Should we use liver grafts repeatedly refused by other transplant teams?

Audrey Winter, Paul Landais, Daniel Azoulay, Mara Disabato, Philippe Compagnon, Corinne Antoine, Christian Jacquelinet, Jean-Pierre Daurès, Cyrille Féray

Table of contents

Supplementary methods

Comparison between PA and CA populations.

A comparison of CA and PA recipients was performed using the χ^2 test, Fisher's exact test and a Student test or Wilcoxon-Mann-Whitney test, when appropriate. For each test, an effect-size (ES) was also reported. For χ^2 tests and Fisher tests, Cramer's V (ϕ_C) was calculated (i.e. magnitude of the ES; small: $0.1 \leq ES < 0.3$, medium $0.3 \leq ES < 0.5$ and large $ES > 0.5$), and for the Student and Wilcoxon-Mann-Whitney tests r^2 and Cohen's r^2 (i.e. magnitude of the ES; small: $0.01 \leq ES < 0.09$, medium $0.09 \leq ES < 0.25$ and large $ES > 0.25$), respectively (Cohen, 1988; Fritz *et al.*, 2012; Tomczak and Tomczak, 2014). The larger the ES, the greater was the impact of the findings, all other things being equal.

Propensity Score and weighted Cox model

To deal with any potential selection bias, we used a propensity score approach and a weighted Cox model using the Inverse Probability of Treatment Weighting method to compare recipients and grafts survival between PA and CA

Propensity Score

The propensity score (PS) defining the type of allocation was determined using logistic regression, according to the method described by Rosenbaum and Rubin (Rosenbaum and Rubin, 1983). This PS included covariates to explain the type of allocation, so that the two recipient groups would be comparable.

The following recipient covariates were used: age, gender, MELD score at transplant and the presence of HCC. The PS was the probability of a patient being grafted with a CA graft given these covariates. Thus:

$$PS_i = \Pr(Z = 1 | X = x_i),$$

where $Z = 1$, CA type and $X = x_i$ explanatory variables for patient i.

We then validated this PS in two steps in order to assess its ability to balance the groups (with and without a CA transplant). Patients presenting with the same PS were matched (without any replacements):

$$Pr(Z = 1|X = x_i) = Pr(Z = 0|X = x_i).$$

This led to the creation of a dataset of 672 patients (including 336 CA and 336 PA recipients). Student's test/Wilcoxon-Mann-Whitney test and the χ^2 test were then performed on the explanatory variables. This two-step procedure was repeated 200 times. If the tests were non-significant, whatever the correction used (i.e. Bonferroni or a less conservative correction), then the covariates were considered to be well balanced after adjustment to the PS.

Weighted Cox model

This PS was used to implement the Inverse Probability of Treatment Weighting (IPTW) method. A CA transplant was considered as the variable of interest. All patients were included in the model. For CA recipients we therefore noted:

$$W_{IPTW-CA} = \frac{1}{PS},$$

and for PA recipients:

$$W_{IPTW-PA} = \frac{1}{1-PS}.$$

The adjustment covariates for recipients that were chosen for the weighted Cox model were: re-transplantation, MELD exceptions, status at transplant (hospital, ICU, home), diabetes, on dialysis at LT, decompensated cirrhosis, non-cirrhotic liver disease, body mass index, encephalopathy, ascites, waiting time and ABO compatibility. The adjustment covariates for donors were height and DQI (Winter *et al.*, 2018). These covariates were selected *a priori*. An inter-region effect was also tested and if significant, an inter-region stratification was applied. The proportional hazard assumption was checked using Schoenfeld residuals.

Survival benefit: Sequential Stratification method

The survival benefit associated with a CA graft was estimated using sequential stratification derived from the method described in Schaubel et al. (Schaubel *et al.*, 2006, 2008; Schaubel and Kalbfleisch, 2014). This method essentially reorganises the observed data, and as close as possible reproduces the conditions of a randomised controlled trial. The aim was to determine in a given patient, whether it was better to be grafted without delay with a CA graft rather than remaining on the WL and possibly later receiving a PA graft (Winter *et al.*, 2020).

According to Schaubel et al (Schaubel *et al.*, 2006), for each liver transplantation (LT), a stratum was created, which included the transplanted patient (index patient) and all the matched "control" patients, who:

- were active on the waiting list (WL) (not grafted, still alive, not removed from the WL, not under temporary contraindication, not lost to follow-up),
- had more or less than 2 Model for End-stage Liver Disease (MELD) points apart from the index patient (the MELD is calculated at the time of the transplantation of reference for the index patient and the control patients),
- had more or less than 5 years old apart from the index patient,
- had the same sex as the index patient,
- had the same hepatocellular carcinoma status as the index patient,
- and were ABO compatible with the graft received by the index patient.

“Control" patients were censored only if they received a patient allocation (PA) graft. Indeed, if they received a PA graft then they can no longer be considered as part of the control group. Then they were censored at their own transplant date.

According to randomized controlled trial data, all patients have "the same t_0". In a stratum, follow-up began at the index patient’s LT. We then computed, for each patient included in a stratum, the time spent on the WL and the time spent on the WL before the date of the reference LT (these durations are identical for the index patient). Finally, we created a covariate “centre-matching” coded 0 if the index patient and the matched patients belonged to the same centre; and 1 otherwise. These covariates have been used as adjustment covariates in the final model.

Strata were then combined, and a stratified Cox regression model was fitted in order to estimate the hazard ratio (HR) specific to the allocation type group, adjusted for covariates. The HR allows to compare patients grafted with a centre allocation (CA) graft with the group of patients remaining on the WL, waiting for a potential PA graft. For the ℓ^{th} strata, the hazard function was:

$$\lambda_{i(\ell)}(t;\beta) = \lambda_{0(\ell)}(t)e^{\beta_A(Allocation\ type)_{i(\ell)} + \boldsymbol{\beta_B^T}\boldsymbol{Z}_{i(\ell)}},$$

where β_A is the parameter of interest to estimate, corresponding to the $(Allocation\ type)_{i(\ell)}$ covariate which, takes the value 1 for CA recipient, and takes the value 0 otherwise (control patients matched: patients remained on the WL waiting for a potential PA graft).

The covariate vector $\boldsymbol{Z}_{i(\ell)}$ included recipient's covariates: re-transplantation, MELD exception, status at transplant (hospital, intensive care unit, home), diabetes, on dialysis at listing, decompensated cirrhosis, non-cirrhotic liver disease, BMI, encephalopathy, ascites, centre-matching covariate, time on WL and time on WL before matching. We assumed that the censorship was conditionally independent of the outcome knowing $\boldsymbol{Z_i}(t)$ and the strata s_i.

The consistency of β_A and the constancy of the HRs over time were checked according to Schaubel et al (Schaubel *et al.*, 2006).

Table S1 – Proportion of non-transplanted grafts in the 2009-2014 period and related causes.

Causes: / **Year:**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**
Macroscopic steatosis (with or without biopsy)	38	41	40	35	28	35
Poor graft quality	19	23	32	25	17	20
Tumor (or suspicion)	2	5	9	10	9	12
Recipient cause	1	1	0	3	1	2
Other Causes*	9	7	4	5	11	4
Total	**69**	**77**	**85**	**78**	**66**	**73**
Proportion among collected grafts (%)	**6.3**	**6.7**	**7**	**6.3**	**5.1**	**5.4**

* logistic impediments, damaged graft, technical difficulties of surgery, anatomical difficulties, or cardiac arrest during organ removal)

References:

Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, N.J: L. Erlbaum Associates; 1988

Fritz CO, Morris PE, Richler JJ. Effect size estimates: Current use, calculations, and interpretation. Journal of Experimental Psychology: General 2012; 141: 2–18.

Rosenbaum PR, Rubin DB. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika 1983; 70: 41–55.

Schaubel D, Sima C, Goodrich N, Feng S, Merion R. The survival benefit of deceased donor liver transplantation as a function of candidate disease severity and donor quality. American journal of transplantation 2008; 8: 419–425.

Schaubel DE, Kalbfleisch JD. Assessing the Effect on Survival of Kidney Transplantation with Higher-Risk Donor Kidneys. In: Statistics in Action: A Canadian Outlook. Jerald F.Lawless; 2014. p. 209–224

Schaubel DE, Wolfe RA, Port FK. A Sequential Stratification Method for Estimating the Effect of a Time-Dependent Experimental Treatment in Observational Studies. Biometrics 2006; 62: 910–917.

Tomczak M, Tomczak E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. 2014; 1: 19–25.

Winter A, Feray C, Antoine C, Azoulay D, Daurès J-P, Landais P. Matching Graft Quality to Recipient's Disease Severity Based on the Survival Benefit in Liver Transplantation. Scientific Reports (in press) 2020

Winter A, Féray C, Audureau E, Azoulay D, Antoine C, Daurès J-P, et al. A Donor Quality Index for liver transplantation: development, internal and external validation. Scientific Reports 2018; 8: 9871–9884.