DR. MANUELA CESARETTI (Orcid ID : 0000-0001-6419-1481)

DR. RAFFAELE BRUSTIA (Orcid ID : 0000-0001-6426-9533)

PROF. SARA MOCCIA (Orcid ID : 0000-0002-4494-8907)

Article type : Original Articles

Use of artificial intelligence as innovative method for liver graft macrosteatosis assessment

Cesaretti M[1,2,3], Brustia R[4], Goumard C[4], Cauchy F[1], Poté N[5], Dondero F[1], Paugam-Burtz C[6], Durand F[7], Paradis V[5], Diaspro A[2], Mattos L[8], Scatton O[4], Soubrane O[1], Moccia S[8,9,10]

ORCIDs: Cesaretti Manuela (0000-0001-6419-1481); Brustia Raffaele (0000-0001-6426-9533); Goumard Claire (0000-0001-9565-7264); Cauchy François (0000-0002-4997-3730), Poté Nicolas (0000-0001-7695-7317); Dondero Federica (0000-0001-7278-5554); Paugam-Burtz Catherine (0000-0002-8168-7152); Durand François (0000-0002-0357-1090); Paradis Valerie (0000-0003-1269-6407); Diaspro Alberto (0000-0002-4916-5928); Mattos Leonardo (0000-0003-4800-0010); Scatton Olivier (0000-0001-8313-3500); Soubranee Olivier (0000-0002-2059-1237); Moccia Sara (0000-0002-4494-8907)

AFFILIATIONS:

[1] Department of HPB Surgery and Liver Transplantation, Hôpital Beaujon AP-HP, Clichy, France

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/LT.25801

[2] IIT - Istituto Italiano di Tecnologia, Nanophysics Department, Genova, Italy

[3] Centre Hospitalier Universitaire de Nice - Digestive Surgery and Liver Transplantation Unit, Archet 2 Hospital, Nice, France.

[4] Department of HPB Surgery and Liver Transplantation, Hôpital de la Pitié-Salpêtrière, AP-HP, Paris, France

[5] Department of Pathology, Hôpital Beaujon, DHU UNITY, AP-HP, Clichy, France. INSERM UMR1149, Paris, France

[6] Department of Anesthesia and Intensive Care, DHU Unity, Paris 7 Diderot University, Sorbonne Paris Cite, 75013, Paris, France

[7] Hepatology & Liver Intensive Care, Hopital Beaujon, APHP, Clichy, France. INSERM U1149, Paris, France; University Paris Diderot, Paris, France.

[8] Department of Advanced Robotics, Istituto Italiano di Tecnologia, Via Morego 30, 16163, Genoa, Italy

[9] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy.

[10] Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy

Corresponding author:

Professor Olivier Soubrane, MD, PhD,

Department of HPB Surgery and Liver Transplantation, Hôpital Beaujon AP-HP,

100 Boulevard du General Leclerc, 92210 Clichy, France.

E-mail: olivier.soubrane@aphp.fr.

Tel 0033 01 40 87 52 52

Fax 0033 01 40 87 52 53

ABBREVIATIONS:

HS: Hepatic Steatosis

AI: Artificial Intelligence

ML: machine-learning

FCNN: fully-convolutional neural networks

LT: liver transplantation

DBD: donor after brain death

DCD: donor after cardiac death

SVM-SIL: support vector machines-single instance learning

LBPs: local binary patterns

Acc: accuracy

Rec: recall

Prec: precision

RGB: Red Green Blue

Keywords: Histopathology; Liver transplantation; Artificial Intelligence; Steatosis; Liver graft

Electronic word count: 2888

Number of figures and tables: 4 and 1

Declaration of interests: We declare no competing interests.

Financial support statement: No funding has been received for the conduct of this study and/or preparation of this manuscript.

Authors Contributions: CM and MS conceived the study. CM was chief investigator. CM, BR, GC, CF, DF obtained data inputs. PBC, DF, OS, DA, ML and OS were responsible for the day-to-day management of the study. PN did histopathological analyses; with supervision of PV. MS did the data and the statistical analyses. CM and MS wrote the first draft of the manuscript; all authors revised this draft. All authors read and approved the final version.

ABSTRACT

Worldwide implementation of liver-graft pool using marginal livers (such as grafts which carry a high risk of technical complications and impaired function, or a risk of transmitting infection or malignancy to the recipient) has led to a growing interest in developing methods for accurate evaluation of graft quality. Liver steatosis is associated with a higher risk of primary nonfunction (PNF), early graft dysfunction (EAD), poor graft survival rate. The present study aimed to analyze the value of artificial intelligence (AI) in the assessment of liver steatosis during

procurement compared to liver biopsy evaluation. One hundred and seventeen consecutive brain-deceased donors' liver grafts were included and classified in 2 cohorts: > vs < 30% hepatic steatosis. AI analysis required the presence of an intraoperative smartphone liver picture, as well as a graft biopsy and donor's data. Firstly, a new algorithm, arising from current visual recognition methods was developed, trained and validated to obtain automatic liver-graft segmentation from smartphone images. Secondly, a fully automated texture analysis and classification of the liver graft was performed by machine learning algorithms. Automatic liver-graft segmentation from smartphone images achieved an accuracy of 98% while the analysis of the liver graft features (cropped picture and donor's data) showed an accuracy of 89% in graft classification (> vs < 30%). This study demonstrates that AI has the potential to assess steatosis in a handy and non-invasive way to reliably identify potential non-transplantable liver grafts and avoid improper grafts utilization.

BACKGROUND

In liver transplantation, percentage of hepatic steatosis (HS) in the liver graft is associated with increased risk of graft dysfunction or early no-function [1,2] and, fast and accurate assessment of HS is of paramount importance in the setting of organ procurement for liver transplantation (LT)..

While a wide range of methods for the assessment of HS are currently available, pathological examination remains the gold standard for both HS diagnosis and grading, and is used as the reference method for any evaluation of any new HS measurement technique[3]. However, routine use of pathological examination is limited by invasiveness, or need of additional time-consuming procedures and instrumentation, often unavailable in remote graft-procurement hospitals[4]. Therefore, the decision to accept or discard a liver graft still relies on indirect parameters such as

clinical history, blood tests, liver/spleen attenuation ratio on imaging (when available), and, more importantly, on the harvesting surgeon's personal evaluation of the liver texture and color. Even-though this latter parameter achieves an accuracy of 86,2%[21] for experienced surgeons, it remains a qualitative and subjective evaluation, source of major biases in case of comparison of repeated measures.

Several recent reports have recently focused on tissue analysis using artificial intelligence (AI), including "machine-learning" (ML) and fully-convolutional neural networks (FCNN) concepts, from intraoperative optical images. These types of computerized analysis may help to obtain standardized basic pathological analysis without interobserver variability, classification biases, or technical constraints[5,6,7,8]. Based on the recent promising strategy reported by our team on HS graft analysis with AI ML technology, the present study aims at automatizing liver-graft segmentation from smartphone images and validate the robustness of this approach in order to assess HS on a larger cohort of liver grafts using different cut-off values.

MATERIALS AND METHODS

Participants

This prospective study involved two academic, high-volume LT centers and was approved by the Institutional Review Board (IRB) of "Assistance Public Hopitaux de Paris" (APHP). From January to June 2018, any liver graft from a donor after brain (DBD) or cardiac death (DCD) that was proposed for LT to one of the two participating centers, regardless the setting of the organ procurement was eligible for inclusion. The decision to accept or decline the proposal, and to subsequently carry on for graft procurement, relied on the on-call transplant surgeon.

Inclusion in the study required the presence of an intraoperative picture (taken with a smartphone from one of the harvesting team member) results of graft biopsy, clinical, biochemical and radiological donor's data.

Data collection

Important data for decision-making in graft acceptance and for LT outcome were recorded and used for ML analysis in order to improve the performance of the algorithm. Donor variables were selected based on the factors used by the surgeon to reach a decision[9] and included age, weight and height. The following biochemical variables, recorded at the time of referral, were included: γ-glutamyl transferase, alanine aminotransferase, aspartate aminotransferase and total bilirubin. Lastly, CT-scan liver/spleen attenuation ratio was also recorded.

Test methods

The test method was a HS tissue classification developed with an AI ML algorithm based on images taken during graft harvesting with a smartphone.

1.1 Artificial Intelligence Algorithms

For HS classification, ML algorithms were used. ML is a commonly used AI method to build algorithms that "learn" from data and make predictions based on these algorithms. In this study, we used the support vector machines-single instance learning (SVM-SIL) system. SVM-SIL are semi-supervised ML algorithms that use both labeled and unlabeled data to build a predictive algorithm for classification analyses. More specifically, the SVM-SIL system, from a set of training examples (i.e. with or without significant HS), using the training algorithm, builds a model that will assign new examples to one category or the other, making it a non-probabilistic binary linear classifier. This process, performed as many times as cases are included, allows assessing the sensibility, specificity and accuracy of the classification algorithm model. This method is currently the most advanced existing system in ML to avoid selection and classification bias. For liver segmentation from intraoperative photos we used fully-convolutional neural networks (FCNN), which is one of the deep-learning strategies that involves some convolutional filters, which can learn hierarchical features from data. The role of the filters consists in extracting some characteristics from the input images and collects them in maps, which include these features. The number of filters for each layer is chosen according to the time necessary for training the network and the complexity of the problem; in general, a higher number of filters will give better results. The rule is applied only up to a certain threshold because, beyond this threshold, an increase in the number of filters does not affect anymore the performance.

1.2 Smartphone images

During the procurement, at least one digital image of the liver graft was taken by the personal smartphone of a member of the harvesting surgical team, using a previously published standardized protocol to obtain the best possible image quality and reproducibility[10,11]. The same smartphone (Apple Iphone 6s) was used. The camera was automatically white-balanced and used in the Macro mode. On these smartphones, exposure and focus area selection modes were nonadjustable.

2. Pathological analysis

The existence and quantification of HS was estimated on frozen section from liver graft biopsies, and was considered the reference control test. Surgical biopsies were performed during liver procurement in case of suspected severe liver steatosis, or after reperfusion as routinely performed in both institutions. Presence and proportion of HS was assessed by an experienced pathologist (NP), supervised by a referenced pathologist (VP) and was expressed in percentage of hepatocytes with macrovesicular steatosis (0% to 100%).

3. Hepatic steatosis: test positivity cut-offs

An HS rate higher than 60% is associated with a high risk of primary graft failure[12]and is widely accepted as a cut-off value to discard liver graft discard. However, early allograft dysfunction was reported even in grafts with HS>30% generating different acceptance policies. For this reason and because 30% is the cut-off value used at our centres the machine-learning method for HS assessment was pre-determined for two categories of grafts: HS<30%, HS>30%.

3.1 Hepatic steatosis, AI cut-off

Contrarily to previous study[18] where a set of images was created by manually cropping the original photos so the target organ (liver) would occupy 100% of the frame with elimination of the background, in this study, a FCNN has been used, composed of several layers as shown in Fig. 1, inspired by U-net[13] and Resnet[14]. In our case, the model of the network consists of a convolutional (descending) and upconvolutional (ascending) paths. The combination of a convolutional block and two identity convolutional blocks is repeated four times (in Fig. 1: from stage 1 to stage 4). In each stage, the number of convolutional kernels per layer is doubled. The upconvolutional path is symmetric to the convolutional one. Each stage, repeated four times (in Fig. 1: from stage 5 to stage 8), presents an upconvolutional block instead of the convolutional one. The ascending path

ends with an upsampling block of (2, 2) size and a convolutional one, in this case with a (3 x 3) kernel and a sigmoid activation. The network was trained and validated with RGB images and with greyscale images, obtained converting the original ones. The results from automatic segmentation are compared to the ones from the manual segmentation, considered the gold standard. To evaluate the segmentation performance of FCNN we calculated the accuracy (Acc), Recall (Rec) and Precision (Prec) on MATLAB software.

Once liver graft masks were obtained by FCNN, ML was used for features classification. Since the liver texture is heterogeneous, the texture analysis was performed on liver image patches. Each image was divided in fifteen non-overlapping patches of 100 × 100 pixels. Each patch was classified in transplantable (HS < 30%) or non-transplantable (HS > 30%) and the features of each patch according to its category were extracted by rotation-invariant local binary patterns (LBPs), which are resistant to light and camera pose variations, and accurately render the liver tissue (Fig. 2). To perform the classification of patches according to donor's data, multiple instance learning (MIL) on SVM-SIL was utilized, which has the strong advantage of allowing the fusion of patch-wise information (such as textural features) with image-wise information (such as donor's data features). The feature classification was implemented with scikit-learn (http://scikit-learn.org). In order to perform robust performance evaluation, we performed leave-one-patient-out cross-validation (i.e. the images from all patients-1 were used for training and the remaining one for testing). The duration of the test classification analysis was 10^{-3} seconds. Donor's features (photos and data) classification was analyzed in term of accuracy (Acc), recall (Rec) and precision (Prec) on MATLAB software.

3.2 Hepatic steatosis, pathology

According to the most commonly used scoring system, HS was categorized in: Normal (grade 0) when the proportion of HS affected cells ranged from 0% to 5%; Mild (grade 1) between 5% and 33%; Moderate (grade 2) between 34% and 66%; Severe (grade 3) when the proportion of affected cells was greater than 67% [15].

4. Blindness

Clinical, biological and radiological information were available to the surgeons performing the organ procurement, to the pathologist performing the graft biopsy analyses and to the team in

charge of the AI assessment. The results of HS based on pathology report were available to AI assessors, but pathologists were blinded to AI HS assessment results.

Statistical analysis

Quantitative continuous variables were expressed in median and interquartile range (IQR; 25th to 75th percentile) for discrete variables, as appropriate. AI liver graft segmentation from intraoperative photo and classification of donor's features were analyzed in term of accuracy (Acc) (true positive+true negative/whole sample), recall (Rec) or sensivity (true positive/true positive+false negative) and precision (Prec) or specificity (true positive/true positive+false positive), compared to the manual photos cropping and to the index test (liver biopsy).

No sample size calculation was needed. For HS classification, a balanced dataset with 1:1 ratio groups (same numbers of liver grafts for each groups) is required for an appropriate analysis by the SVM-SIL system. Therefore, any liver graft with proven HS >30% during the study period was included in the HS>30% group. A control group was settled by randomly including liver grafts with HS <30% procured with the same technique during the same study period, on a 1:1 ratio. No other baseline variables were considered for matching.

The design of this study was based on the Essential Items for Reporting Diagnostic Accuracy Studies guidelines (STARD 2015)[16].

RESULTS

From January to June 2018, 117 consecutive liver grafts from deceased donors were photographed and biopsied with the intention to be transplanted. Of these, 28 had HS >30%, and were included in the analysis without being transplanted. Twenty-eight liver grafts procured during the same study period with HS <30% were randomly selected on a 1:1 ratio and included in the control group. This created a balanced dataset, which is required for an appropriate SVM-SIL analysis. Thus, the final inclusion cohort for HS classification comprised 56 liver grafts.

A liver biopsy was performed without any complication in all 117 cases. Forty Red Green Blue (RGB) liver graft images (size 3264× 2448 pixels, 8 Megapixels), taken during each graft procurement, were analyzed. For liver segmentation analysis (liver graft image extraction), all 117

intraoperative pictures were used: 50 for the training dataset and 67 for the testing dataset. The Prec medians for grey-scale and RGB images are equal to 95% and 97%, respectively; The Acc and Rec medians for grey-scale and RGB images were 92%, 89% and 98%, 97%, respectively (**Fig.3**).

Texture analysis by SVM-SIL was performed on liver image patches, and a balanced dataset of 600 patches was obtained. Flow diagram of participants is summarized in **Figure 4**, and baseline characteristics are summarized in **Table 1**.After analysis of the features of each dataset, the SVM-SIL showed good results in graft classification according to their acceptance or discard for LT, with a recall (Rec) of 93% and a precision (Prec) of 82% for transplanted grafts and with a Rec of 97% and Prec of 83% for discarded grafts.

The classification of donor data from the liver picture according to the proportion of steatosis (superior or inferior to 30% of HS) showed an accuracy of 89%.

Discussion

This is the first multicenter prospective study to assess the performance of AI technologies to evaluate HS through smartphone pictures and donor data with comparison to liver biopsy. AI approach reported a sensibility (Rec) of 97% and 93% for the classification of non-steatotic and steatotic grafts based on acut off value of30% with an accuracy of 89%. These results validate our primary monocentric study where analysis of 40 liver grafts (two classes of 20 grafts) showed a sensibility (Rec) of 80% and 95% for the classification of non-steatotic, mild and moderate steatotic vs. severe steatotic grafts (cut off of HS=60%) with an accuracy of 88%[18]. Furthermore, the use of FCNN for liver graft image extraction from donor's picture, firstly reported in literature, leads to a fully automated HS assessment method. Testing this AI approach with the most frequently used HS cut-off (30% and 60%) for liver grafts acceptance policy supports that this method represents a promising step towards a clinically relevant processing system for automatic, non-invasive and objective HS assessment in the setting of LT.

The main strengths of this study are the non-invasiveness, rapidity, utilization of worldwide available device (i.e. smartphone) for the analysis as compared to the standard reference exam, the liver biopsy. Liver photos and donor's data are processed and classified within seconds with the computer-assisted semi-supervised system. No additional procedure or equipment is needed during graft harvesting, and no invasive liver sampling is necessary. On the contrary, a liver biopsy requires a surgical graft sampling, which may cause complications such as bleeding and involves an examination from a pathologist, which is 1) time consuming and 2) frequently unavailable 24/7 in remote places. In contrast, this AI-based steatosis assessment technique is a real time, non-invasive method to assess hepatic steatosis.

The pictures were acquired with an easy to use and widely available smartphone camera. Although smartphones have a significant cost, they are prevalent in the general population (100% of our surgical team). Moreover, the smartphone camera used in this study (Apple iPhone®) has a medium performance when compared to other smartphones[17] so it would be reasonable to assume that more recent smartphones could also be used for graft picture acquisition. The second advantage of this method is the standardization of HS evaluation. Even though liver biopsy is considered as the standard exam for HS assessment, this technique is subject to inter-observer variability or sampling size limitation, as underlined by previous publications[3]. Furthermore, when pathological examination is not available, the decision to accept or discard a liver graft is usually based on donor's data, liver texture, macroscopic evaluation and subjective clinical experience of the harvesting surgeon, making the graft evaluation highly subjective to bias. By using an AI-based classification, interobserver variability and classification bias are dramatically limited. Indeed, the use of AI allows converting a subjective decision-making process into a standardized, computerized and homogeneous process.

In this experimental study, the highest classification performance was obtained using texture features combined with significant donor's data. Indeed, the inclusion of donor features in the algorithm helped increasing the accuracy of the classification by SVM-SIL, as previously demonstrated[18]. The real innovation of this study, through AI computerization of a human process, is threefold: first, to translate the subjective visual assessment of liver texture into an objective and standardized method (smartphone picture),;second, to develop a fully automated HS assessment by the liver-graft image extraction; and third, to replicate the clinical experience of a transplant surgeon (donor's data analysis). A major limitation of this study is the small sample

size, which is a common problem within the computer-assisted diagnosis community[19] and particularly in LT setting where larger cohort of discarded grafts (without any intention to be transplanted) are available only for machine perfusion studies[20]. On the contrary, our sample of discarded grafts were accepted with the intention to be transplanted but then discarded for HS>30% (measured by intraoperative biopsy). A second limit is the impossibility of SVM-SIL analysis to differentiate macro- and microsteatosis by liver picture and biology when only macrosteatosis is predictive of liver graft dysfunction. Therefore, the present results will require confirmation in larger multicentric studies probably in the machine perfusion study setting. Enlarging the training dataset would also allow investigating more advanced machine learning methods. Nevertheless, it has been reported that SVM-SIL achieves competitive results as compared with other more sophisticated semi-supervised methods[6]. The third limit is that the HS assessment method described here, as all methods reported in literature[21], could not be a substitute to liver biopsy, as no methods are set to predict other liver conditions (such as balloon degeneration, centrolobular necrosis, or chronic hepatitis), which (more rarely) can also adversely affect graft outcome. Lastly the use of 30% steatosis as a cut-off value could not reflect the worldwide liver graft acceptance policy and the old smartphone technology (the Iphone 6s was firstly officially released on September 25, 2015) could limit algorithm performance.

As the LT scientific community claims for new standards on HS assessment, researches to find alternatives to histopathology are highly encouraged. Interesting results are available on MRI and elastography [22, 23], but none of them is either practical or performing in the very specific setting of organ procurement.

This research showed that liver texture analysis from pictures and donor's data features, analyzed by AI approach, could represent a promising step towards a helpful processing system to support the surgeon's decision – particularly the younger surgeons of the future- to accept or discard a liver graft during procurement.

Acknowledgments

We thank Evelyne Monmignot, Anne-Gaëlle Ceres, Anne Buisine and Smices® for providing the technical support for the study.

References

1. Canelo R, Braun F, Sattler B, Klinge B, Lorf T, Ramadori G, Ringe B. Is a fatty liver dangerous for transplantation? Transplant Proc. 1999 Feb-Mar;31(1-2):414-5
2. de Graaf EL, Kench J, Dilworth P, Shackel NA, Strasser SI, Joseph D, Pleass H, Crawford M, McCaughan GW, Verran DJ. Grade of deceased donor liver macrovesicular steatosis impacts graft and recipient outcomes more than the Donor Risk Index. J Gastroenterol Hepatol. 2012 Mar;27(3):540-6. doi:10.1111/j.1440-1746.2011.06844.x. PubMed PMID: 21777274.
3. El-Badry AM, Breitenstein S, Jochum W, Washington K, Paradis V, Rubbia-Brandt L, Puhan MA, Slankamenac K, Graf R, Clavien PA. Assessment of hepatic steatosis by expert pathologists: the end of a gold standard. Ann Surg. 2009 Nov;250(5):691-7. doi: 10.1097/SLA.0b013e3181bcd6dd. PubMed PMID: 19806055.
4. Caussy C, Alquiraish MH, Nguyen P, Hernandez C, Cepin S, Fortney LE, Ajmera V, Bettencourt R, Collier S, Hooker J, Sy E, Rizo E, Richards L, Sirlin CB, Loomba R. Optimal threshold of controlled attenuation parameter with MRI-PDFF as the gold

standard for the detection of hepatic steatosis. Hepatology. 2018 Apr;67(4):1348-1359. doi: 10.1002/hep.29639. Epub 2018 Feb 19. PubMed PMID: 29108123; PubMed Central PMCID: PMC5867216.

5. Li B., Meng M. Q.-H. Texture analysis for ulcer detection in capsule endoscopy images. Image and Vision Computing. 2009;27(9):1336–1342. doi: 10.1016/j.imavis.2008.12.003.
6. Quellec G, Cazuguel G, Cochener B, Lamard M. Multiple-Instance Learning for Medical Image and Video Analysis. IEEE Rev Biomed Eng. 2017;10:213-234. doi:10.1109/RBME.2017.2651164.
7. Moccia S, De Momi E, Guarnaschelli M, Savazzi M, Laborai A, Guastini L, Peretti G, Mattos LS. Confident texture-based laryngeal tissue classification for early stage diagnosis support. J Med Imaging (Bellingham). 2017 Jul;4(3):034502. doi: 10.1117/1.JMI.4.3.034502.
8. de Cunha DA, Eadie LH, Barbur JL, Hawkes DJ, Seifalian AM. The effect of image colour distortion on evaluation of donor liver suitability for transplantation. Comput Biol Med. 2004 Oct;34(7):615-32. PubMed PMID: 15369712.
9. Neuberger J. Transplantation: Assessment of liver allograft steatosis. Nat Rev Gastroenterol Hepatol. 2013 Jun;10(6):328-9. doi: 10.1038/nrgastro.2013.74
10. Cesaretti M, Poté N, Cauchy F, Dondero F, Dokmak S, Sepulveda A, Schneck AS, Francoz C, Durand F, Paradis V, Soubrane O. Noninvasive assessment of liver steatosis in deceased donors: A pilot study. Liver Transpl. 2018 Apr;24(4):551-556. doi: 10.1002/lt.25002.
11. Gaujoux S, Ceribelli C, Goudard G, Khayat A, Leconte M, Massault PP, Balagué J, Dousset B. Best practices to optimize intraoperative photography. J Surg Res. 2016 Apr;201(2):402-7. doi: 10.1016/j.jss.2015.11.048. Epub 2015 Nov 30. PubMed PMID: 27020825.
12. Le Naour F, Gadea L, Danulot M, Yousef I, Vibert E, Wavelet M, Kaščáková S, Castaing D, Samuel D, Dumas P, Guettier C. Quantitative assessment of liver steatosis on tissue section using infrared spectroscopy. Gastroenterology. 2015 Feb;148(2):295-7. doi: 10.1053/j.gastro.2014.11.038.
13. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical

Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham

14. He K, Zhang X, Ren S, SunJ. Deep residual learning for image recognition, IEEE Conference on Computer Vision and Pattern Recognition. 2016:770–778. arXiv:1512.03385

15. Bedossa P; FLIP Pathology Consortium. Utility and appropriateness of the fatty liver inhibition of progression (FLIP) algorithm and steatosis, activity, and fibrosis (SAF) score in the evaluation of biopsies of nonalcoholic fatty liver disease. Hepatology. 2014 Aug;60(2):565-75. doi: 10.1002/hep.27173.

16. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar DA, Cohen JF; STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015 Oct 28;351:h5527.

17. Morrison AO, Gardner JM. Microscopic Image Photography Techniques of the Past, Present, and Future. Arch Pathol Lab Med. 2015 Dec;139(12):1558-64. doi:10.5858/arpa.2014-0315-RA. Epub 2015 May 19. Review. Erratum in: Arch Pathol Lab Med. 2016 Jul;140(7):618. PubMed PMID: 25989285.

18. Moccia S, Mattos LS, Patrini I, Ruperti M, Poté N, Dondero F, Cauchy F, Sepulveda A, Soubrane O, De Momi E, Diaspro A, Cesaretti M. Computer-assisted liver graft steatosis assessment via learning-based texture analysis. Int J Comput Assist Radiol Surg. 2018 Sep;13(9):1357-1367. doi:10.1007/s11548-018-1787-6

19. Maier-Hein L, Vedula SS, Speidel S, Navab N, Kikinis R, Park A et al. Surgical data science for next-generation interventions. Nature Biomedical Engineering. 2017 Sep 1;1(9):691-696. DOI: 10.1038/s41551-017-0132-7

20. Vogel T, Brockmann JG, Quaglia A, Morovat A, Jassem W, Heaton ND, Coussios CC, Friend PJ. The 24-hour normothermic machine perfusion of discarded human liver grafts. Liver Transpl. 2017 Feb;23(2):207-220. doi: 10.1002/lt.24672.

21. Cesaretti M, Addeo P, Schiavo L, Anty R, Iannelli A. Assessment of liver graft steatosis: where do we stand? Liver Transpl. 2018 Oct 31. doi: 10.1002/lt.25379.

22. Raptis DA, Fischer MA, Graf R, Nanz D, Weber A, Moritz W, Tian Y, Oberkofler CE, Clavien PA. MRI: the new reference standard in quantifying hepatic steatosis? Gut. 2012 Jan;61(1):117-27. doi: 10.1136/gutjnl-2011-300155.

23. Hong YM, Yoon KT, Cho M, Chu CW, Rhu JH, Yang KH, Lee JW. Clinical usefulness of controlled attenuation parameter to screen hepatic steatosis for potential donor of living donor liver transplant. Eur J Gastroenterol Hepatol. 2017 Jul;29(7):805-810. doi: 10.1097/MEG.0000000000000876

FIG.1 FCNN work-flow: Left - The whole path can be divided in two parts: the descending path and the ascending one, each consisting of four stages. Each stage of the descending path is made of a convolutional block (red boxes) and two identity blocks (blue boxes), whereas in the

ascending path there are an upconvolutional block (green boxes) and two identity blocks. The ”ZeroPadding” block represents a zero padding layer (P, P), with padding P x P; the ”Convolut.” block represents a convolutional layer (C, N, S), with channels C, kernel size N x N, and stride S. Each convolutional layer is followed by a batch normalization layer and a ReLU activation function. The ”Max Pooling” block denotes a max pooling operation (N, S) over N x N patches and with stride S; the ”UpSampling” denotes an upsampling operation (K x K) with size K. The blue dashed arrows indicate the concatenation of the feature map from the descending path to the ascending one. Right - Example of convolutional, identity and upconvolutional blocks.

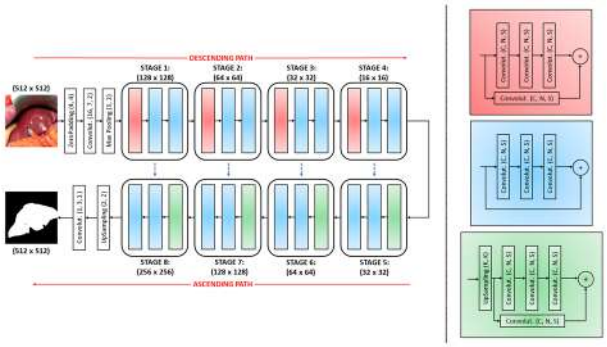
FIG.2 ML work-flow: Image patches are extracted from RGB liver photos acquired during liver procurement. From each patch, textural and intensity features are extracted and added to clinical, biological and radiological donor’s features. The features extracted from training patches are used to train a semi-supervised support vector machine (SVM)-single instance learning (SIL) model (green boxes). The approach is semi-supervised as the ground-truth label is assigned to the whole image and not to the single patch.

FIG. 3 Sample of segmentation outcomes. The orange, light-blue and green lines refer to the manual mask, mask from grey-scale original image and mask from RGB original image, respectively

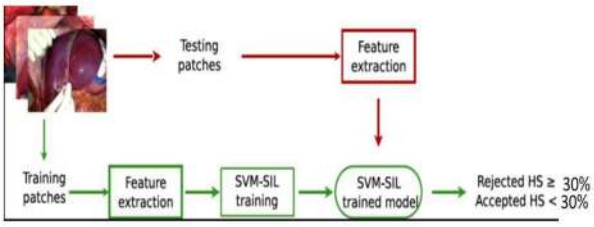
FIG.4 Flow diagram for Machine Learning data analysis

Table I. Donor's data

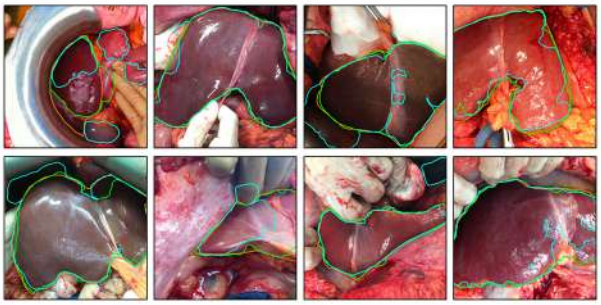
	Transplanted grafts (n=28)	**Not transplanted grafts (n=28)**
Preoperative		
Age	52 [18 ; 88]	62 [17 ; 86]
Sex Ratio M/F	16/12	17/11
BMI (kg/m^2)	23.6 [17 ; 31]	30 [16;40]
Height (m)	170 [150 ; 190]	170 [150 ; 190]
Weight (kg)	69 [52 ; 90]	90 [44 ; 178]
ICU (days)	3 [1 ; 10]	3 [1 ; 11]
AST (UI/l)	57 [17 ; 177]	193 [24 ; 2000)
ALT (UI/l)	37 [11 ; 136]	196 [14 ; 2000]
gGT (UI/l)	69 [15 ; 318]	95 [16 ; 569]
Bilirubine (µmol/l)	11 [3 ; 23]	17 [3.6; 57]
Lactate (m.mol/l)	2 [0.5 ; 6.2]	3.4 [0.8 ; 14]
L/S density (HU)	16 [2 ; 70]	17 [5 ; 79]
Macrovacuolar steatosis (%, mean)	15 [5 ; 30]	40 [30 ; 90]
Microvacuolar steatosis (% mean)	2 [0 ; 5]	50 [20 ; 80]
Fibrosis (stage)	0	0



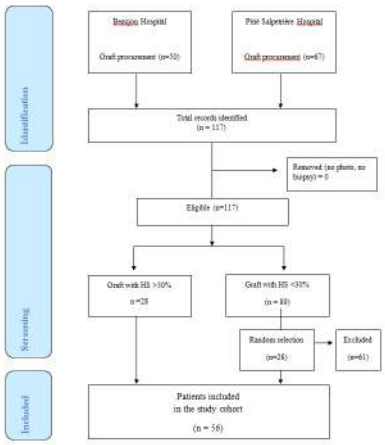
lt_25801_f1.tif



lt_25801_f2.tif



lt_25801_f3.tif



lt_25801_f4.tif