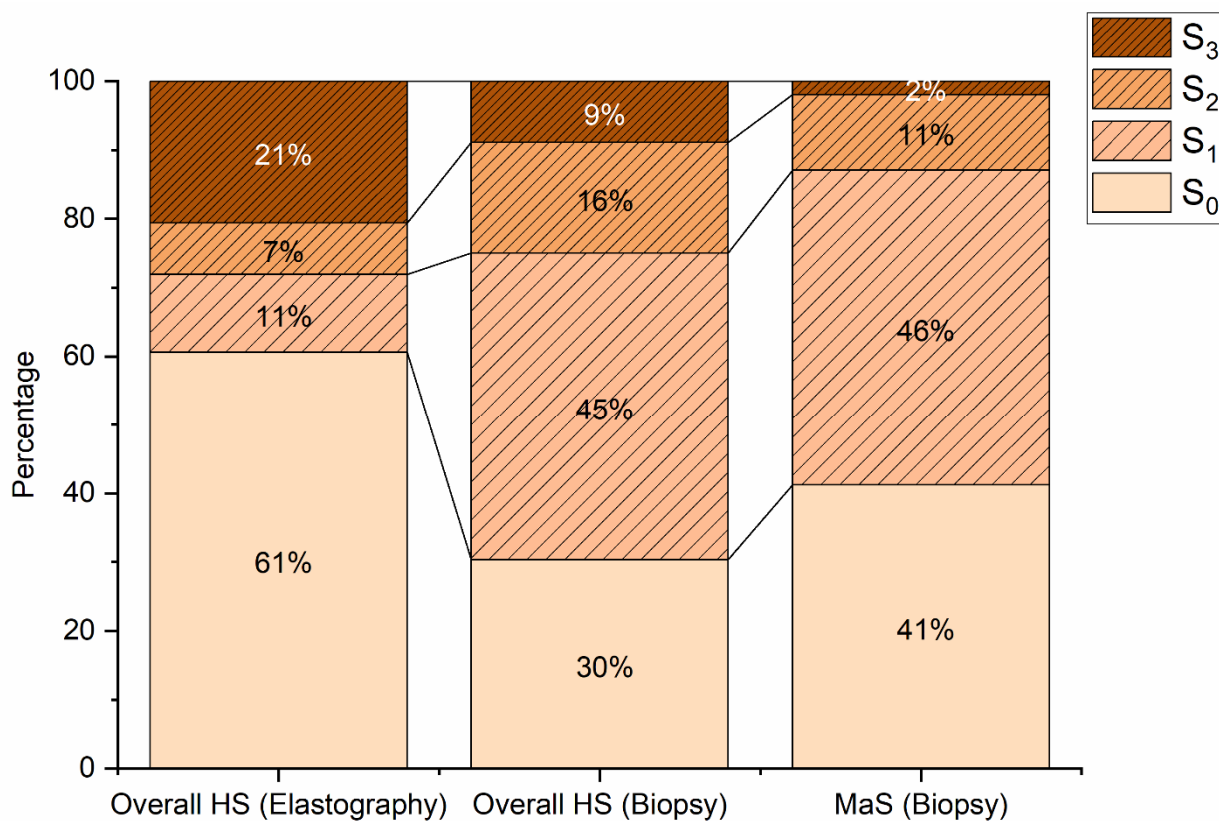


Figure 2. stacked column chart the prevalence of steatosis in the post-matching groups.  $S_0$ , non-steatosis;  $S_1$ , mild steatosis;  $S_2$ , moderate steatosis;  $S_3$ , severe steatosis. The median CAP scores and biopsy records classify  $S_0$ - $S_3$ . Elastography classified more HS into  $S_0$  and  $S_3$ , while biopsy classified less HS into the “risky” group  $S_2$  &  $S_3$

The pre- and post-matching population characteristics of the included donors registered in OPTN with and without biopsy records are shown in Table 4, with the reporting of pre- and post-matching p-values of the “pstest”. Since most donors (9,529, 67%) did not have biopsy records, only 4,264 donors without biopsy records were matched. The post-matching donors with and without biopsy records had no statistically significant difference ( $p>0.05$ ) in the distribution of covariate between BMIs, indicating excellent BMI matching in this intra-group comparison. Although this less rigorous matching solely targeted BMI, the values of other risk factors, no matter their statistical differences were significant or not, were all within the safe ranges because the donations including unfavored characteristics had already been excluded before matching.

Table 4. Pre- and post-matching population characteristics of donors registered in OPTN.

Characteristics	Pre-matching		P	Post-matching		
	Biopsy	Non-biopsy		Biopsy	Non-biopsy	P
Gender, N (%)						
Female	2327 (50.1)	3791 (39.8)		2077 (48.7)	1864 (43.7)	
Male	2316 (49.9)	5738 (61.2)		2187 (51.3)	2400 (56.3)	
Age, yr, mean $\pm$ SD						
Mean $\pm$ SD	47.2 $\pm$ 12.1	39.6 $\pm$ 13.6	<0.001	47.2 $\pm$ 12.1	41.1 $\pm$ 13.3	<0.001
Median (IQR)	49 (39-57)	39 (27-51)		49 (39-57)	42 (30-53)	
BMI, kg/m <sup>2</sup>						
Mean $\pm$ SD	32.0 $\pm$ 8.2	28.0 $\pm$ 6.5	<0.001	30.8 $\pm$ 7.1	30.8 $\pm$ 7.1	0.934
Median (IQR)	30.9 (26.0-36.5)	26.9 (23.5-31.2)		30.1 (25.6-34.9)	30.1 (25.6-34.9)	
ALT, U/L, median (IQR)						
Female	28 (18-46)	28 (18-46)	0.907	28 (18-45)	29 (18-48)	0.052
Male	29 (19-46)	29 (19-46)	0.194	29 (19-46)	30 (20-48)	0.630
AST, U/L, median (IQR)						
Female	29 (20-45)	31 (20-47)	0.039	30 (19-45)	32 (21-48)	0.002
Male	31.5 (22-50)	34 (22-52)	<0.001	31 (20-48)	33 (22-51)	0.002
Total Bilirubin, mg/dL, median (IQR)						
Female	0.6 (0.4-0.8)	0.5 (0.4-0.8)	0.339	0.6 (0.4-0.9)	0.5 (0.375-0.8)	0.136
Male	0.7 (0.4-1.0)	0.6 (0.4-1.0)	0.010	0.7 (0.5-1.0)	0.6 (0.4-1.0)	0.041
Sodium, mg/dL, median (IQR)						
Female	149 (144-154)	148 (142-153)	<0.001	148 (143-153)	148 (142-153)	0.032
Male	149 (143.5-154)	148 (143-154)	0.041	149 (144-154)	148 (143-154)	0.531

S<sub>0</sub>, non-steatosis; S<sub>1</sub>, mild steatosis; S<sub>2</sub>, moderate steatosis; S<sub>3</sub>, severe steatosis; SD, standard deviation; BMI, body mass index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; IQR, interquartile range. The median CAP scores classify S<sub>0</sub>-S<sub>3</sub>. Null hypothesis: there is a

---

statistically significant difference in the distribution of a covariate between the treated and control groups after matching.

Figure **3(a)** compares the distributions of post-matching BMIs between the groups of all donors with and without biopsy records. The kernel density curves almost fully overlap each other, which confirms that inherent discrepancy of BMI distributions does not impact the comparisons between donors with and without biopsy records in the following sub-groups.

Figures **3(b)** shows a comparison of the distributions of post-matching BMIs between the group of donors with biopsy records (N=3,929) and those without biopsy (N=3,421) records in the group “For transplant.” Among the donors livers “for transplant”, those with biopsy records were more frequent than those without biopsy records in the high-BMI region (where BMI is higher than the cross point at 32.5 kg/m<sup>2</sup>). 4.4% donor livers with biopsy records similarly high statistical risks of HS were precured for transplant (the area difference under the curves on the right side of the dashed line).

Meanwhile, among the discarded donor livers, those in the high-BMI region (where BMI is higher than the cross point at 37.0 kg/m<sup>2</sup>) were more likely without biopsy records. Figures **3(c)** shows comparisons of the distributions of post-matching BMIs between the group of donors with biopsy records (N=335) and those without biopsy (N=843) records in groups “Not for transplant.” 11.9% of donor livers “Not for transplant” with similarly high statistical risks of HS did not have biopsy data recorded. Converting the percentage numbers to the pre-matching OPTN records, 11.9% of 1,487 livers “not for transplant” without biopsy records (out of 9,529 liver donations without biopsy records for the data matching) led to 177 unrecovered donor livers. Enlarging the applied scope to the 24,365 eligible organ donors registered in the United States in 2017-March 2020, 11.9% of 3,407 livers “not for transplant” without biopsy records led to 373 unrecovered donor livers.

Figures 3(d) compares the distributions of post-matching BMIs between the group of donors with biopsy records (N=142) and ruled out after evaluation in the operating room and the group of donors without biopsy records (N=140) and ruled out after evaluation in the operating room. Applying 28.4 kg/m<sup>2</sup> (the corresponding BMI value at the lower cross point of the kernel density curves) as cut-off, comparing those with BMI lower than 28.4 kg/m<sup>2</sup>, donors ruled out after evaluation in the operating room and with BMI larger than 28.4 kg/m<sup>2</sup> were 15.9% more frequent to be without biopsy results.

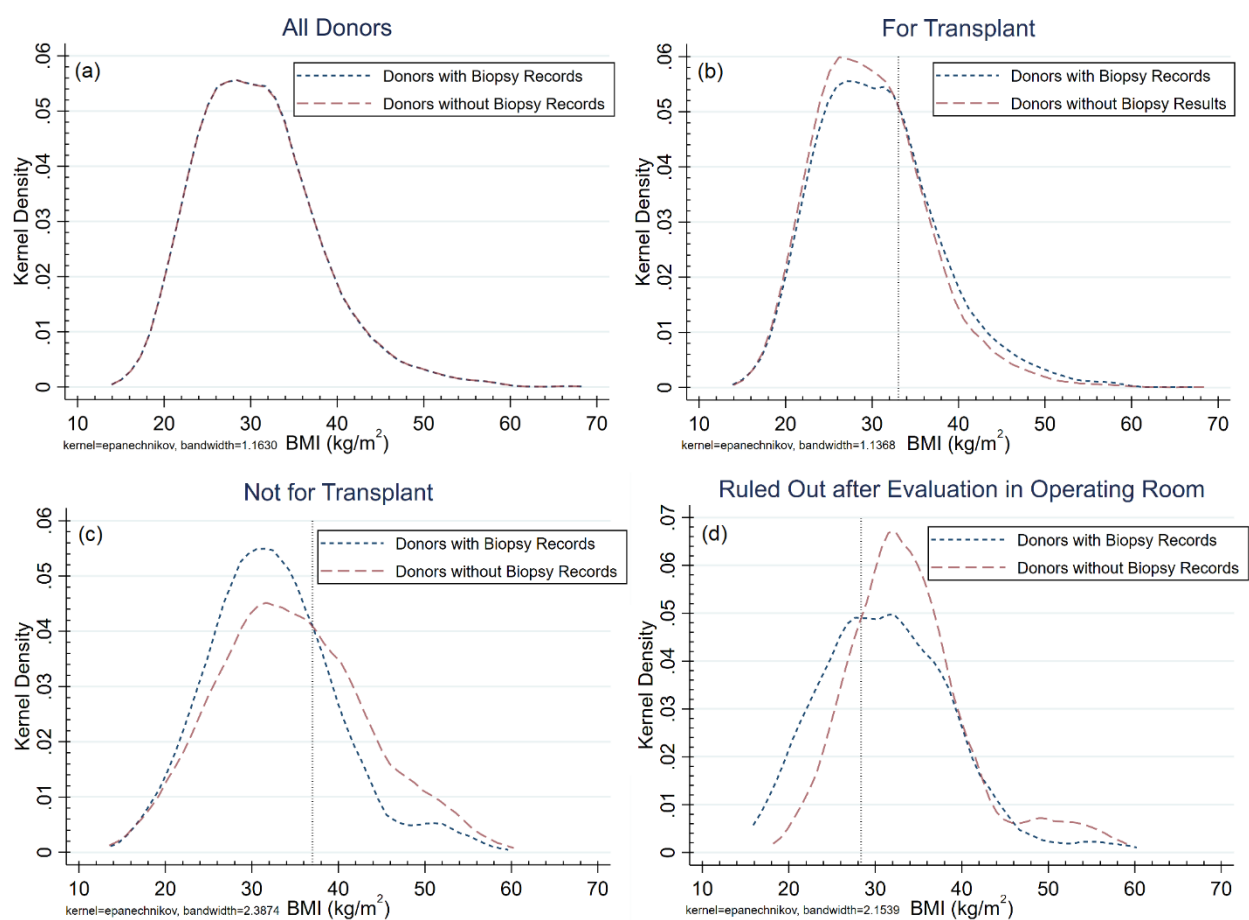


Figure 3. Univariate kernel density estimations of post-matching BMIs among subjects of different groups. (a) Comparison of BMIs between the group donors with and without biopsy records. (b) Comparison of BMIs between the group of donors without biopsy records and those with biopsy records in the group "For Transplantation." The vertical dashed line at 32.5 kg/m<sup>2</sup> indicated the corresponding BMI value at the cross point of the kernel density curves. (c) Comparison of BMIs between the group of donors without biopsy records and those with biopsy records in the group "Not for Transplant." The vertical dashed line at 37.0 kg/m<sup>2</sup> indicated the corresponding BMI value at the cross point of the kernel density curves. (d) Comparison of BMIs between the group of donors without biopsy records and ruled out after evaluation in the operating room and those with biopsy records ruled out after evaluation in the operating room. The vertical dashed line at 28.4 kg/m<sup>2</sup> indicated the corresponding BMI value at the lower cross point of the kernel density curves.



## 4 Discussions and Conclusion

This present study adopted BMI as an independent indicator of the statistical risk of HS and tested two hypotheses: (1) transplant centers tend to overestimate the degree of HS and make the decision of LT using a conservative biopsy, and (2) non-biopsy donor liver assessments are more conservative than biopsy-based assessments.

Tests on the first hypothesis were performed by comparing the population prevalence of HS in the United States in 2017-March 2020 and the prevalence of HS and MaS in American liver donors in the same timeframe. Potential interference of risk factors other than HS and BMI were excluded according to the definition of a liver donor marginal by EASL Clinical Practice Guidelines.<sup>37</sup> To enable cross-database comparisons, rigorous data matching eliminated significant statistical differences between the NHANES population data of participants and the OPTN data of donors in gender, age, BMI, ALT, AST, total bilirubin, and sodium. Although biopsy classified less HS to the degree  $S_0$  and more to the degree  $S_1$ , this finding did not support that transplant centers assessed donor HS in a statistically conservative approach by biopsy, as both  $S_0$  and  $S_1$  are considered “non-risky” in LT. On the contrary, 13% less donors were classified into the “risky steatosis” category according to the percentage of MaS, compared with classifications based on CAP scores.

The second hypothesis was evaluated utilizing a less stringent BMI matching strategy. Given that both the biopsy and non-biopsy data originated from the same OPTN registry, and that liver donors on the marginal fringe had been excluded prior to matching, the matching process was solely dependent on the donors' BMI. As BMI is a positively correlated independent risk factor of HS,<sup>17-19</sup> it was utilized as a surrogate to estimate the statistical risks of HS impacting the liver discard rates, specifically in the context of non-biopsy assessments of steatosis.

The analysis of post-matching OPTN data corroborated that on-site non-biopsy evaluations of donor livers tended to be more conservative than their biopsy-inclusive counterparts. Notably, 4.4% of donor livers with biopsy records exhibited comparably high statistical risks of HS but were still selected for transplantation. Conversely, among the discarded donor livers, biopsy data was absent for 11.9% of those presenting similar high statistical risks of HS. Translating these percentages into real numbers based on OPTN records, the conservative on-site non-biopsy assessment of HS inadvertently led to the loss of 177 to 373 donor livers in the United States from 2017 to March 2020. This suggests that an additional 177 to 373 livers could potentially have been recovered without surpassing the accepted risk threshold for HS.

A prominent disparity caused by on-site non-biopsy assessments HS could be observed in the operating room. Analysis of the post-matching BMI distributions between donors with and without biopsy records revealed that donors without biopsy records were discarded 15.9% more frequently compared to donors with biopsy records and comparable high BMIs.

While HS typically presents as a mixed form of MaS and MiS, the negative impact of MiS on post-LT dysfunction is significantly less pronounced than MaS.<sup>39,40</sup> Therefore, determining HS in biopsied donor livers based on the degree of MaS has greater clinical relevance. However, the HS defined by CAP scores is not directly comparable to biopsy-defined MaS, despite the referenced meta-analysis being based on biopsy-controlled CAP datasets.<sup>32</sup> Nevertheless, both overall HS (25%) and MaS (13%) assessments by biopsy classified fewer donor livers as "high-risk HS" compared to population prevalence (28%), effectively refuting the hypothesis that transplant centers tend to conduct overly conservative biopsies on donor livers.

A limitation of this study is that elastography measurements from NHANES were not biopsy-validated, although the CAP cut-offs were optimized using biopsy-controlled CAP datasets.

Although CAP cut-offs varied according to cause,<sup>32</sup> we believe our utilized cut-offs were appropriate, given the exclusion of viral hepatitis and marginal donor livers. This assumption is indirectly supported by a 2020 histology-based study reporting a 38% prevalence of non-alcoholic fatty liver disease in a middle-aged US cohort,<sup>41</sup> closely matching our 39% prevalence of CAP-based HS.

Another constraint is that to fulfill data matching criteria, most of the data (58%) remained untreated, which negatively affected the data volume. This study also purposefully omitted considerations of compound factors influencing donor liver quality. Although we excluded non-recoveries due to non-steatosis related reasons, livers classified as "not for Transplant" could be attributed to numerous factors, including poor quality of blood vessels, malignancies, fibrosis, technical procurement issues, and HS. However, these factors might not be thoroughly considered when extrapolating statistical HS risks to BMIs.

In this study, we translated comparisons between HS degrees in donor livers with and without biopsy records into comparisons of HS risks, using BMI as an indirect intermediate parameter to statistically depict HS risks. One advantage of this inferential approach is that it allowed us to quantify the impact of conservative on-site steatosis estimations on unrecovered donor livers on a statistical level. Nevertheless, the current research did not disclose how on-site assessments, including both biopsies and non-biopsy evaluations, impact liver recovery for individual donors. Individual case discrepancies might exceed statistical discrepancies due to uncertainties inherent in current subjective and semi-subjective HS assessment methods. Most donors (9,529, 67%) did not have biopsy records. Although the OPTN data did not enable us to track if a biopsy was operated before or during the organ procurement surgeries, the majority of the recorded biopsies were believed intraoperative biopsies, since most organ procurement organizations reported that

5% to 10% of their donors received a prerecovery liver biopsy, according to a study in 2017.<sup>42</sup>

The development of accurate and objective intraoperative techniques for assessing HS and MaS in the future will be vital to mitigate the impacts of conservative on-site estimations.

#### Reference:

1. Mergental H, Laing RW, Kirkham AJ, et al. Transplantation of discarded livers following viability testing with normothermic machine perfusion. *Nat Commun* 2020 111. 2020;11(1):1-12. doi:10.1038/s41467-020-16251-3
2. Moosburner S, Gassner JMGV, Nösser M, et al. Prevalence of Steatosis Hepatis in the Eurotransplant Region: Impact on Graft Acceptance Rates. *HPB Surg*. 2018;2018. doi:10.1155/2018/6094936
3. Orman ES, Barritt AS, Wheeler SB, Hayashi PH. Declining Liver Utilization for Transplant in the United States and the Impact of Donation after Cardiac Death. *Liver Transpl*. 2013;19(1):59. doi:10.1002/LT.23547
4. Spitzer AL, Lao OB, Dick AAS, et al. The biopsied donor liver: Incorporating macrosteatosis into high-risk donor assessment. *Liver Transplant*. 2010;16(7):874-884. doi:10.1002/LT.22085
5. Jackson KR, Motter JD, Haugen CE, et al. Temporal trends in utilization and outcomes of steatotic donor livers in the United States. *Am J Transplant*. 2020;20(3):855-863. doi:10.1111/AJT.15652
6. Linares I, Hamar M, Selzner N, Selzner M. Steatosis in Liver Transplantation: Current Limitations and Future Strategies. *Transplantation*. 2019;103(1):78-90. doi:10.1097/TP.0000000000002466
7. Mikolasevic I, Filipec-Kanizaj T, Mijic M, et al. Nonalcoholic fatty liver disease and liver transplantation - Where do we stand? *World J Gastroenterol*. 2018;24(14):1491. doi:10.3748/WJG.V24.I14.1491
8. McCormack L, Dutkowski P, El-Badry AM, Clavien PA. Liver transplantation using fatty livers: Always feasible? *J Hepatol*. 2011;54(5):1055-1062. doi:10.1016/J.JHEP.2010.11.004
9. Deroose JP, Kazemier G, Zondervan P, IJzermans JNM, Metselaar HJ, Alwayn IPJ. Hepatic steatosis is not always a contraindication for cadaveric liver transplantation. *HPB*. 2011;13(6):417-425. doi:10.1111/j.1477-2574.2011.00310.x
10. Imber CJ, St. Peter SD, Lopez I, Guiver L, Friend PJ. Current practice regarding the use of fatty livers: A trans-Atlantic survey. *Liver Transplant*. 2002;8(6):545-549. doi:10.1053/jlts.2002.31747
11. Neil DAH, Minervini M, Smith ML, Hubscher SG, Brunt EM, Demetris AJ. Banff consensus recommendations for steatosis assessment in donor livers. *Hepatology*. 2022;75(4):1014-1025. doi:10.1002/HEP.32208
12. Ratzu V, Charlotte F, Heurtier A, et al. Sampling Variability of Liver Biopsy in Nonalcoholic Fatty Liver Disease. *Gastroenterology*. 2005;128(7):1898-1906. doi:10.1053/J.GASTRO.2005.03.084
13. El-Badry AM, Breitenstein S, Jochum W, et al. Assessment of hepatic steatosis by expert pathologists: the end of a gold standard. *Ann Surg*. 2009;250(5):691-696. doi:10.1097/SLA.0B013E3181BCD6DD
14. Volk ML, Roney M, Merion RM. Systematic bias in surgeons' predictions of the donor-specific risk of liver transplant graft failure. *Liver Transplant*. 2013;19(9):987-990. doi:10.1002/LT.23683
15. Yersiz H, Lee C, Kaldas FM, et al. Assessment of hepatic steatosis by transplant surgeon and expert pathologist: A prospective, double-blind evaluation of 201 donor livers. *Liver Transplant*. 2013;19(4):437-449. doi:10.1002/lt.23615
16. Franzén LE, Ekstedt M, Kechagias S, Bodin L. Semiquantitative evaluation overestimates the degree of steatosis in liver biopsies: a comparison to stereological point counting. *Mod Pathol* 2005 187. 2005;18(7):912-916. doi:10.1038/modpathol.3800370
17. Fan R, Wang J, Du J. Association between body mass index and fatty liver risk: A dose-response

- analysis. *Sci Reports* 2018 81. 2018;8(1):1-7. doi:10.1038/s41598-018-33419-6
18. Katrina Loomis A, Kabadi S, Preiss D, et al. Body mass index and risk of nonalcoholic fatty liver disease: Two electronic health record prospective studies. *J Clin Endocrinol Metab.* 2016;101(3):945-952. doi:10.1210/JC.2015-3444/SUPPL\_FILE/JC-15-3444.PDF
  19. Miyake T, Kumagi T, Hirooka M, et al. Body mass index is the most useful predictive factor for the onset of nonalcoholic fatty liver disease: A community-based retrospective longitudinal cohort study. *J Gastroenterol.* 2013;48(3):413-422. doi:10.1007/S00535-012-0650-8/TABLES/10
  20. Angulo P, Keach JC, Batts KP, Lindor KD. Independent predictors of liver fibrosis in patients with nonalcoholic steatohepatitis. *Hepatology.* 1999;30(6). doi:10.1002/hep.510300604
  21. Rehm JL, Connor EL, Wolfgram PM, Eickhoff JC, Reeder SB, Allen DB. Predicting hepatic steatosis in a racially and ethnically diverse cohort of adolescent girls. *J Pediatr.* 2014;165(2). doi:10.1016/j.jpeds.2014.04.019
  22. Abeysekera KWM, Orr JG, Gordon FH, et al. Evaluating future risk of NAFLD in adolescents: a prediction and decision curve analysis. *BMC Gastroenterol.* 2022;22(1):1-13. doi:10.1186/S12876-022-02401-Y/TABLES/6
  33. Bohm MK, Liu Y, Esser MB, et al. Binge Drinking Among Adults, by Select Characteristics and State — United States, 2018. *MMWR Morb Mortal Wkly Rep.* 2021;70(41):1441-1446. doi:10.15585/MMWR.MM7041A2
  34. Wong VWS, Petta S, Hiriart JB, et al. Validity criteria for the diagnosis of fatty liver by M probe-based controlled attenuation parameter. *J Hepatol.* 2017;67(3):577-584. doi:10.1016/j.jhep.2017.05.005
  35. Oeda S, Tanaka K, Oshima A, Matsumoto Y, Sueoka E, Takahashi H. Diagnostic Accuracy of FibroScan and Factors Affecting Measurements. *Diagnostics* 2020, Vol 10, Page 940. 2020;10(11):940. doi:10.3390/DIAGNOSTICS10110940
  36. White J. Public Comment Proposal Data Collection to Evaluate Organ Logistics and Allocation OPTN Operations and Safety Committee. Accessed June 14, 2023. [https://bodandcommittees.unos.org/Staff/\\_layouts/15/WopiFrame2.aspx?sourcedoc=/Staff/Liaison Manual/Data Colle](https://bodandcommittees.unos.org/Staff/_layouts/15/WopiFrame2.aspx?sourcedoc=/Staff/Liaison Manual/Data Colle)
  37. Association for the Study of the Liver E. EASL Clinical Practice Guidelines: Liver transplantation q. *J Hepatol.* 2016;64:433-485. doi:10.1016/j.jhep.2015.10.006
  38. Leuven E, Sianesi B. PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Published online 2018. <https://econpapers.repec.org/RePEc:boc:bocode:s432001>
  39. Han S, Ko JS, Kwon G, et al. Effect of pure microsteatosis on transplant outcomes after living donor liver transplantation: A matched case-control study. *Liver Transplant.* 2014;20(4). doi:10.1002/lt.23824
  40. Croome KP, Lee DD, Croome S, et al. Does Donor Allograft Microsteatosis Matter? Comparison of Outcomes in Liver Transplantation With a Propensity-Matched Cohort. *Liver Transplant.* 2019;25(10). doi:10.1002/lt.25583
  41. Harrison SA, Gawrieh S, Roberts K, et al. Prospective evaluation of the prevalence of non-alcoholic fatty liver disease and steatohepatitis in a large middle-aged US cohort. *J Hepatol.* 2021;75(2). doi:10.1016/j.jhep.2021.02.034