University of Louisville

ThinkIR: The University of Louisville's Institutional Repository

Electronic Theses and Dissertations

5-2012

Propensity score matching for surgical outcomes with observational data.

Robert M. Cannon 1983-
University of Louisville

Follow this and additional works at: https://ir.library.louisville.edu/etd

Recommended Citation

PROPENSITY SCORE MATCHING FOR SURGICAL OUTCOMES WITH OBSERVATIONAL DATA

By

Robert M Cannon

MD, Medical College of Georgia, 2008

A Thesis

Submitted to the Faculty of the

School of Public Health and Information Sciences of the University of Louisville

in Partial Fulfillment of the Requirements

for the Degree of

Master of Sciences

Department of Biostatistics and Informatics

University of Louisville

Louisville, Kentucky

May 2012

PROPENSITY SCORE MATCHING FOR SURGICAL OUTCOMES WITH OBSERVATIONAL DATA

By Robert M Cannon

MD, Medical College of Georgia, 2008

A Thesis Approved On

3/26/2012

by the following Thesis Committee

Guy Brock, Ph.D   .

Thesis Director

John Myers, Ph.D.

_____

_____

TABLE OF CONTENTS

ABSTRACT

PROPENSITY SCORE MATCHING FOR SURGICAL OUTCOMES WITH OBSERVATIONAL DATA

Robert M. Cannon, 3/26/2012

Because of limitations in randomized controlled trials, medical researchers are often forced to rely upon studies of observational data. Confounding is a major difficulty encountered in such studies that can create considerable bias in estimates of treatment effects. Propensity score analysis was developed by Rosenbaum & Rubin in 1983 to overcome these difficulties. In essence, a propensity score allows balance to be achieved on confounding covariates in treatment and control groups, thus creating a 'quasi-randomized' trial from observational data. In this study, I illustrate the use of propensity matching to demonstrate that African American race is a significant risk factor for receiving a lower quality donor kidney using a national database on transplantation. I then use propensity matching to demonstrate the benefits of laparoscopic resection for hepatic colorectal metastases. In doing so, the great value of propensity matching in reducing bias in observational studies is demonstrated.

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION TO PROPENSITY SCORE METHODS

The goal of much of medical research is to evaluate the effect of a particular intervention or exposure (which, hereafter I will refer to as a treatment) on a given outcome of interest. For example one may wish to determine whether a new chemotherapy regimen has a superior response rate compared to the current standard. In the hierarchy of medical evidence, the randomized controlled trial (RCT) is considered by many to be the gold standard for the assessment of such questions.

There are a number of important reasons for the value given to the RCT in studies of causal inference. By the nature of randomization, the potential of bias in the allocation of patients to either the treatment or control group is eliminated. As a corollary to this property, a properly executed randomization scheme tends to produce treatment and control groups that are balanced (that is, they have similar distributions) on relevant covariates, both those that are measured *and* those that are unmeasured[1]. As I will demonstrate, these characteristics allow for unbiased estimation of the effect of the particular treatment under study.

Despite these considerable advantages, randomized controlled trials are subject to a number of limitations. In the first place, the costs associated with designing and carrying out an RCT can be expensive, particularly when a large number of subjects is required to achieve the desired power. Randomized controlled trials frequently contain stringent inclusion and exclusion criteria speciying whom is allowed to participate. Thus, the trial population may differ from the general population to a degree that the findings of the trial may not be broadly applicable, which is to stay that the study lacks external validity[2]. Finally, there are a number of circumstances where a randomized trial may be impractical, unethical, or even impossible. For example, if a researcher is interested in studying the effect of race on survival after diagnosis of a particular cancer, it would obviously be impossible to randomize the race of the subjects. If one wanted to study the effect of smoking on mortality, it would clearly be unethical to randomize subjects to smoking or non-smoking.

Because of the limitations of RCTs, researchers are often forced to rely upon studies that are observational in nature. In an observational study, the researcher has no control over the treatment assignment for the subjects, instead they are "self-selected" into the treatment and control groups[2]. Subjects in the treatment and control groups thus may differ widely in terms of baseline covariates. A covariate that is correlated with both the treatment and the outcome is known as a confounder. If subjects in the treatment and control groups differ significantly on confounding variables, then subsequent estimates of the treatment effect may be subject to considerable bias if steps are not taken to control for the confounding factors.

Rubin gives an example of the effect confounding can have on estimates of treatment effect, which I will summarize here[3]. Consider the case of a researcher wishing to study the effects of smoking on mortality. He analyzes a cohort that consists

of frequent smokers and nonsmokers. Consider now that the non-smokers are, on average, older than the smokers. If the researcher were to simply compare the mortality rates of the smokers to those of the nonsmokers, he may very well find that the mortality rate is higher in the nonsmokers. This observation does not account for the confounding effect of age, in that older subjects are likely to have higher mortality rates regardless of smoking status. If the researcher were to then re-analyze the data in a manner that adjusts for the age difference between groups, he would then correctly observe that smoking is indeed a risk factor for increased mortality.

If the number of confounding variables were relatively few, as in the example above, it would be fairly straightforward to control for their effect and thus reduce the bias observed in estimation of the treatment effect. One of the most common methods is stratification, also referred to as subclassification by Cochran[4]. In brief, stratification involves dividing the treatment and control groups into categories, or strata, based on the levels of the confounding variable. In the example given by Rubin above, one may divide the smokers and nonsmokers into categories based on age, such as younger, middle-aged, and older[3]. The outcome variable can then be assessed within each strata, and the results combined to get the overall estimated treatment effect. By performing comparisons using 5 to 6 strata, Cochran has demonstrated that 90% or more of the bias present in unadjusted analysis may be removed[4].

Frequently, however, there are many potential confounders for which investigators need to control. As the number of covariates increases, Cochran has demonstrated that the number of strata grows in an exponential fasion.[5] For example, a study with k dichotomous covariates would have 2^k strata. Propensity score methods, initially developed by Rosenbaum and Rubin[6] in 1983, provide a convenient means for addressing the problem of confounding by multiple covariates in observational studies

3

and have become increasingly popular over the past two decades in the medical literature.

Before specifying the derivation and use of the propensity score, it will first be helpful to consider the conceptual framework upon which it operates, which is known as the potential outcomes framework, or Rubin Causal Model[7, 8]. In this model, there are two treatments (labeled 0 and 1) and an outcome of interest. For a sample of N subjects, the ith subject has two potential outcomes, which are $Y_i(0)$ and $Y_i(1)$, which are the outcomes when the ith subject receives treatment 0 and when the ith subject receives treatment 1, respectively. In reality, each subject only receives one treatment, either the active or the control.

Now let Z be an indicator variable for the treatment assignment, where $Z=0$ for the control treatment and $Z=1$ for the active treatment. The outcome observed for each subject under the treatment received thus becomes:

$$Y_i = Z_i Y_i(1) + (1 - Z_i)Y_i(0).$$

The average treatment effect (ATE) thus becomes the difference between the expectations between $Y_i(1)$ and $Y_i(0)$, that is to say:

$$ATE = Y_i(1) - Y_i(0)$$

The average treatment effect is thus the effect of moving the entire population from the untreated to the treated condition[7]. In other words, the ATE is an estimate of the difference in outcome if those who received one treatment instead received the other. In the RCT setting, randomization allows for direct comparison of the treatment and control groups. This can't be done in observational studies, however, because the subjects in one treatment group differ systematically from those in the other treatment group. In

order for direct comparisons to be meaningful in such a setting, a balancing score is needed to even out the differences in covariates between the two treatment groups[6].

A balancing score, *b(x)* (of which the propensity score is a single example), is a function of the observed covariates *x* such that the conditional distribution of *x* given *b(x)* is the same for both the treated *(z=1)* and control *(z=0)* groups. That is to say, *x* and *z* are independent conditional on *b(x)*. There are many balancing scores, the finest of which is the vector of covariates itself, x^6.

Having defined a balancing score, now we consider the distribution of the treatment assignments, which is as follows:

$$e(x) = pr(z = 1|x)$$

and it is also assumed that

$$pr(z_1, \ldots, z_n | x_1, \ldots, x_n) = \prod_{i=1}^{n} e\,(x_i)^{z_i}\{1 - e(x_i)\}^{1-z_i}$$

In the above equation, *n* is the number of subjects, while z_i is the indicator for treatment assignment of patient *i*, and x_i is the vector of covariates for the *i*th subject.

The probability *e(x)* is known as the propensity score, and represents the conditional probability of assignment to the treatment group, given the observed covariates. If two subjects have the same propensity score, then the probability that either of them would be assigned to the treatment group, conditional on the observed covariates, is also equal. Thus it is as if a coin flip was performed to determine to which treatment group the subject would actually be assigned. Herein lies the primary utility of propensity score methods, in that they create what has been termed a "quasi-randomized" experiment from observational data[9].

An important assumption of the propensity score is that of strongly ignorable treatment assignment[6], which states that treatment assignment and outcome are conditionally independent given the vector of covariates, and that for each subject, there is a possibility of assignment to either treatment group. In mathematical notation:

$$Y_i(0), Y_i(1) \perp z|x, and\ 0 < pr(z = 1|x) < 1$$

The first condition of strongly ignorable treatment assignment is also known as the "no unmeasured confounders assumption"[7]. Stated yet another way, the assumption is that all variables that affect treatment assignment have been included in the propensity score. The second condition of strongly ignorable treatment assignment $(0 < pr(z=1|x) < 1)$, may be stated alternatively that there must be sufficient overlap in estimated propensity score between the treated and control groups. The region of overlap is known as the zone of common support[2]. At a given value of the propensity score, the difference between the treatment and control means is an unbiased estimate of the average treatment effect when treatment assignment is strongly ignorable.

In a randomized controlled trial, the propensity score is known, as the assignment into treatment groups is controlled by the investigators. Remember that the propensity score is simply the probability that a particular subject will be assigned to a treatment group. In an RCT, treatment assignment is solely a function of the randomization scheme, and thus propensity is known and subject to control by the investigators. In most cases, the propensity score for each subject in an RCT would thus be 0.5. In observational studies, on the other hand, the propensity score is unknown and must be estimated using the measured covariates. Although several methods for estimating the propensity score have been suggested[10-12], by far the most commonly used method is that of logistic regression[6,7].

In the logistic regression framework, the propensity score (p) is estimated as follows:

$$\log \frac{p}{1-p} = \beta' x$$

where β' is the vector of regression coefficients and x is the vector of measured covariates. The major question in deriving the estimated propensity score then becomes which covariates to include in the logistic regression model. Possible sets of covariates include all measured baseline variables, all baseline covariates that affect the treatment assignment, all covariates that affect the outcomes (potential confounders), and all covariates that affect both the treatment assignment and the outcome (true confounders)[7]. Although there is no universal agreement on which set is the best to choose, there have been some studies to suggest that inclusion of only the potential or true confounders leads to more precise estimation of the treatment effect[13, 14]. Which variables meet these requirements will often require subject matter expertise and/or reference to the available literature. Austin has noted, however, that most baseline characteristics are likely to affect both treatment assignment and outcome, so inclusion of all measured baseline characteristics is generally safe[7]. There has been discussion in the literature of using goodness of fit estimates such as the c-statistic to determine the correct specification of the propensity score model, however such methods have been shown to be ineffective in reducing confounding, and may actually result in a decreased zone of common support between the treated and control groups[15]. Thus, selection of covariates to include in the propensity score model should be guided as stated above by expert knowledge, reference to the literature, and empirical observation[15, 16].

After the propensity score has been estimated, there are three common methods for its use: matching on the propensity score, stratification on the propensity score, and

regression on the outcome of interest with the propensity score included as a covariate[2, 6, 7, 9]. In general, matching has been demonstrated to result in greater balance between the treated and untreated groups than the techniques of stratification and regression[7, 13, 17]. As matching (generally in a 1:1 fashion) is the most commonly employed method, and is the method I will feature in the following chapters, the discussion here will be limited to matching methodologies.

Having decided to match on propensity scores, the investigator is faced with a number of choices as to how matching should actually be performed. The first choice is whether to match with or without replacement[2, 7]. A potential drawback of matching with replacement is that a single untreated subject may be matched with multiple treated subjects. Another decision is whether to match using a simple nearest neighbor algorithm (in which treated subjects are matched to untreated control with the closest propensity score), or whether to impose a caliper on the maximum difference in propensity score between the treated subject and untreated control for the match to be considered valid.

If the zone of common support is limited, then simple nearest neighbor matching will tend to result in poor matches being made for the treated subjects at the higher end of the estimated propensity score distribution. In order to prevent this problem, imposition of a caliper that specifies the maximum difference in propensity scores has been proposed[7, 18-20]. There is no generally accepted caliper width, so the ultimate judgment must be based on the degree of balance achieved between the matched samples, as will be discussed later. Having specified an appropriate caliper, the investigator must then choose between "greedy" matching and "optimal" matching.

In a greedy matching scheme, once matches are made they are not broken[7]. Thus, once a control is matched to a treated subject, that control is removed from the pool of potential controls for others in the treated group. An optimal matching strategy, on the other hand, is formulated so that matches are made which minimize the total distance in propensity scores between pairs. Comparison between the two schemes have demonstrated that optimal matching performs no better than greedy matching in producing balanced samples, which is the ultimate goal of propensity matching[21]. Luo and colleagues present a nice graphical depiction of the choices facing the investigator when implementing propensity score techniques[2].

As stated above, the ultimate goal of propensity score matching is to create a cohort of treated and control subjects that are balanced on relevant covariates. Checks of the balance between the matched samples then becomes important to determine whether the estimated propensity score and matching methodology have been appropriately specified. In this area, there has been some controversy as to whether significance testing is an appropriate method for determining balance on covariates in a propensity matched sample.

On the one hand, Austin and others have argued against the appropriateness of significance testing on the grounds that significance testing is confounded by sample size[7, 20, 22]. In Austin's view, significance tests of balance may lack power to detect important differences between covariates if the matched sample is small, while trivial differences may be declared significant if the matched sample size is large[7, 20]. As an alternative, Austin has suggested the standardized difference for continuous variables as a more appropriate measure of balance. The formulation for the standardized difference (d) is as follows[20]:

$$d = \frac{100 * |mean_{treatment} - mean_{control}|}{\sqrt{\dfrac{s^2_{treatment} + s^2_{control}}{2}}}$$

Thus we see that the standardized difference places the difference in means between the treatment and control in terms of the pooled variance of the two groups. There is no widely agreed upon cutoff to define an acceptable standardized difference, though Normand has defined a difference of less than 0.1 in absolute value to represent appropriate balance has been achieved[23]

On the other hand, Hansen has argued in favor of significance testing for the assessment of covariate balance[24]. In Hansen's view, hypothesis tests for balance tend to reject the null hypothesis of balance when the difference in covariates is enough to introduce significant bias into the subsequent causal inferences. On the other hand, significance tests of balance fail to reject the null hypothesis when differences between covariates are small enough that any subsequent bias in the causal estimates is ignorable. The use of significance testing for balance does have a tradition in the literature[2, 9, 25]. Luo and colleagues make the reasonable suggestion that, if hypothesis testing is to be used, then tests appropriate for matched pairs such as the paired t-test or Mcnemar's test should be chosen[2].

Having estimated propensity scores, formed matched cohorts, and appropriately assessed balance between the groups, the investigator is left with the final choice of how to perform significance testing on outcomes. The basic distinction to be made is whether the matched samples should be treated as independent. Schafer has argued that the treated and control groups in the matched sample should be regarded as independent[7, 16], which would lead to tests such as the two sample t-test for continuous outcomes and

Pearson's chi-squared test for categorical outcomes. Austin, on the other hand, has made the argument that subjects matched on important baseline covariates are likely to have similar outcomes, and thus methods that account for the lack of independence between groups are more appropriate. In this manner, Austin suggests methods such as the paired t-test and Mcnemar's test for assessment of continuous and categorical outcomes, respectively[7].

Having now described the basis of the propensity score and suggestions for its use in observational studies, I will now turn to specific examples of how the propensity score can be used in settings where randomized trials are either impossible or impractical. In the first case, the disparity in donor organ quality between African American and Caucasian recipients will be examined. This is a case where randomization is clearly impossible, as recipient race is an intrinsic characteristic of the patient not subject to investigator control. In the next case, outcomes after laparoscopic versus open resection of hepatic colorectal metastases will be evaluated. A randomized controlled trial in this setting is theoretically possible, but rendered infeasible given that patients are unlikely to agree to be randomized to a more invasive procedure[26].

CHAPTER II

THE IMPACT OF DONOR QUALITY ON THE OBSERVED DISPARITY IN GRAFT SURVIVAL BETWEEN AFRICAN AMERICAN AND CAUCASIAN KIDNEY TRANSPLANT RECIPIENTS

Kidney transplantation remains the gold standard for treatment of patients with end stage renal disease when compared to long term maintenance dialysis, which has been demonstrated across ethnic groups[27]. As the field of transplantation has matured a number of disparities in access[28-30] to transplantation, as well as in post-transplant outcomes[30-34] between African American and Caucasian recipients have come to light. Several authors have proposed factors contributing to these disparities, including socioeconomic[31, 34, 35], genetic[36], immunologic[37, 38], and even pharmacokinetic[39] issues.

Issues of donor quality also play a major role in post-transplant outcomes. Kidneys from African American donors, for example, have been shown to be associated with an increased risk of graft loss[32, 40, 41], especially when transplanted into other African Americans[42]. The importance of donor quality in post-transplant outcomes is especially clear when looking at kidneys from expanded criteria donors, which pose a 70% greater risk of graft failure than lower risk organs[43]. Despite the known importance of donor quality, there is a relative paucity of large scale studies examining racial disparities in terms of donor quality in a systematic manner [44].

12

We undertook this present study to determine the presence of and any possible contributing factors to differences in donor quality between African American and Caucasian first time recipients of deceased donor kidney transplants, as well as to ascertain the effect of any donor quality disparity on post-transplant outcomes.

Patients and Methods

Selection of Patients for Analysis

After receiving an institutional review board exemption, an analysis of the United Network for Organ Sharing/Organ Procurement and Transplantation Network database as of 3/31/2011 was performed. The target population for this study was Caucasian or African American first time recipients of deceased donor kidney alone transplants from 1/1/2000 through 12/31/2009.

Assessment of Donor Quality

The primary goal of this study was to determine disparity in donor quality between African American and Caucasian recipients. The kidney donor risk index (DRI) derived by Rao, et al. [45] was chosen as the measure for donor quality for this study as it avoids the creation of arbitrary cutpoints in continuous data, and that it provides more granular information than more traditional measures such as the expanded donor criteria. The DRI includes donor age, race, history of hypertension and diabetes, serum creatinine, cause of death, height and weight, donation after cardiac death status, hepatitis C status, B and DR locus mismatch level, cold ischemic time, and whether an enbloc or double kidney transplant is to be performed. The score represents the estimated risk of graft failure relative to a hypothetical standard donor with a DRI of 1.0. For example, a DRI of 1.3 confers a 30% greater risk of graft failure than the reference donor.

Statistical Analysis

Comparisons were made between the African American and Caucasian recipient cohorts. Continuous variables were summarized as mean/standard deviation and analyzed using Student's t-test while categorical variables were summarized as count/percentage and analyzed using the chi-squared or Fisher's exact test, where appropriate. Graft survival was calculated as the time from transplant to graft failure, with censoring at the time of death with a functioning graft or last follow-up (right-censoring) in the case of Cox proportional hazards analysis[46]. The cumulative incidence of graft failure in African Americans versus Caucasians was calculated by competing risks analysis[47] and compared using a modified chi-squared statistic as described by Gray[48]. In this analysis, death and graft failure were treated as competing risks, while patients still alive at the end of follow-up with a functioning graft were censored.

To control for confounders related to recipient race that may influence observed differences in DRI, a 1:1 propensity matched[49] cohort of African American and Caucasian recipients was created using a nearest neighbor algorithm[50]. Propensity scores were estimated through logistic regression, with immunologic (HLA-A, -B, and – DR antigens, peak PRA, and ABO antigen), socioeconomic (payment source, educational achievement, employment status, and UNOS Region), and medical (age, recipient bmi, etiology of renal failure, recipient diabetes status, weight, gender, recipient HCV status, recipient CMV status, and days on the waiting list) factors as independent variables and recipient race as the dependent variable. The result of this model is to assign each patient a propensity score that describes the probability of the patient either being African American or Caucasian based upon the independent variables entered in the model. The form of the model is as described above in Chapter I:

$$log\frac{e(x)}{1 - e(x)} = \beta^{-1}x$$

Where e(x) is the propensity toward being African American, and **x** is the vector of medical, immunologic, and socioeconomic baseline covariates described above. These covariates were chosen on the empirical assumption that they are correlated with recipient race and/or donor selection (which is the desired outcome to ultimately analyze). Matching was on a 1:1 nearest neighbor fashion based on the estimated propensity score, with a caliper of 0.1 imposed to ensure that matching was within the zone of common support. Furthermore, the matching algorithm was "greedy" in that once made, matches were not broken. African American recipients without a corresponding Caucasian control who had a propensity score within the specified caliper were discarded.

After creation of the propensity matched cohort, balance on the covariates included in the model estimating the propensity score was assessed using matched pair techniques. Specifically, continuous covariates were compared using the paired t-test and the Spearman signed rank test, where appropriate. Categorical covariates were compared using McNemar's test or its extension for larger dimension square tables, or conditional logistic regression for non-square tables. Differences in the overall DRI as well as individual components were then assessed in the propensity matched cohort, again using paired data techniques as described for assessment of balance.

To explore the contribution of DRI to disparities in graft survival between African Americans and Caucasians, several Cox regression models were examined in sequence. First univariable models were fitted with recipient race and DRI respectively as the sole predictor variables to determine the baseline unadjusted hazard for graft failure posed by these factors. Next, to control for other potentially relevant covariates, a

multivariable Cox regression model was fitted using forward selection of variables significant at the $p<0.05$ level. Factors considered in the model were the same covariates used to create the propensity score described above. Donor risk index was then included as a covariate in the third and final multivariable model. The difference between the hazard ratios corresponding to African Americans was then examined in the progressively adjusted models. Reductions in the hazard ratio would indicate that the additional covariates partially explain the observed disparity in graft survival between African American and Caucasian recipients.

As a final test of the contribution of DRI disparity to the difference in graft survival between African Americans and Caucasians, graft survival was analyzed in a cohort of African Americans and Caucasians matched on DRI. The matching process was similar to that used in the propensity matching scheme, namely 1:1 matching based on a nearest neighbor algorithm. The hazard ratio for African American race in this matched cohort was then compared to the unadjusted hazard ratio for African American race in the overall cohort. The percent of excess hazard explained by matching on DRI was calculated as $(HR_{unmatched}-HR_{matched})/(HR_{unmatched}-1)*100\%$.[51] The interaction between DRI and recipient race was also examined by performing Cox proportional hazards regression of graft survival stratified by DRI in the overall cohort. The three strata analyzed were DRI <1, $1 \leq DRI < 1.5$, and DRI ≥ 1.5.

Sensitivity analysis was also performed to determine whether significant bias may have been introduced into the analysis by exclusion of patients with missing data that did not allow for calculation of the DRI. First, chi squared analysis was performed to determine whether patients with missing DRI data differed significantly by race. DRI for the patients with missing component variables was then calculated using multiple imputation, whereby five imputations were performed in which missing DRI component

variables were predicted based on remaining DRI variables that were available for each patients. The five imputations were then combined to give estimates of the mean and standard deviation for the DRI in subset of patients with missing data. The imputed DRIs were then compared to the DRI for African Americans and Caucasians with complete data to determine whether the patients with missing data were significantly different from the subset with complete data. Cox proportional hazards analysis was then performed to determine whether graft failure in the patients with missing DRI data was significantly different from the patients who had complete data. All statistical analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC), with the exception of competing risks analysis, which was performed in R (The R Project for Statistical Computing).

Results

Overall Recipient and Donor Characteristics

There were 55,982 recipients included in the study with complete DRI data collection, of whom 33,405 (59.67%) were Caucasian and 22,577 (40.33%) were African American. There were 3,916 Caucasians and 2,534 African Americans in the original dataset that had missing data for DRI and thus were not included in the analysis. The majority of patients analyzed were male (61.02%) at a mean age of 50.74 years at the time of transplant. Mean time on the waiting list was 734.02 days. Peak PRA was a mean of 12.79%. Remaining descriptive statistics of the recipients in this study are detailed in Table 2.1.

The average age of donors was 38.14 years, with 59.70% being males (n=33,423) and 7,496 (13.39%) African Americans. Terminal creatinine was a mean of 1.11 mg/dl. Cause of death was cerebrovascular accident in 21,717 (38.79%). Donation

17

after cardiac death was uncommon (8.19% of all donors) as was utilization of HCV

positive donors (2.91%). Mean cold ischemic time was 18.59 hours. The mean DRI was

1.21. Remaining donor characteristics are summarized in Table 2.2.

Differences in Donor Risk Index By Recipient Race

African American recipients received kidneys from higher risk donors as

measured by most components of the donor risk index(table 2.1, 2.2). Specifically,

African Americans were more likely to receive a kidney from an African American donor

(20.37% vs. 8.67%; p<0.001). Utilization of hepatitis C positive donors was also

significantly higher among African American compared to Caucasian recipients (5.43%

vs. 1.20%; p<0.001). There was also a statistically significant, though likely not clinically

significant, increase in cold ischemic time for African American recipients (18.86 vs.

18.41 hours; p<0.001). Overall, donor risk index for African American Recipients was

significantly higher than for Caucasian recipients (1.27 vs. 1.17; p<0.001). The

distribution of the DRI in African American and Caucasian recipients is presented in

figure 2.1. There was no significant difference in the percentage of African Americans

and Caucasians who received organs from extended criteria donors (17.66%, n=3986

vs. 17.98%, n=6005; p=0.330).

Propensity Matched Analysis

Matching on propensity score yielded a cohort of 2446 Caucasian and 2446

African American Recipients (distribution of propensity scores pre and post matching are

presented in Figure 2.2 and Figure 2.3, respectively). The African American and

Caucasian groups were well matched on etiology of renal failure (p=1.0), ABO type

(p=1.0), UNOS region (p=0.95), A1 (p=0.99), A2 (p=0.92), B1 (p=0.76), B2 (p=1.0),

DR1 (p=0.81), and DR2 (p=0.99) haplotypes, payment source (p=1.0), HCV positivity

(6.17% vs. 6.13%; p=0.95), cytomegalovirus positivity(7.15% vs. 7.85%; p=0.94), gender (62.2% male vs. 63.0% male; p=0.55), educational achievement (p=0.91), diabetes status (p=0.81), age (mean difference -0.61 years, 95% CI -1.32 -0.11; p=0.10), peak PRA (mean difference - 0.20, 95% CI -1.74 - 1.33; p=0.979), weight (mean difference - 0.58kg, 95% CI -1.69 - 0.52; p=0.30), body mass index (mean difference -0.32, 95% CI - 0.78 - 0.14; p=0.17), pretransplant creatinine (mean difference -0.08mg/dl, 95% CI -0.27 - 0.10; p=0.38), and days on the waiting list (mean difference -11.5 days, 95% CI -47.5 - 24.5; p=0.53).

In the propensity matched cohort, there remained a minor difference in mean DRI between the African American (1.28) and Caucasian (1.25) cohorts (mean difference 0.03, 95% CI .005-0.06; p=0.02)(Figure 2.4). Comparison of donor risk index components between the two groups are outlined in table 2.3. Notably, African Americans are still significantly more likely than Caucasians to receive a kidney from an African American donor (15.74% vs. 13.41%, OR 1.21, 95% CI 1.03 – 1.41; p=0.02). Furthermore, utilization of HCV positive donors remained significantly higher in African American recipients (3.97% vs. 1.96%, OR 2.06, 95% CI 1.45-2.93; p<0.001).

Effect of Recipient Race and Donor Risk Index on Graft Survival

Both DRI and recipient ethnicity significantly correlated with graft survival in the univariable models. The hazard ratio for each one point increase in DRI in the over the baseline of 1.0 in the unadjusted model was 2.2 (95% CI 2.071-2.239; p<0.001). One, five, and ten year cumulative incidence of graft failure for African Americans was 7.5% (95% CI 7.2%-7.9%), 25.2% (95% CI 24.5%-25.9%), and 43.9% (95% CI 42.4%-45.3%) compared to 5.6% (95% CI 5.3%-5.8%), 14.8% (95% CI 14.4%-15.3%), and 25.7%

(95% CI 24.8%-26.6%) (p<0.001) for the Caucasian cohort. The hazard ratio for African Americans relative to Caucasians in the unadjusted model was 1.8 (95% CI 1.7-1.9; p<0.001). The median followup for the overall, African American, and Caucasian cohorts was 37.2 months, 35.9 months, and 41.5 months, respectively. The number of graft failures (not including death with a functioning graft) in the overall, African American, and Caucasian cohorts were 10,132, 5,245, and 4,887, respectively.

The multivariable Cox regression model for graft survival after forward selection included recipient race, age, and weight, etiology of renal failure, employment status, UNOS region, payment source, and recipient hepatitis C status. The hazard ratio for African Americans relative to Caucasians in this model was reduced to 1.5 (95% CI 1.3-1.6; p<0.001). After entering DRI as a covariate in the above multivariable model, the hazard ratio for African Americans was further reduced to 1.3 (95% CI 1.2-1.4; p<0.001). Hazard ratios for all other covariates included in the model are presented with the supplementary material (Table 2.4).

As a further test of the importance of DRI disparities on the differential graft survival seen with African American and Caucasian recipients, matching for donor risk index yielded a cohort of 22,466 African Americans and 22,466 Caucasians with a mean DRI of 1.26 for each group (p=1.0). The hazard ratio for African Americans relative to Caucasians in this matched cohort was reduced to 1.6 (95% CI 1.5-1.7; p<0.001), giving a percent of excess hazard explained of 25%.

Finally, examination of the interaction between DRI and race in terms of graft survival was undertaken using stratified analysis as outlined above. In the lowest DRI stratum (DRI <1.0), the HR associated with African American Race was 2.1 (95% CI 1.9-2.2; p<0.001). In the middle stratum (1.0≤DRI<1.5), the HR associated with African

American race decreased to 1.6 (95% CI 1.5-1.7; p<0.001). In the highest DRI stratum (DRI≥1.5), the HR associated with African American race was further reduced to 1.4 (95% CI 1.3-1.5; p<0.001).

Sensitivity Analysis for Missing DRI Data

There were 2534 (10.09%) African Americans with missing DRI data compared to 3916 (10.49%) Caucasians (p=0.106), indicating that patients with missing DRI were equally distributed by race. Using multiple imputation as described above, the mean DRI for African Americans with missing data was predicted to be 1.26 (0.465) compared to 1.27 (0.445) for those with complete data (p=0.486). The mean imputed DRI for Caucasians with missing data was 1.15 (0.482) compared to 1.17 (0.542) for those with complete data (p=0.058). Thus, the patients with missing data did not differ significantly in terms of DRI from those with complete data. In terms of graft failure, Caucasians with missing DRI had statistically similar risk of graft failure compared to those with complete DRI (HR for missing DRI = 1.015, 95% CI 0.936-1.101; p=0.716). African Americans with missing DRI data had significantly worse graft failure than their counterparts with complete DRI data (HR for missing DRI = 1.114, 95% CI 1.033-1.200; p=0.005). Thus, by not including the patients with missing data, our analysis of disparity in graft survival between African Americans and Caucasians may actually be conservative.

Discussion

This study demonstrates that African American recipients receive significantly "riskier" organs as defined by the DRI than their Caucasian counterparts, and that this disparity cannot completely be explained by socioeconomic, immunologic, or other medical factors. Furthermore, this disparity on organ donor quality has a significant impact on graft survival. The negative effects of the donor quality disparity are illustrated

by our analysis of graft survival, in which the percent of excess hazard explained by matching on DRI was 25%. The analysis of graft survival between African Americans and Caucasians indicates that the contribution of DRI to the observed higher risk of graft failure in African Americans is most prominent in the higher DRI strata. Gaining a clear understanding of the reasons underlying the lack of access to quality organs faced by African Americans thus becomes crucial if outcomes are to be improved.

The propensity matched model of DRI demonstrates that multiple immunologic, medical, and socioeconomic factors account for most, but not all, of the disparity in DRI. Two of the most important contributors to the disparity in organ quality appear to be increased use of organs from African American and hepatitis C positive donors in African American recipients. Although matching for HLA type and region did reduce the disparity in usage of African American donor kidneys, the odds of an African American vs. a Caucasian receiving a kidney from an African American donor remained 1.2 after matching. That such a difference remains after extensive matching suggests the influence of unmeasured factors which merit further study, such as the role of minor blood group antigens. Other potential explanations include that African Americans recipients and donors may live in closer proximity to each other, or perhaps that transplant surgeons preferentially allocate organs to recipients of the same race.

The fact that African Americans remain more likely to receive a kidney from a hepatitis C positive donor than their Caucasian counterparts after matching for recipient hepatitis C status presents a similar quandary. Given that the patients were also matched on educational achievement, it may be assumed that there existed a similar level of sophistication on the part of recipients in both groups. Despite this, African Americans apparently were more likely to receive a kidney from a hepatitis C positive

donor than their Caucasian counterparts. This continued differential suggests an opportunity for a more complete informed consent process.

The underlying causes for the disparity in kidney donor quality faced by African Americans are likely multifactorial. Adjustment for socioeconomic, immunologic, and clinical factors closes the gap somewhat, giving evidence to their significant role in the disparity. Even when taking these factors into consideration, African Americans still receive significantly more "risky" organs than their Caucasian counterparts. Given the known barriers to transplantation faced by African Americans[30, 52], transplantation with a lower quality organ may be preferable to remaining on dialysis; however, such a decision requires thoughtful and thorough informed consent and patient education. Furthermore, the disparities in deceased donor organ quality faced by African American recipients underscores the importance of efforts to increase living donation among the African American community. Successful examples of such educational efforts can be seen in the work of Callender and Foster [34, 53].

If the disparity in post-transplant outcomes between African Americans and Caucasians is to be remedied, then we clearly need to improve our efforts to understand and address the donor quality gap. It is upon the transplant community to ensure that the distribution of higher risk organs to African Americans as illustrated above is not the result of a failure of transparency on our part. Patients place a tremendous amount of trust in their physicians because of our "expert" status. We must ensure that such trust is not misplaced by taking extraordinary efforts to ensure that patients are properly educated to make the best decisions for themselves. Organ donor quality and its effect on post-transplant outcomes should be a mandatory part the discussion held with the patient when determining whether to accept a particular organ offer. Failure to clearly outline the risks posed by donor factors would be to perpetuate the disparity in the

quality of organs accepted by and for African American patients. As a prescriptive measure, we echo Rao's call to include the DRI, or a similar measure of organ quality, in both the allocation and consent processes[45]. In the end, the decision of whether or not to accept a particular organ must lie with the patient, and be based on full disclosure of all relevant information and informed consent. Only when we have fully disclosed and discussed issues of organ quality with the patient can we say that informed consent has truly taken place. That such an informed consent process can be effective is evidenced by the fact that African Americans and Caucasians accepted organs from extended criteria donors at similar rates in this study. Were potential recipients made as aware of the risks of high DRI kidneys as they are of the risks of ECD kidneys, then perhaps the gap in DRI between African Americans and Caucasians would be mitigated.

In this study we have outlined several factors that contribute to this gap in what is, to our knowledge, the largest study to systematically address donor quality issues in African American recipients. The disparities underscored by our study, such as use of hepatitis positive and African American donors, certainly have more subtle effects on graft survival than more traditional measures of donor quality such as the expanded donor criteria, which were not significantly different in the matched cohort. When taken all together, however, these small differences combine to produce a significant effect. It is through acknowledging and paying appropriate attention to all the small details of patient care that optimal outcomes are achieved. Furthermore, it may be preferable in the future to preferentially allocate higher risk organs to patients with shorter life expectancy in order to mitigate the effects of shorter expected graft longevity.

Despite this data, there remain other unmeasured factors yet to be determined that merit future study. It will require further vigilance on the part of the transplant community going forward to ensure that one of these factors is not the failure to disclose

important issues of donor quality. Limitations of this study are largely a product of its observational nature. We have included a number of potential confounders that can potentially explain observed differences in DRI in our model; however, there may be other factors which were not included and may contribute to residual confounding. The most important of these is the inability to control for specific immunosuppression protocols, which is the factor which likely has the greatest impact on graft survival, and may have differing effectiveness across racial groups. Other factors may include minor blood group antigens and willingness on the part of the patient(of either race) to accept suboptimal organs as an alternative to remaining on dialysis[54]. We have also used multivariable modeling of graft survival in order to control for the effects of confounders. Though this approach has been widely used in the literature, there is the possibility that such models do not completely remove the effects of confounding variables. Another weakness of our Cox regression model is that the hazard associated with DRI and African American race appear to change with time, thus making our estimates less precise than a model where race and DRI were treated as time dependent. Though such analysis would add precision to a model intended to *predict* survival, that was not the goal of the current manuscript. The Cox regression analysis was of death censored graft failure, which tends to overestimate risk in the setting of competing risks. Were death instead treated as a competing risk, the observed difference in graft survival between African American and Caucasians may differ from the findings in the current manuscript.

Another potential weakness is our use of the DRI as a surrogate for organ quality, as it differs from traditional markers such as the expanded donor criteria. One particular potential weakness of the DRI is that it does not completely isolate donor factors, in that HLA mismatch is a product of both the donor and the recipient. Nonetheless, we feel that DRI is the most appropriate proxy for donor quality for the

reasons outlined previously in this manuscript. Finally, there were a number of patients in the original dataset for whom complete information was not available to calculate the DRI, and inclusion of these patients may potentially lead to somewhat different outcomes. Although our sensitivity analysis indicates that the patients with missing values were similar to those with complete data, the possibility of bias introduced by exclusion of patients with missing data still remains.

Table 2.1: Overall Recipient Characteristics

Variable	Caucasian (n=33,405)	African American (n=22,577)	P-value
Male Gender	20582 (61.6%)	13577 (60.1%)	**<0.001**
Hepatitis C Positive	1022 (3.1%)	2178 (9.8%)	**<0.001**
Cytomegalovirus Positive	17373 (54.2%)	16079 (74.4%)	**<0.001**
EnBloc Transplant	478 (1.5%)	321 (1.4%)	0.726
Double Kidney Transplant	484 (1.5%)	374 (1.7%)	0.050
Age (mean, std)	48.3 (14.1)	52.4 (15.0)	**<0.001**
Peak Panel Reactive Antibody (median, IQR)	0 (7)	0 (14)	**<0.001**
A Locus Mismatch (median, IQR)	1 (1)	2 (1)	**<0.001**
B Locus Mismatch (median, IQR)	1 (1)	2 (1)	**<0.001**
DR Locus Mismatch (median, IQR)	1 (2)	1 (1)	**<0.001**
Total HLA Mismatch (median, IQR)	4 (3)	5 (1)	**<0.001**
Height (cm, mean, std)	170.2 (14.5)	171.1 (13.6)	**<0.001**
Days on the waitlist (median, IQR)	478 (702)	756 (973)	**<0.001**

Table 2.1: Comparison of recipient characteristics between African Americans and Caucasians. Categorical variables are presented as count(percentage). Abbreviations: std(standard deviation), iqr(interquartile range), HLA(human leukocyte antigen). Components of the DRI are highlighted in bold.

Table 2.2: Donor Characteristics

Variable	Caucasian Recipients (n=33,405)	African American Recipients (n=22,577)	P-value
Donor Risk Index	1.17(0.43)	1.27(0.45)	**<0.001**
Age (years,mean, std)	38.1 (17.2)	38.3 (16.8)	0.186
Terminal Cr (mg/dl, mean, std)	1.08 (0.85)	1.14 (0.97)	**<0.001**
Weight (kg, mean, std)	77.4 (23.6)	77.7 (23.7)	0.175
Height (cm, mean, std)	169.3(17.7)	169.2(17.7)	0.450
Cold Ischemic Time (hours, mean, std)	18.4 (8.9)	18.9 (9.5)	**<0.001**
Male Gender	19848 (59.4%)	13575 (60.1%)	0.092
African American Race	2896 (8.67%)	4600 (20.37%)	**<0.001**
History of Hypertension	8077 (24.18%)	5963 (26.41%)	**<0.001**
History of Diabetes	1884 (5.64%)	1405 (6.22%)	**0.004**
Death From Stroke	12865 (38.51%)	8952 (39.21%)	0.097
Donation after Cardiac Death	2631 (7.88%)	1954 (8.65%)	**0.001**
Hepatitis C Positive	402 (1.20%)	1226 (5.43%)	**<0.001**

Table 2.2: Comparison of donor characteristics by recipient race. Categorical variablesare presented as count(percentage). Components of the Donor Risk Index are highlighted in bold. Abbreviations: std(standard deviation).

Table 2.3: Differences in Donor Risk Index Components in a Propensity Matched Cohort of African American and Caucasian Recipients

	Caucasian (n=2,446)	African American (n=2,446)	P-value
Donor Age (mean, std)	39.45(16.71)	39.67(16.76)	0.64
Donor Weight(kg, mean, std)	79.00(24.42)	79.17(24.41)	0.80
Donor Height (cm, mean, std)	169.4(17.7)	169.1(18.0)	0.61
Cold Ischemic Time (hours, mean, std)	19.12(10.99)	18.75(9.56)	0.21
B Locus Mismatch (median, iqr)	2 (1)	2 (1)	**0.006**
DR Locus Mismatch (median,iqr)	1 (1)	1 (1)	**0.03**
Donor Terminal Creatinine (mg/dl, mean, std)	1.16(0.83)	1.14(0.70)	0.45
African American Donor	328(13.41%)	385(15.74%)	**0.02**
Diabetic Donor (n, %)	164(6.70%)	196(8.01%)	0.08
Hypertensive Donor (n, %)	704(28.78%)	706(28.86%)	0.95
CVA as cause of death (n, %)	984(40.23%)	974(39.82%)	0.77
Hepatitis C Positive Donor (n, %)	48(1.96%)	97(3.97%)	**<0.001**

DCD Donor (n, %)	245(10.02%)	273(11.16%)	0.20
Double Kidney Transplant (n, %)	33(1.35%)	50(2.04%)	0.06
En Bloc Transplant (n, %)	33(1.35%)	35(1.43%)	0.81

Table 2.3 (cont). Abbreviations: DCD (donation after cardiac death), CVA (cerebrovascular accident),std (standard deviation)

Parameter	Parameter Estimate	P-value	HR	Lower 95% CI	Upper 95% CI
DRI	0.87	<0.001	2.38	2.23	2.54
African American Race	0.26	<0.001	1.30	1.21	1.40
AGE (years)	-0.02	<0.001	0.98	0.98	0.99
Weight (kg)	0.01	<0.001	1.01	1.01	1.01
Etiology of Renal Failure (Reference: Hypertensive Nephrosclerosis)					
OTHER	0.14	0.04	1.15	1.01	1.32
CRESCENTIC GLOMERULONEPHRITIS	-0.96	0.05	0.38	0.14	1.02
MEMBRANOUS GLOMERULONEPHRITIS	0.08	0.51	1.09	0.85	1.40
MESANGIO-CAPILLARY 1 GLOMERULONEPHRITIS	-0.01	0.98	0.99	0.41	2.38
MESANGIO-CAPILLARY 2 GLOMERULONEPHRITIS	1.19	0.24	3.28	0.46	23.33
IGA NEPHROPATHY	-0.15	0.20	0.87	0.69	1.08
ANTI-GBM	0.36	0.47	1.44	0.54	3.84
FSGS	0.04	0.56	1.04	0.92	1.18
REFLUX NEPHROPATHY	-0.20	0.26	0.82	0.58	1.16
POLYCYSTIC KIDNEYS	-0.34	<0.001	0.71	0.62	0.82
NEPHRITIS	0.19	0.23	1.21	0.89	1.65
NEPHROPHTHISIS	0.80	0.17	2.22	0.71	6.91
DIABETES - TYPE I INSULIN DEPENDENT /JUVENILE ONSET	0.18	0.81	1.19	0.29	4.90
DIABETES - TYPE II NON-	-0.40	0.49	0.67	0.22	2.08

INSULIN DEPENDENT/ADULT ONSET					
OXALATE NEPHROPATHY	1.42	<0.001	4.13	2.06	8.31
CYSTINOSIS	-0.78	0.43	0.46	0.06	3.25
FABRY'S DISEASE	-0.94	0.35	0.39	0.06	2.76
AMYLOIDOSIS	-0.27	0.58	0.76	0.28	2.03
SYSTEMIC LUPUS ERYTHEMATOSUS	0.24	0.01	1.27	1.07	1.50
PROGRESSIVE SYSTEMIC SCLEROSIS	-0.34	0.73	0.71	0.10	5.05
RENAL CELL CARCINOMA	-0.42	0.18	0.65	0.35	1.22
MYELOMA	-0.26	0.79	0.77	0.11	5.47
HEMOLYTIC UREMIC SYNDROME	0.31	0.30	1.37	0.75	2.48
HYPOPLASIA/DYSPLASIA/DYSG ENSIS/AGENESIS	-0.21	0.61	0.81	0.36	1.81
CORTICAL NECROSIS	0.44	0.66	1.56	0.22	11.08
ACUTE TUBULAR NECROSIS	0.49	0.23	1.64	0.73	3.66
MEDULLARY CYSTIC DISEASE	0.59	0.12	1.81	0.86	3.81
SICKLE CELL ANEMIA	-0.03	0.94	0.97	0.46	2.05
ACQUIRED OBSTRUCTIVE NEPHROPATHY	0.32	0.21	1.37	0.84	2.25
ALPORT'S SYNDROME	-0.70	0.01	0.50	0.29	0.86
FAMILIAL NEPHROPATHY	-0.40	0.49	0.67	0.22	2.08
GOODPASTURE'S SYNDROME	-0.15	0.66	0.86	0.45	1.67
MALIGNANT HYPERTENSION	0.03	0.76	1.03	0.85	1.24
HENOCH-SCHOENLEIN PURPURA	0.15	0.84	1.16	0.29	4.64
PRUNE BELLY SYNDROME	-0.58	0.56	0.56	0.08	3.98
DIABETES - TYPE I NON-INSULIN DEPENDENT /JUVENILE ONSET	0.99	0.32	2.70	0.38	19.21

DIABETES - TYPE II INSULIN DEPENDENT /ADULT ONSET	0.10	0.86	1.11	0.36	3.44
CHRONIC GLOMERULONEPHRITIS UNSPECIFIED	-0.20	0.03	0.82	0.68	0.98
MEMBRANOUS NEPHROPATHY	0.24	0.33	1.27	0.79	2.05
CHRONIC GLOMERULOSCLEROSIS UNSPECIFIED	-0.44	0.13	0.65	0.37	1.14
ANALGESIC NEPHROPATHY	0.45	0.07	1.57	0.97	2.53
RADIATION NEPHRITIS	0.56	0.43	1.75	0.44	7.01
ANTIBIOTIC-INDUCED NEPHRITIS	1.07	0.06	2.91	0.94	9.05
CANCER CHEMOTHERAPY INDUCED NEPHRITIS	0.03	0.96	1.03	0.33	3.20
CALCINEURIN INHIBITOR NEPHROTOXICITY	-0.11	0.55	0.90	0.63	1.28
HEROIN NEPHROTOXICITY	0.16	0.75	1.17	0.44	3.14
RENAL ARTERY THROMBOSIS	-0.83	0.41	0.44	0.06	3.11
CHRONIC NEPHROSCLEROSIS-UNSPECIFIED	-0.16	0.57	0.86	0.50	1.48
CONGENITAL OBSTRUCTIVE UROPATHY	0.14	0.45	1.15	0.80	1.66
SCLERODERMA	0.54	0.15	1.72	0.82	3.63
WEGENERS GRANULOMATOSIS	-0.84	0.01	0.43	0.23	0.84
POLYARTERITIS	-0.69	0.49	0.50	0.07	3.58
SARCOIDOSIS	-0.26	0.61	0.77	0.29	2.06
LYMPHOMA	1.60	0.11	4.97	0.70	35.38
NEPHROLITHIASIS	0.20	0.45	1.23	0.72	2.08
DRUG RELATED INTERSTITIAL NEPHRITIS	-0.10	0.69	0.90	0.54	1.50
CHOLESTEROL EMBOLIZATION	-0.51	0.61	0.60	0.09	4.29
RAPID PROGRESSIVE GLOMERULONEPHRITIS	0.54	0.13	1.71	0.85	3.44
DIABETES MELLITUS - TYPE I	-0.04	0.56	0.96	0.83	1.11
DIABETES MELLITUS - TYPE II	-0.02	0.69	0.98	0.90	1.08
DIABETES MELLITUS - TYPE	-0.31	0.15	0.73	0.48	1.12

OTHER / UNKNOWN					
Employment Status (reference: Employed)					
Unemployed	0.15	0.00	1.16	1.07	1.26
Unknown	0.08	0.19	1.08	0.96	1.21
UNOS Region (reference: Region 11)					
Region 1	0.08	0.44	1.08	0.89	1.31
Region 2	0.09	0.12	1.09	0.98	1.22
Region 3	-0.19	0.00	0.83	0.74	0.93
Region 4	0.06	0.39	1.06	0.93	1.22
Region 5	-0.11	0.14	0.89	0.77	1.04
Region 6	-0.10	0.38	0.91	0.73	1.13
Region 7	-0.07	0.28	0.93	0.81	1.06
Region 8	-0.04	0.59	0.96	0.82	1.12
Region 9	0.05	0.46	1.05	0.92	1.21
Region 10	-0.04	0.57	0.96	0.85	1.09
Payment Source (reference: Private Insurance)					
Public insurance - Medicaid	0.17	0.03	1.19	1.02	1.39

Public insurance - Medicare FFS (Fee for Service)	0.13	0.00	1.14	1.05	1.24
Public insurance - Medicare & Choice	0.15	0.00	1.16	1.06	1.28
Public insurance - Department of VA	0.18	0.31	1.20	0.85	1.69
Public insurance - Other government	0.03	0.85	1.03	0.73	1.47
Free Care	-1.20	0.23	0.30	0.04	2.15
Public insurance - Medicare other detail not specified	1.06	<0.001	2.88	1.74	4.79
Hepatitis C Status (reference = unknown)					
Negative	-0.10	0.21	0.90	0.77	1.06
Not Determined	-0.03	0.73	0.97	0.80	1.17
Positive	0.02	0.81	1.02	0.85	1.24

Table 2.4: Multivariable Cox regression model of death censored graft failure including recipient race, donor risk index, and other potentially relevant covariates using a forward selection model. Hazard ratios are presented relative to the reference class for categorical covariates and for each one unit increase for continuous covariates. Parameter estimates and corresponding 95% confidence intervals are given for factors that reached statistical significance and thus were included in the final model. Factors that did not reach significance and thus were not included in the model were as follows: recipient body mass index, recipient creatinine, days on the waiting list, peak panel reactive antibody, recipient ABO blood group, recipient HLA antigens (a locus, b locus, and dr locus), recipient diabetes status, recipient cytomegalovirus status, recipient gender, and recipient educational achievement.
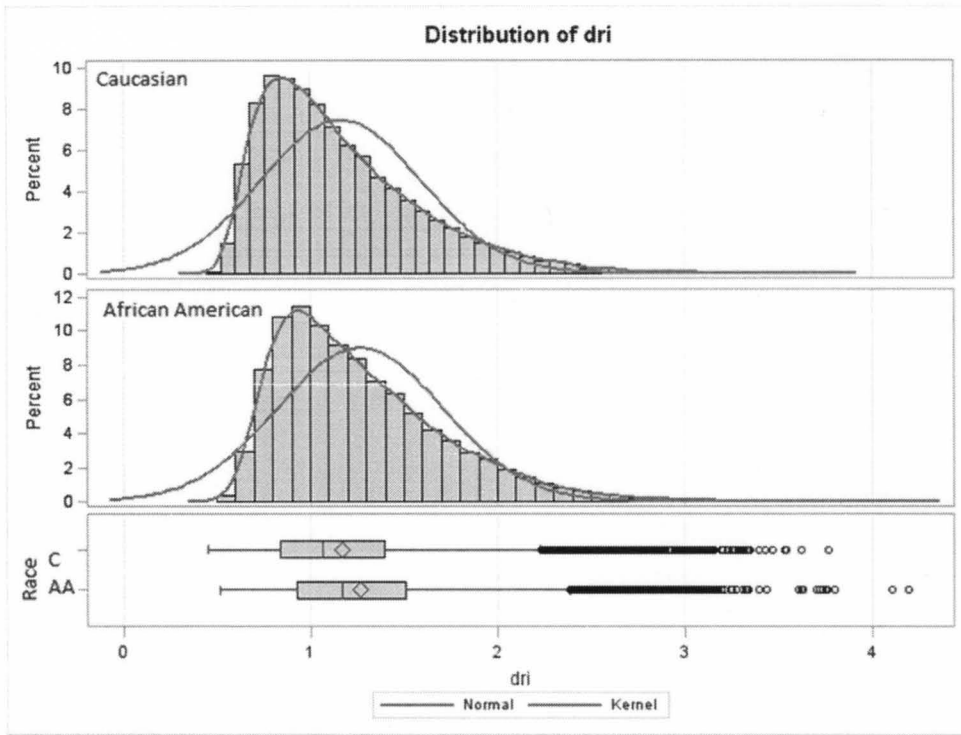
Figure 2.1: Distribution of Donor Risk Index (DRI) in Overall Cohort of Caucasian and African American Recipients.
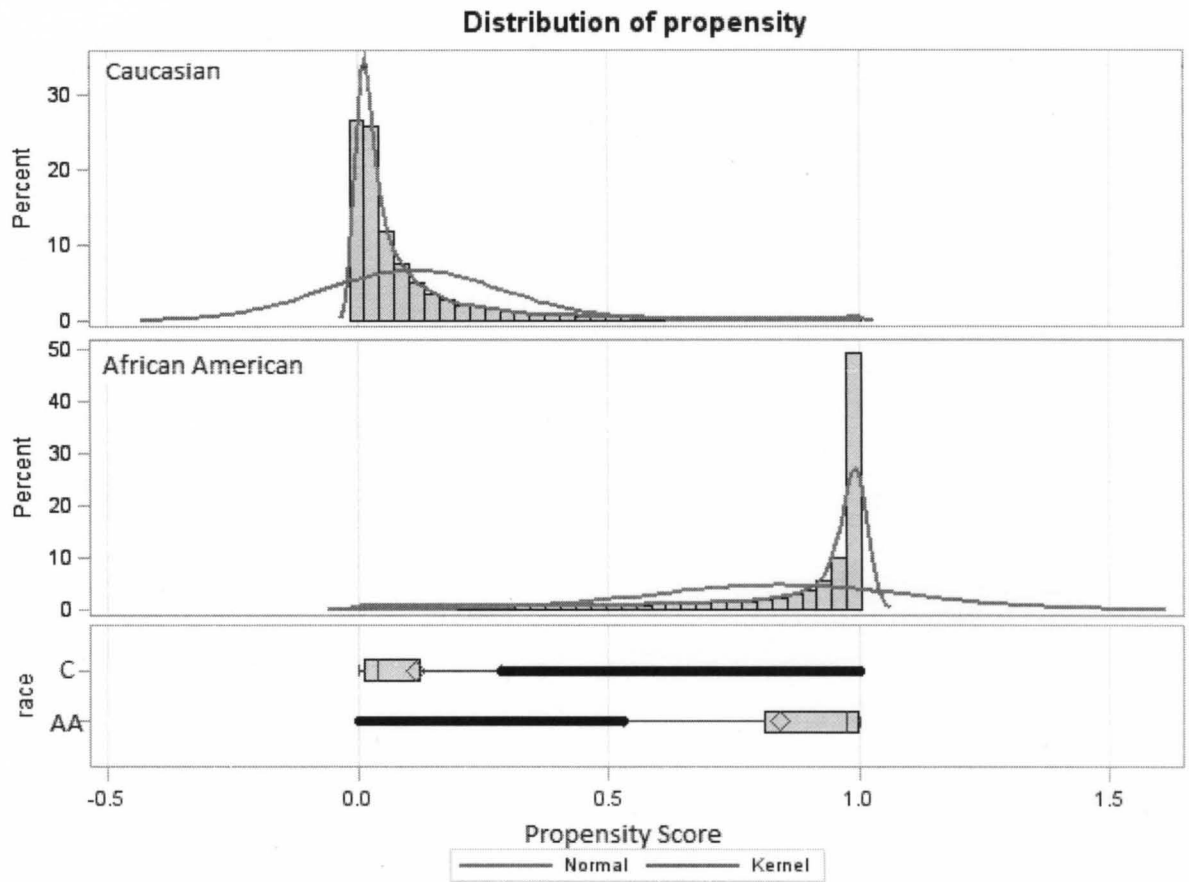
Figure 2.2: Distribution of Propensity Score in Caucasian and African American Recipients Prior to Matching
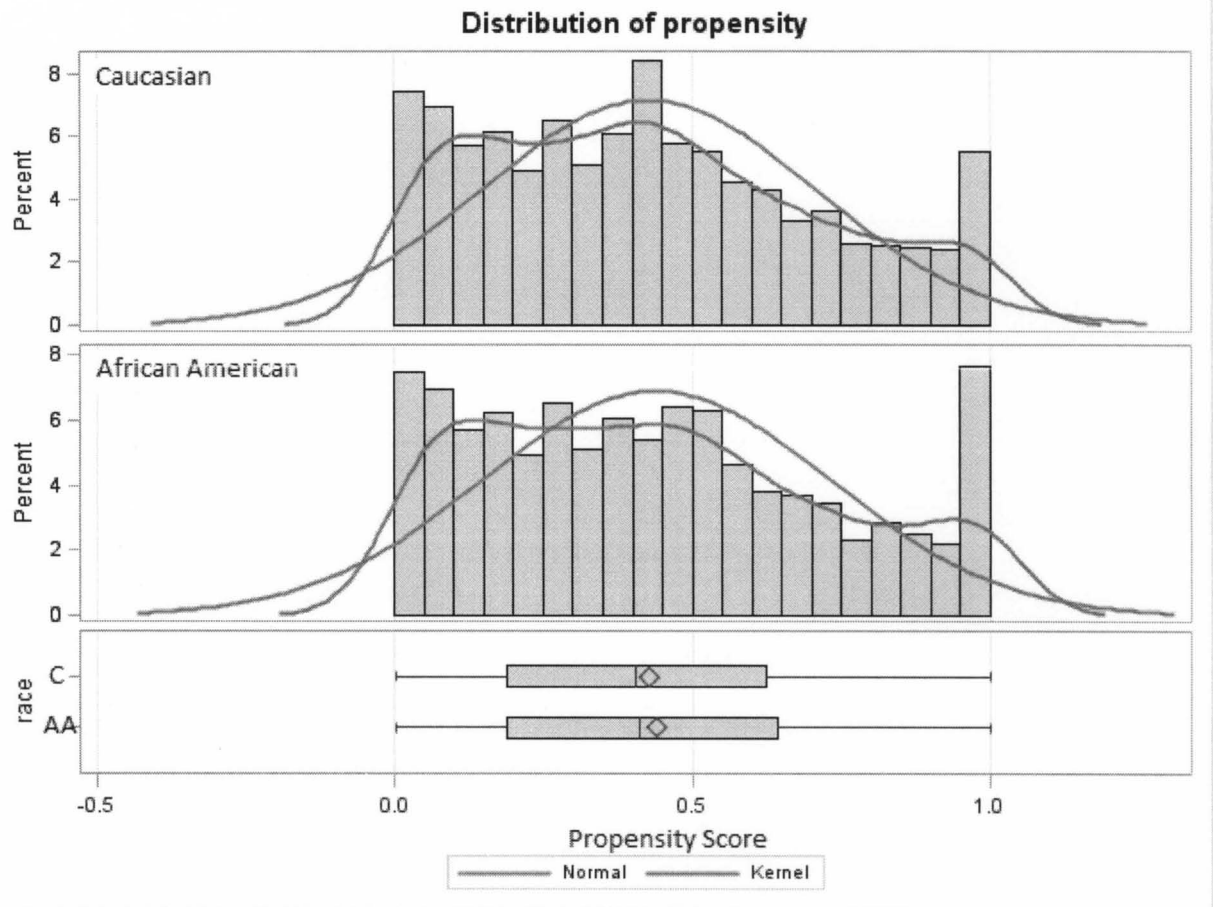
Figure 2.3: Distribution of Propensity Score in Caucasian and African American Recipients After Matching
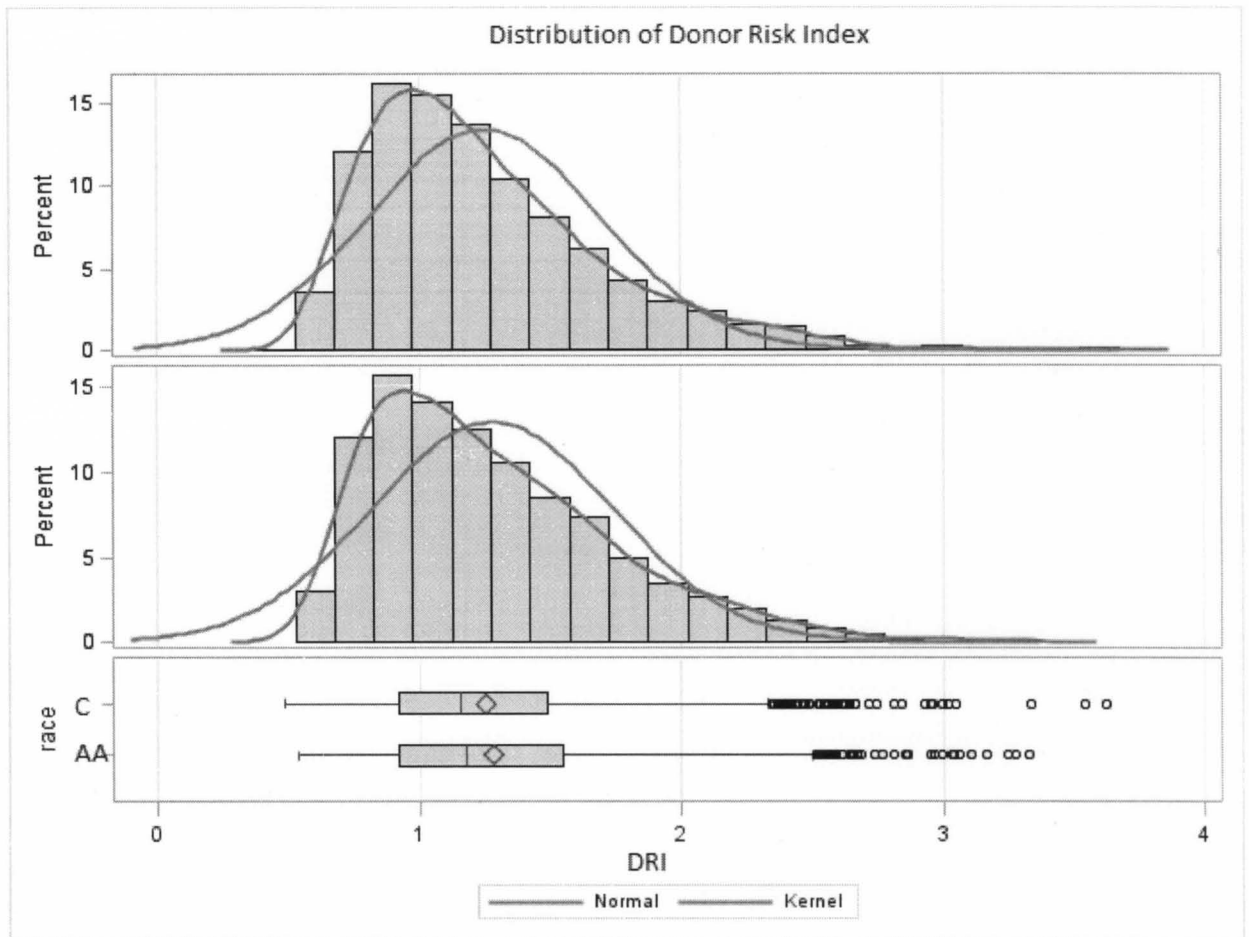
Figure 2.4: Distribution of DRI in Caucasian and African American Recipients After Matching for Propensity Score

CHAPTER III

A PROPENSITY MATCHED ANALYSIS OF LAPAROSCOPIC VERSUS OPEN RESECTION OF HEPATIC COLORECTAL METASTASES

Since the first reports of limited laparoscopic liver resection in the first half of the 1990s, the application of minimally invasive techniques to hepatic surgery has been relatively slow[55]. Acceptance of the technique within specialized centers has been steadily gaining ground since Cherqui's initial report of 30 laparoscopic hepatectomies. A recent review by Nguyen demonstrated 2,804 cases reported in the world literature[56, 57]. Experienced centers are now performing the majority of liver resections in a minimally invasive fashion, with several now reporting series of 300 patients or more[58, 59].

With respect to indication, laparoscopic resection for benign lesions gained acceptance relatively early on[26, 60]. Enthusiasm for laparoscopic resection of malignant lesions has been slower to develop, owing to concerns about the oncologic adequacy of laparoscopic hepatectomy[61]. In the realm of colorectal cancer, laparoscopic resection of the primary tumor has been demonstrated by randomized controlled trials to be oncologically equivalent to open surgery, with the benefit of shorter postoperative length of stay[62]. We have recently demonstrated the significant benefits of laparoscopic hepatic lobectomy in comparison to open hepatic lobectomy with benefits of reduction in adverse events, incision related adverse events, and reduction in hospital stay[63]. There have now been a handful of retrospective reports demonstrating the equivalence of laparoscopic and open hepatic resection for colorectal metastases as well[64-66]. This study was undertaken to evaluate the safety and efficacy of

laparoscopic resection of hepatic colorectal metastases as compared to a matched group of patients undergoing open resection.

Patients and Methods

With IRB approval, a retrospective review of all patients undergoing laparoscopic first resection of hepatic colorectal metastases by the division of surgical oncology at the University of Louisville from 1995 to 2010 was undertaken. These patients were compared to a cohort of patients undergoing open resection during the same period that were matched on a 1:1 basis. Matching was by propensity scoring[6], with scores based on patient age, size and number of lesions, performance of major hepatectomy or synchronous colectomy, and Fong score[67].

Propensity scores were estimated by logistic regression, with the propensity toward laparoscopic treatment [e(x)] being specified as follows:

$$log \frac{e(x)}{1 - e(x)} = \beta^{-1}x$$

Where β^{-1} is the vector of regression coefficients and **x** is the vector of covariates as listed above.

After estimation of propensity scores, matched cohorts were created using greedy nearest neighbor matching within a caliper of 0.1, as described in Chapter II. Two propensity matched groups were constructed. In the "inclusive" cohort, controls undergoing open resection were selected from the group of all patients undergoing open resection from 1995 through 2010. A "restricted" cohort was also created in which the controls undergoing open resection were selected from a group limited to those undergoing open resection from 2004 (when the first laparoscopic resection was performed) through 2010. The rationale for analyzing the inclusive and restricted cohorts

separately is to limit both era bias and selection bias. Use of controls from the earlier time period in the inclusive cohort has the limitation that there have been improvements in patient care and adjuvant therapy over the past two decades. Thus, utilization of controls undergoing open operation may introduce bias due to the improved adjuvant therapy in the laparoscopic cohort, which represents a more recent group of patients. On the other hand, using controls who underwent operation prior to the introduction of laparoscopy minimizes selection bias, in that even the "easy" cases underwent open resection prior to the introduction of laparoscopy. In the restricted cohort, the patients undergoing open and laparoscopic resection are contemporaneous, thus minimizing the effect of changes in adjuvant therapy. On the other hand, the restricted cohort is subject to greater selection bias as the more difficult cases preferentially receive an open operation.

Balance of the baseline covariates was assessed using the paired t-test for continuous covariates and Mcnemar's test for categorical covariates. Further comparisons were made between the matched groups in terms of patient demographics, tumor characteristics, operative factors, short term outcomes, and overall (OS) and disease free (DFS) survival. Resection margins were defined as either micsrocopically positive (R1) or negative (R0). Baseline comorbidities were assessed using the Charlson Comorbidity Index[68]. Continuous variables were summarized as mean(standard deviation) or median(interquartile range) and analyzed using the paired t-test or the Spearman signed rank test, where appropriate. Categorical variables were summarized as count(percentage) and compared using Mcnemar's test. Overall survival was calculated from the time of resection to death or last followup according to Kaplan-Meier, while disease free survival was the time from resection to recurrence, death, or last followup. Differences in survival curves were compared using the log rank test or Cox

proportional hazards regression. All statistical analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC), and p-values less than 0.05 were considered significant.

Results

Analysis of the Inclusive Cohort

There were 35 patients in the laparoscopic cohort and 35 patients in the open cohort. The groups were well matched on age, tumor size and number, body mass index (BMI), Fong score, and Charlson Comorbidity Index[68](CCI) (Table 1.3). In terms of operative procedure performed, 19(54.3%) of patients in the laparoscopic cohort underwent major hepatectomy compared to 18(51.4%) in the open cohort (p=0.808).Synchronous colectomy was undertaken in 3(8.6%) patients in the laparoscopic cohort and 2(5.7%) patients in the open cohort (p=0.655).

Patients in the laparoscopic group had significantly less operative blood loss (202ml vs. 457ml; p=0.002) than those undergoing open resection. Transfusions were required in 4 (15.4%) patients undergoing laparoscopic resection compared to 8 (30.8%; p=0.206) in the open group. At least one postoperative complication was experienced significantly less often by patients in the laparoscopic versus open group (22.9% vs. 48.6%; p=0.020). Among patients who did experience a complication, there was no significant difference in the median grade of worst complication between the laparoscopic and open groups (2.5 vs. 2.0; p=1.0). Length of stay was an average of 4.8 days versus 7.6 days in the laparoscopic versus open groups (p=0.001). Ninety day mortality was 2.9% in both cohorts (p=1.0).

In terms of oncologic outcome, R0 resection was achieved in 97.1% (n=34) of laparoscopically resected patients compared to 82.9% (n=29) of patients in the open

cohort (p=0.059). Overall survival at 1, 3, and 5 years in the laparoscopic and open cohorts was 96.9%, 62.6%, and 35.8% versus 100.0%, 66.1%, and 40.8% (Figure 3.1; p=0.484). Disease free survival in the laparoscopic and open cohorts was 79.3%, 36.96%, and 15.4% versus 90.7%, 43.6%, and 26.2% at 1, 3, and 5 years (Figure 3.2; p=0.117). There were no recurrences at laparoscopic port sites or at the laparotomy incision.

Analysis of the Restricted Cohort

There were 35 patients undergoing laparoscopic resection and 35 matched patients in the open cohort. The groups were well matched on age, tumor size and number, body mass index (BMI), Charlson Comorbidity Index[68](CCI), and Fong score (Table 3.1). Major hepatectomy was performed in 54.3% (n=19) of patients undergoing laparoscopic resection compared to 42.9% (n=15; p=0.371) of patients undergoing open resection. A synchronous colectomy was performed in 8.6% (n=3) of the laparoscopic cohort and 8.6% (n=3; p=1.0) of the open cohort.

Average blood loss was statistically similar in the laparoscopic and open groups (202ml vs. 327ml; p=0.213), as were rates of blood transfusion (15.4% vs. 34.6%; p=0.096). At least one postoperative complication was experienced in a significantly smaller percentage of the patients undergoing laparoscopic compared to open resection (22.9% vs. 48.6%; p=0.039). Of the patients who did experience a complication, the median grade of the worst complication was similar in the laparoscopic and open cohorts (2.5 vs. 2.5; p=1.0). Length of stay was significantly shorter in the laparoscopic vs. open cohort (4.8 days vs. 7.3 days; p=0.042).

In terms of oncologic outcomes, patients undergoing laparoscopic resection were significantly more likely to have a margin negative resection than patients in the open

44

cohort (97.1% vs. 80.0%; p=0.014).Overall survival at 1, 3, and 5 years in the laparoscopic and open cohorts was 96.9%, 62.6%, and 35.8% vs. 91.4%, 63.0%, and 26.3% (Figure 3.3; p=0.535). Disease free survival at 1, 3, and 5 years in the laparoscopic and open cohorts was 79.3%, 36.9%, and 15.4% vs. 79.3%, 45.0%, and 15.0% (Figure 3.4; p=0.637), respectively.

Discussion

Results with open resection of hepatic colorectal metastases have been excellent, with several centers reporting 5 year survival in excess of 50%[69, 70]. The group at Memorial Sloan Kettering have reported ten year cancer specific survival to be 23%[67] in a series of over 1,000 patients. Given these excellent results, the bar has been set relatively high for the application of laparoscopic techniques to this disease. Specific concerns about the oncologic adequacy of laparoscopy in general include port site metastases, the trophic effect of pneumoperitoneum on malignant cells, and inability to adequately inspect the peritoneal cavity, and lack of tactile sensation when inspecting the liver[61, 71-75].

Experience has shown these fears to be largely unfounded; however. In Nguyen's review of the world literature on laparoscopic liver resection, there were no port site recurrences reported following laparoscopic resection of hepatic colorectal metasases[76]. In the current study, there were no port-site recurrences. Randomized trials of laparoscopic colon resection for cancer also failed to demonstrate increased rates of wound recurrence with laparoscopy[77]. As such, fears over port site recurrence appear to be unsupported by the available evidence.

Rates of margin negative resection in this series were also similar or better in the laparoscopic versus open cohorts in this study, demonstrating that adequate delineation

of intrahepatic tumor anatomy can be obtained during laparoscopic resection. It is worth noting here the importance of facility with laparoscopic ultrasound as a key factor in obtaining such results. The ability to perform intraoperative ultrasound has been noted as one of the pre-requisite skill sets necessary before embarking on a program of laparoscopic liver resection[26]. Finally, overall and disease free survival were equivalent in the laparoscopic and open cohorts of this study. These results demonstrate that laparoscopic resection for hepatic colorectal metastasis is oncologically equivalent to open surgery.

In addition to oncologic equivalency, we have also demonstrated a number of benefits to laparoscopic hepatectomy. Patients undergoing a minimally invasive procedure experienced significantly less operative blood loss, and experienced postoperative complications at less than half the rate of their counterparts in the open cohort. Furthermore, length of stay was significantly shorter in the laparoscopic cohorts. Another potential benefit of laparoscopic liver resection is for patients who present with synchronous disease. Simultaneous resection of both the primary tumor and hepatic metastases allows for less overall hospitalization by avoiding a second operation[78], and eliminates any delay in hepatectomy while patients receive adjuvant chemotherapy. The ability to perform laparoscopic hepatectomy extends the benefits of minimally invasive surgery to this patient population. These results confirm our prior study evaluating laparoscopic hepatic lobectomy versus open hepatic lobectomy in which there were less adverse events[63].

Given the demonstrated patient benefits of laparoscopy across the field of surgery, it is unlikely that a randomized controlled trial of laparoscopic versus open hepatic resection will ever be undertaken. Thus, comparisons will be limited to observational studies that are potentially confounded by selection bias. A strength of the

current study is the utilization of propensity score based matching, which simulates randomization in eliminating the confounding effects of variables used to create the model. In utilizing this design, the treatment effect of laparoscopy versus open resection in this study is more accurately estimated[49]. There remains the potential, however, that residual confounding by unmeasured variables continues to exist. Another limitation of this study is its relatively small sample size, with only 35 patients undergoing laparoscopic resection. Despite these drawbacks, it appears that laparoscopic resection is an effective and beneficial alternative to open resection for appropriately selected patients with hepatic colorectal metastases. Further study will be needed as experience increases to confirm these findings.

	Inclusive Cohort			Restricted Cohort		
	Open	Lap	P-value	Open	Lap	P-value
Age (years, mean, std)	61.3 (9.6)	62.3 (9.5)	0.629	62.7 (10.5)	62.3 (9.5)	0.858
CCI (median, IQR)	4 (1)	4 (1)	0.526	5 (1)	4 (1)	0.128
BMI (kg/m^2, mean, std)	28.6 (6.0)	28.1 (8.0)	0.865	27.7 (6.6)	28.1 (8.0)	0.956
Size of Largest Tumor (cm, mean, std)	4.7 (3.2)	4.2 (3.1)	0.474	3.7 (2.5)	4.2 (3.1)	0.438
Number of Lesions (median, IQR)	1 (1)	1(1)	0.162	1 (1)	1 (1)	0.648
CEA at time of Liver Resection (mean, std)	45.2 (69.1)	52.0 (126.4)	0.790	71.6 (134.3)	52.0 (126.4)	0.152
Node Positive Primary Tumor (n, %)	21 (60.0%)	20(57.1%)	0.808	23 (65.7%)	20 (57.1%)	0.467
Fong Score (median, IQR)	1 (1)	2(2)	0.306	2 (1)	2 (2)	0.860

Table 3.1: Baseline patient and tumor characteristics in the matched laparosopic and open cohorts. The inclusive cohort includes patients from 1995 through 2010, while the restricted cohort includes patients from 2002 through 2010. Continuous variables are presented as mean(standard deviation) or median(interquartile range) where appropriate, and categorical variables as count(percentage).
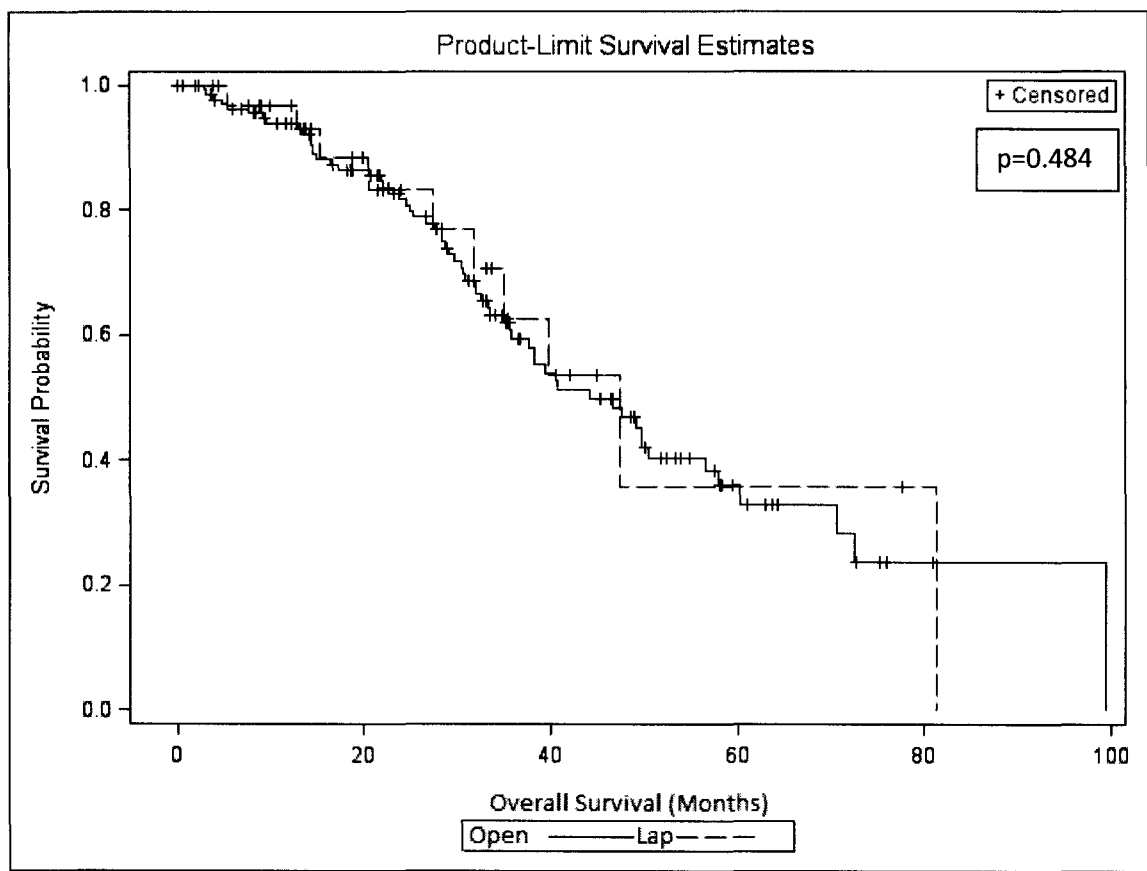
Figure 3.1: Overall Survival (in months) in the matched open and laparoscopic groups of the inclusive cohort.
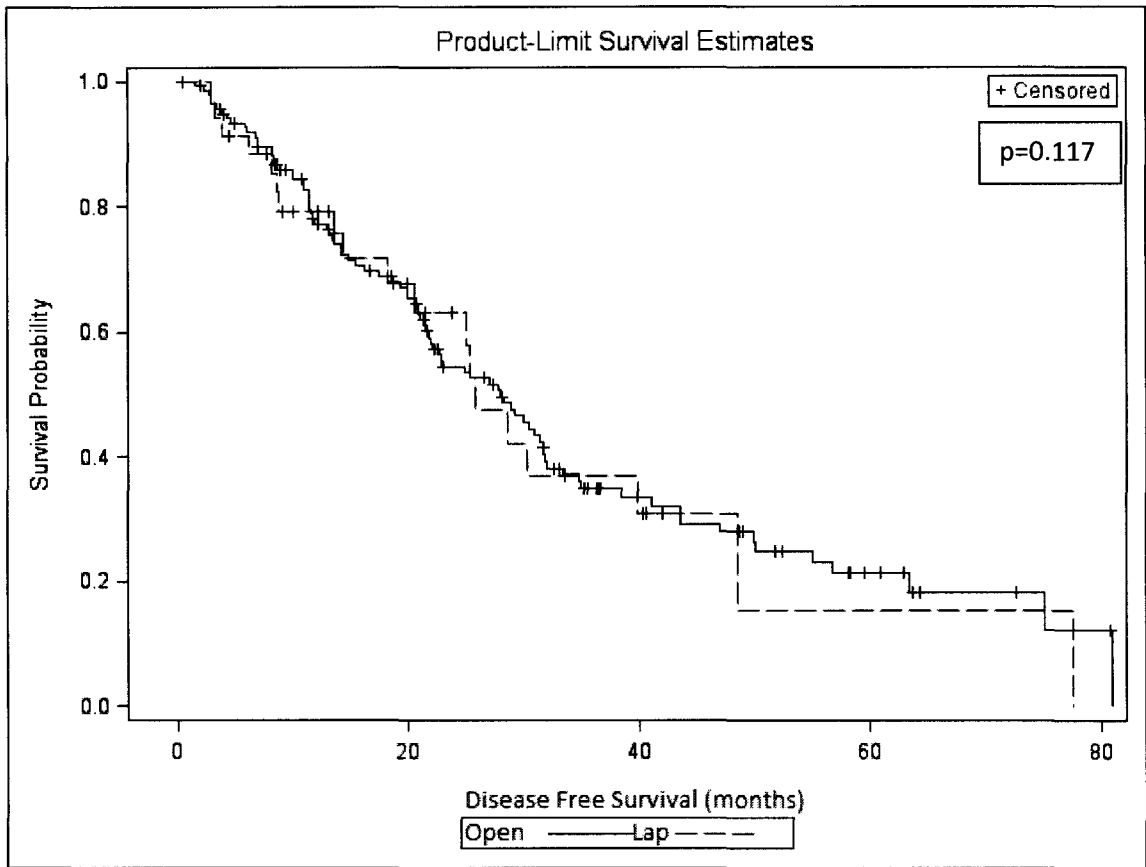
Figure 3.2: Disease free survival (in months) in the matched open and laparoscopic groups of the inclusive cohort.
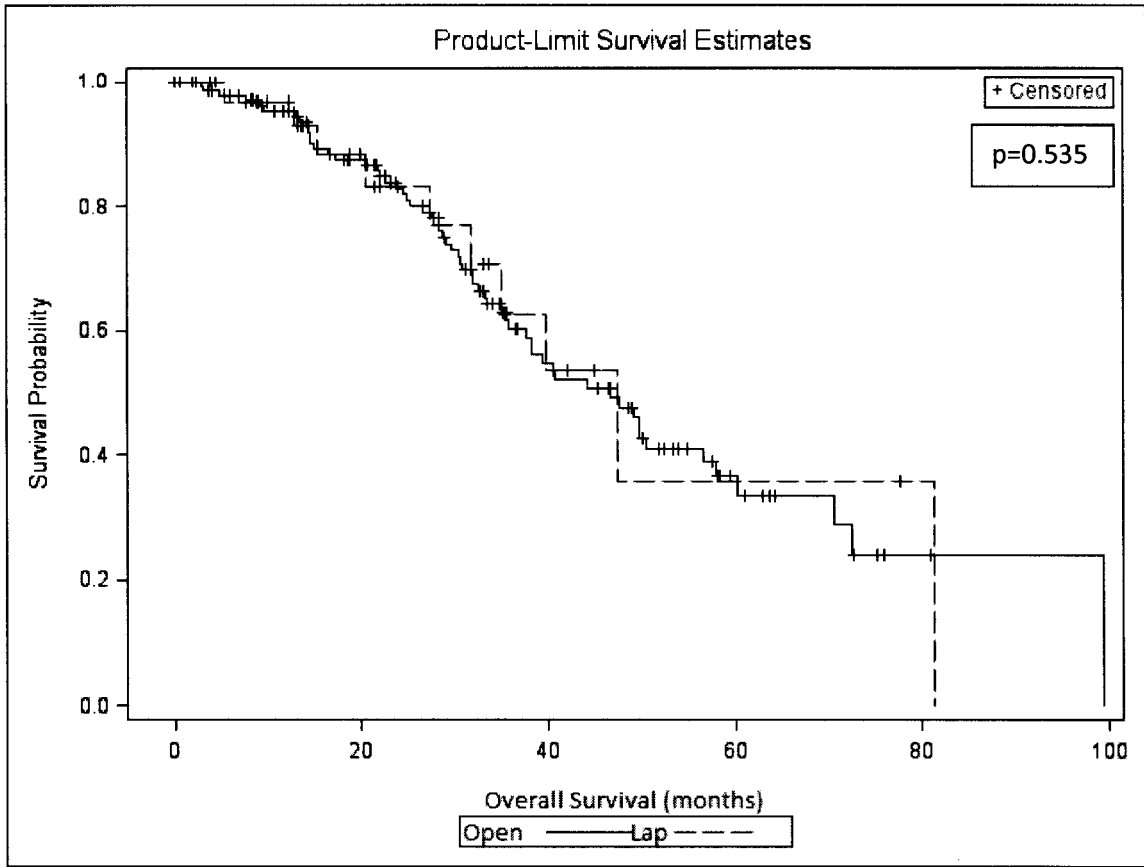
Figure 3.3: Overall Survival (in months) in the matched open and laparoscopic groups of the restricted cohort.
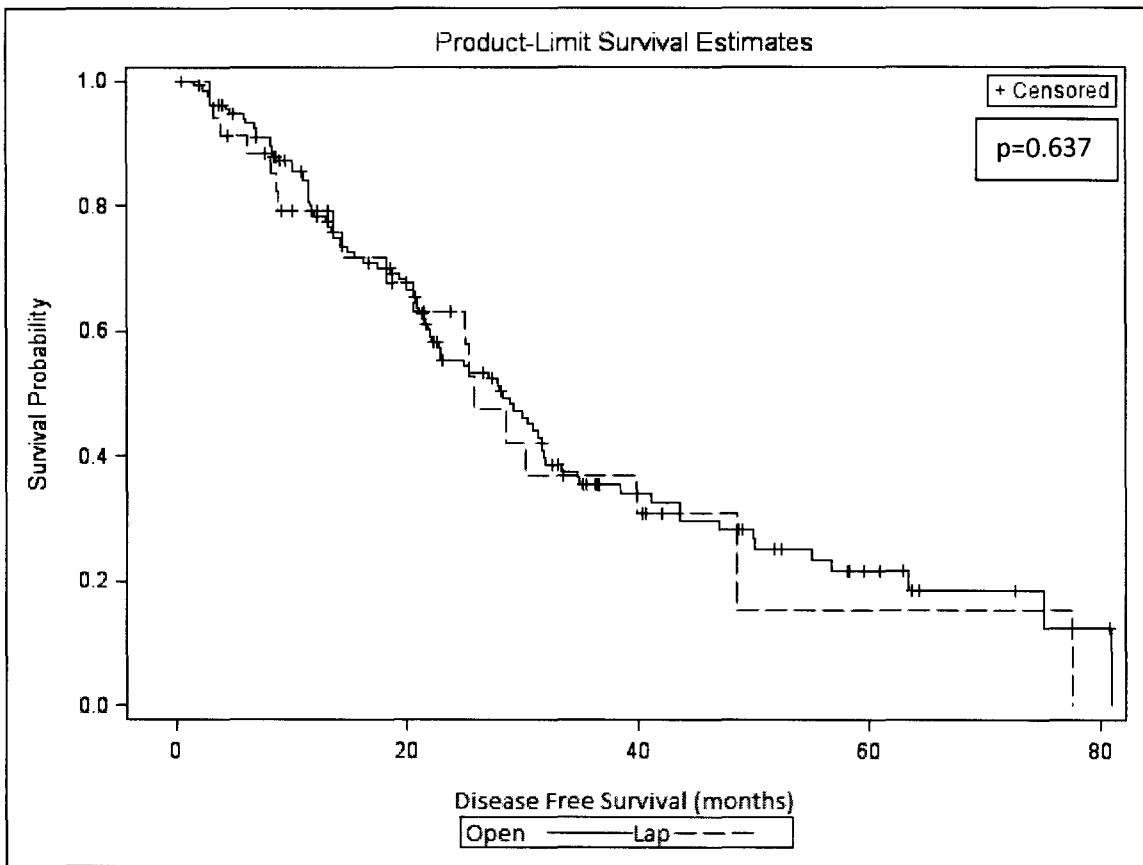
Figure 3.4: Disease free survival (in months) in the matched open and laparoscopic groups of the restricted cohort.

CHAPTER IV

CONCLUSIONS

Herein I have presented an overview of the utilization of propensity score matching to reduce confounding in observational studies in the field of surgery. Although the preeminent role of the randomized controlled trial for causal inference is acknowledged, the field of surgery is unique in the limitations placed on conducting such studies. On the one hand, the invasive nature of surgical procedures prevents proper blinding, in that it would be unethical to perform a sham operation on patients in order to have a control group. Another difficulty encountered is patient willingness to be randomized. This is especially problematic in the evaluation of minimally invasive procedures.

Often, minimally invasive procedures are developed prior to the conduct of any formal study of their efficacy. The problem encountered is that after a number of centers have established expertise in a particular minimally invasive procedure, patients will seek out those surgeons specifically so that they can have the less invasive procedure. Such patients are unlikely to consent to randomization that would potentially have them undergo a more invasive operation. This is the case with laparoscopic liver resection[26], which is why we are limited to observational studies such as the one presented in Chapter III of

this work. In other cases, the exposure to be studied is not modifiable by the investigators, such as patient race in the analysis presented in Chapter II.

Given the limitations on the performance of randomized controlled trials in the field of surgery, analysis based on the propensity score allows the investigator to create a "quasi-randomized" study from observational data, with subsequently reduced bias in estimation of the average treatment effect[6, 9]. Propensity score analysis is not a cure all for the problem of bias, however, and investigators and readers of such studies must keep some caveats in mind.

The principle limitation of propensity score analysis is that of the strongly ignorable treatment assignment assumption[6], which Austin has also named the requirement of "no unmeasured confounders"[7]. A propensity matched cohort represents selection of patients on observable covariates. Thus, any propensity analysis is subject to bias introduced by potential confounders that were not measured. This is a particular problem with retrospective studies such as those presented in Chapters II and III above. For example, patients are now selected for laparoscopic or open liver resection based on the surgeon's experience and perceived difficulty of the procedure. These two factors are not distinctly quantifiable. Though the analysis in Chapter III was confined to a single center, differences in surgeon experience are not likely to affect treatment selection. Percieved difficulty of the operation, on the other hand, could not be measured directly and thus controlled for. In this case, we are forced to use surrogates for difficulty, such as tumor size and number, and requirement for a major hepatectomy in the case of the analysis in Chapter III. It remains likely though, that these surrogates did not completely capture the perceived difficulty of the operation, thus leaving the potential for residual bias in the study.

Another requirement of the strong ignorability assumption is that there must be enough overlap in propensity scores between the treated and control groups. As noted in Chapter I, a simple nearest neighbor matching algorithm may lead to poor matches for treated subjects in the upper tail of the propensity score distribution. Imposition of an appropriate caliper on the maximum difference in propensity scores for a match to be valid has thus become a popular and important method of ensuring that the matched samples are limited to the region of common support[7, 18, 20].

Finally, a propensity score matching scheme is only as good as the balance attained between the treated and control groups on measured covariates. Thus, ensuring proper assessment of balance is critical. Although the most appropriate method for evaluating balance remains a controversial subject, it is critical for investigators to explicitly state the methodology used and results of such analysis in order for readers to properly evaluate the study. Currently, it appears that either the method of standardized difference suggested by Austin[20] or the use of hypothesis tests for paired data[2, 24] are both acceptable. With the above mentioned caveats, propensity matching is a valuable tool for creating nearly unbiased estimates of the treatment effect in cases where proper randomized studies are either infeasible or impossible.

REFERENCE LIST

(1) Friedman LM, Furberg CD, DeMets DL. Basic Study Design. In: Friedman LM, Furberg CD, DeMets DL, editors. *Fundamentals of Clinical Trials*. 4th ed. New York: Springer; 2010. 67.

(2) Luo Z, Gardiner JC, Bradley CJ. Applying Propensity Score Methods in Medical Research: Pitfalls and Prospects. *Medical Research and Review* 2010;67(5):528-554.

(3) Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med* 1997;127(8 Pt 2):757-763.

(4) Cochran WG. The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics* 1968;24(2):295-313.

(5) Cochran WG. The planning of observational studies of human populations. *Journalof the Royal Statistical Society, Series A* 1965;128:234-255.

(6) Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70(1):41-55.

(7) Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res* 2011;46(3):399-424.

(8) Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 1974;66:788-701.

(9) D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* 1998;17(19):2265-2281.

(10) Lee BK, Lessler J, Stuart EA. Improving propensity score weighting using machine learning. *Stat Med* 2010;29(3):337-346.

(11) McCaffrey DF, Ridgeway G, Morral AR. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychol Methods* 2004;9(4):403-425.

(12) Setoguchi S, Schneeweiss S, Brookhart MA, Glynn RJ, Cook EF. Evaluating uses of data mining techniques in propensity score estimation: a simulation study. *Pharmacoepidemiol Drug Saf* 2008;17(6):546-555.

(13) Austin PC, Grootendorst P, Anderson GM. A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a Monte Carlo study. *Stat Med* 2007;26(4):734-753.

(14) Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Sturmer T. Variable selection for propensity score models. *Am J Epidemiol* 2006;163(12):1149-1156.

(15) Westreich D, Cole SR, Funk MJ, Brookhart MA, Sturmer T. The role of the c-statistic in variable selection for propensity score models. *Pharmacoepidemiol Drug Saf* 2011;20(3):317-320.

(16) Schafer JL, Kang J. Average causal effects from nonrandomized studies: a practical guide and simulated example. *Psychol Methods* 2008;13(4):279-313.

(17) Austin PC, Mamdani MM. A comparison of propensity score methods: a case-study estimating the effectiveness of post-AMI statin use. *Stat Med* 2006;25(12):2084-2106.

(18) Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician* 1985;39:33-38.

(19) Cochran WG, Rubin DB. Controlling bias in observational studies: A review. *The Indian Journal of Statistics, Series A* 1973;35:417-466.

(20) Austin PC. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Stat Med* 2008;27(12):2037-2049.

(21) Gu XS, Rosenbaum PR. Comparison of multivariate matching methods: Structures, distances, and algorithms. *Journal of Computational and Graphical Statistics* 1993;2:405-420.

(22) Imai K, King G, Stuart EA. Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society, Series A* 2008;171:481-501.

(23) Normand ST, Landrum MB, Guadagnoli E et al. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. *J Clin Epidemiol* 2001;54(4):387-398.

(24) Hansen BB. The essential role of balance tests in propensity-matched observational studies: comments on 'A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003' by Peter Austin, Statistics in Medicine. *Stat Med* 2008;27(12):2050-2054.

(25) Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association* 1984;79:516-524.

(26) Buell JF, Cherqui D, Geller DA et al. The international position on laparoscopic liver surgery: The Louisville Statement, 2008. *Ann Surg* 2009;250(5):825-830.

(27) Wolfe RA, Ashby VB, Milford EL et al. Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant. *N Engl J Med* 1999;341(23):1725-1730.

(28) Epstein AM, Ayanian JZ, Keogh JH et al. Racial disparities in access to renal transplantation--clinically appropriate or due to underuse or overuse? *N Engl J Med* 2000;343(21):1537-44, 2.

(29) Isaacs RB, Lobo PI, Nock SL, Hanson JA, Ojo AO, Pruett TL. Racial disparities in access to simultaneous pancreas-kidney transplantation in the United States. *Am J Kidney Dis* 2000;36(3):526-533.

(30) Fan PY, Ashby VB, Fuller DS et al. Access and outcomes among minority transplant patients, 1999-2008, with a focus on determinants of kidney graft survival. *Am J Transplant* 2010;10(4 Pt 2):1090-1107.

(31) Butkus DE, Meydrech EF, Raju SS. Racial differences in the survival of cadaveric renal allografts. Overriding effects of HLA matching and socioeconomic factors. *N Engl J Med* 1992;327(12):840-845.

(32) Chakkera HA, O'Hare AM, Johansen KL et al. Influence of race on kidney transplant outcomes within and outside the Department of Veterans Affairs. *J Am Soc Nephrol* 2005;16(1):269-277.

(33) Eckhoff DE, Young CJ, Gaston RS et al. Racial disparities in renal allograft survival: a public health issue? *J Am Coll Surg* 2007;204(5):894-902.

(34) Foster CE, III, Philosophe B, Schweitzer EJ et al. A decade of experience with renal transplantation in African-Americans. *Ann Surg* 2002;236(6):794-804.

(35) Goldfarb-Rumyantzev AS, Sandhu GS, Baird B et al. Effect of education on racial disparities in access to kidney transplantation. *Clin Transplant* 2010.

(36) Israni AK, Li N, Sidhwani S et al. Association of hypertension genotypes and decline in renal function after kidney transplantation. *Transplantation* 2007;84(10):1240-1247.

(37) Leffell MS, Steinberg AG, Bias WB, Machan CH, Zachary AA. The distribution of HLA antigens and phenotypes among donors and patients in the UNOS registry. *Transplantation* 1994;58(10):1119-1130.

(38) Kerman RH, Kimball PM, Van Buren CT, Lewis RM, Kahan BD. Possible contribution of pretransplant immune responder status to renal allograft survival differences of black versus white recipients. *Transplantation* 1991;51(2):338-342.

(39) First MR, Schroeder TJ, Monaco AP, Simpson MA, Curtis JJ, Armenti VT. Cyclosporine bioavailability: dosing implications and impact on clinical outcomes in select transplantation subpopulations. *Clin Transplant* 1996;10(1 Pt 1):55-59.

(40) Swanson SJ, Hypolite IO, Agodoa LY et al. Effect of donor factors on early graft survival in adult cadaveric renal transplantation. *Am J Transplant* 2002;2(1):68-75.

(41) Ojo AO, Hanson JA, Meier-Kriesche H et al. Survival in recipients of marginal cadaveric donor kidneys compared with other recipients and wait-listed transplant candidates. *J Am Soc Nephrol* 2001;12(3):589-597.

(42) Brown KL, Doshi MD, Singh A et al. Does donor race still make a difference in deceased-donor African-American renal allograft recipients? *Am J Surg* 2010;199(3):305-309.

(43) Port FK, Bragg-Gresham JL, Metzger RA et al. Donor characteristics associated with reduced graft survival: an approach to expanding the pool of kidney donors. *Transplantation* 2002;74(9):1281-1286.

(44) Schold JD, Kaplan B, Chumbler NR et al. Access to quality: evaluation of the allocation of deceased donor kidneys for transplantation. *J Am Soc Nephrol* 2005;16(10):3121-3127.

(45) Rao PS, Schaubel DE, Guidinger MK et al. A comprehensive risk quantification score for deceased donor kidneys: the kidney donor risk index. *Transplantation* 2009;88(2):231-236.

(46) Breslow N. Analysis of survival data under the proportional hazards model. *International Statistics Review* 1975;43:45-58.

(47) Gray R. cmprsk: Subdistribution Analysis of Competing Risks. 2011.
Ref Type: Online Source

(48) Gray R. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *Annals of Statistics* 1988;16:1141-1154.

(49) Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70(1):41-55.

(50) Bergstralh E, Kosanke J. GMATCH Mayo Clinic Division of Biomedical Statistics and Informatics. 2003.
Ref Type: Generic

(51) Schempf AH, Kaufman JS. On the percent of excess risk explained. *J Epidemiol Community Health* 2011;65(2):190.

(52) Alexander GC, Sehgal AR. Barriers to cadaveric renal transplantation among blacks, women, and the poor. *JAMA* 1998;280(13):1148-1152.

(53) Callender CO, Hall MB, Branch D. An assessment of the effectiveness of the Mottep model for increasing donation rates and preventing the need for transplantation-- adult findings: program years 1998 and 1999. *Semin Nephrol* 2001;21(4):419-428.

(54) Grams ME, Womer KL, Ugarte RM, Desai NM, Montgomery RA, Segev DL. Listing for expanded criteria donor kidneys in older adults and those with predicted benefit. *Am J Transplant* 2010;10(4):802-809.

(55) Lefor AT, Flowers JL. Laparoscopic wedge biopsy of the liver. *J Am Coll Surg* 1994;178(3):307-308.

(56) Cherqui D, Husson E, Hammoud R et al. Laparoscopic liver resections: a feasibility study in 30 patients. *Ann Surg* 2000;232(6):753-762.

(57) Nguyen KT, Gamblin TC, Geller DA. World review of laparoscopic liver resection-2,804 patients. *Ann Surg* 2009;250(5):831-841.

(58) Koffron AJ, Auffenberg G, Kung R, Abecassis M. Evaluation of 300 minimally invasive liver resections at a single institution: less is more. *Ann Surg* 2007;246(3):385-392.

(59) Cannon RM, Brock GN, Marvin MR, Buell JF. Laparoscopic Liver Resection: An Examination of Our First 300 Patients. *J Am Coll Surg* 2011;213(4):501-507.

(60) Ardito F, Tayar C, Laurent A, Karoui M, Loriau J, Cherqui D. Laparoscopic liver resection for benign disease. *Arch Surg* 2007;142(12):1188-1193.

(61) Gagner M, Rogula T, Selzer D. Laparoscopic Liver Resection: Benefits and Controversies. *Surg Clin North Am* 2004;84:451-462.

(62) Koopmann MC, Heise CP. Laparoscopic and minimally invasive resection of malignant colorectal disease. *Surg Clin North Am* 2008;88(5):1047-72, vii.

(63) Martin RC, Scoggins CR, McMasters KM. Laparoscopic hepatic lobectomy: advantages of a minimally invasive approach. *J Am Coll Surg* 2010;210(5):627-6.

(64) Abu HM, Underwood T, Zuccaro M, Primrose J, Pearce N. Short- and medium-term results of totally laparoscopic resection for colorectal liver metastases. *Br J Surg* 2010;97(6):927-933.

(65) Mala T, Edwin B, Gladhaug I et al. A comparative study of the short-term outcome following open and laparoscopic liver resection of colorectal metastases. *Surg Endosc* 2002;16(7):1059-1063.

(66) Castaing D, Vibert E, Ricca L, Azoulay D, Adam R, Gayet B. Oncologic results of laparoscopic versus open hepatectomy for colorectal liver metastases in two specialized centers. *Ann Surg* 2009;250(5):849-855.

(67) Fong Y, Fortner J, Sun RL, Brennan MF, Blumgart LH. Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases. *Ann Surg* 1999;230(3):309-318.

(68) Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40(5):373-383.

(69) House MG, Ito H, Gonen M et al. Survival after hepatic resection for metastatic colorectal cancer: trends in outcomes for 1,600 patients during two decades at a single institution. *J Am Coll Surg* 2010;210(5):744-745.

(70) Choti MA, Sitzmann JV, Tiburi MF et al. Trends in long-term survival following liver resection for hepatic colorectal metastases. *Ann Surg* 2002;235(6):759-766.

(71) Johnstone PA, Rohde DC, Swartz SE, Fetter JE, Wexner SD. Port site recurrences after laparoscopic and thoracoscopic procedures in malignancy. *J Clin Oncol* 1996;14(6):1950-1956.

(72) Paolucci V, Schaeff B, Schneider M, Gutt C. Tumor seeding following laparoscopy: international survey. *World J Surg* 1999;23(10):989-995.

(73) Ishida H, Murata N, Yamada H et al. Effect of CO_2 pneumoperitoneum on growth of liver micrometastases in a rabbit model. *World J Surg* 2000;24(8):1004-1008.

(74) Gutt CN, Riemer V, Kim ZG, Erceg J, Lorenz M. Impact of laparoscopic surgery on experimental hepatic metastases. *Br J Surg* 2001;88(3):371-375.

(75) Hsu TC. Intra-abdominal lesions could be missed by inadequate laparoscopy. *Am Surg* 2008;74(9):824-826.

(76) Nguyen KT, Gamblin TC, Geller DA. World review of laparoscopic liver resection-2,804 patients. *Ann Surg* 2009;250(5):831-841.

(77) Lacy AM, Garcia-Valdecasas JC, Delgado S et al. Laparoscopy-assisted colectomy versus open colectomy for treatment of non-metastatic colon cancer: a randomised trial. *Lancet* 2002;359(9325):2224-2229.

(78) Martin RC, Augenstein V, Reuter NP, Scoggins CR, McMasters KM. Simultaneous versus staged resection for synchronous colorectal cancer liver metastases. *J Am Coll Surg* 2009;208(5):842-850.

CURRICULUM VITAE

Robert Martyn Cannon, MD

Biographical Inforamation

Date of Birth: January 13, 1983

Place of Birth: Atlanta, Ga

Wife: Lindsley Brooke Cannon

Children: William Lovick Cannon

Home Address: 1827 Tyler Parkway Louisville, Ky 40204

Phone: (404) 405-9329

e-Mail: rmcann03@louisville.edu

<u>EDUCATION</u>

-Master of Science, Biostatistics

University of Louisville, Louisville, Ky

Expected Graduation Date May 2012

-M.D.

Medical College of Georgia, Augusta, GA (May, 2008)

-Bachelor of Science, Biology

University of Georgia Honors Program, Athens, GA (May, 2004)

-High School Diploma

The Walker School, Salutatorian, Marietta, GA (May, 2000)

HONORS AND AWARDS

-National Merit Scholarship

-Phi Eta Sigma Freshman Honor Society

-Golden Key International Honor Society

-National Society of Collegiate Scholars

-Phi Kappa Phi Honor Society

-Phi Beta Kappa

-Medical College of Georgia Faculty Scholarship

-National Association of Pathology Chairs Pathology Honor Society

-AOA, inducted in spring 2007 as Junior Medical Student

-William B Dion Outstanding Student in Surgery Award

-Kentucky Chapter American College of Surgeons Outstanding Resident Paper 2009

-Kentucky Chapter American College of Surgeons Resident Paper Competition: Second Place 2010, 2011

-Vascular Surgery Award, University of Louisville Dept of Surgery. 2010

-Research!Louisville Outstanding Resident Poster. 2011

Publications

Published or Accepted for Publication

1. **Cannon R.**, Shah, M. Suydam, E., Gucwa A., Bowden T., Holsten S. Early Thrombosis of the Superior Vena Cava in a Patient with a Central Venous Catheter and Adenocarcinoma of the Ampulla of Vater. American Surgeon 2008;74:1195

2. Buell JF, Tranchart H, **Cannon R**, Dagher I. Management of Benign Hepatic Tumors. Surg Clin North Am. 2010;90:719

3. Paramesh A, **Cannon R**, Buell JF. Malignancies in Pediatric Solid Organ Transplant Recipients: Epidemiology, Risk Factors, and Prophylactic Approaches. Curr Opin Organ Transplant. 2010;15:621

4. **Cannon RM**, Killackey MT, Buell JF. Laparoscopic Hepatectomy for Colorectal Metastases: Ready for Prime Time?. Current Colorectal Cancer Reports. 2011;7:161

5. Shafaee Z, Kazaryan AM, Marvin MR, **Cannon R**, Buell JF, Edwin B, Gayet B. Is Laparoscopic Repeat Hepatectomy Feasible? A Tri-institutional Analysis. J Am Coll Surg. 2011;212(2):171-9

6. **Cannon RM**, Eng M, Marvin MR, Buell JF. Laparoscopic Living Kidney Donation at a Single Center: An Examination of Donor Outcomes with Increasing Experience. The American Surgeon. 2011; 77(7):911-915

7. **Cannon RM**, et al. Laparoscopic Liver Resection: An Examination of Our First 300 Patients. J Am Coll Surg. 2011;213(4):501-7

8. **Cannon RM**, et al. Safety and Efficacy of Hepatectomy for Colorectal Metastases in the Elderly. Journal of Surgical Oncology. 2011; 104(7): 804

9. **Cannon RM**, et al. A Novel Classification System to Address Financial Impact and Referral Decisions for Bile Duct Injury in Laparoscopic Cholecystectomy. HPB Surgery, vol 2011, Article ID 371245

10. **Cannon RM**, et al. Flail Chest Injury: Are We Making Any Progress? The American Surgeon. Accepted for Publication

11. **Cannon RM**, et al. Analysis of BK Viral Infection After Alemtuzumab Induction For Renal Transplant. Transplant Infectious Disease. Accepted for Publication

12. Franklin G, **Cannon R**, et al. Impact of Withdrawal of Care and Futile Care on Trauma Mortality. Surgery. 2011; 150(4):854.

13. **Cannon RM**, et al. A Resident Perspective on Increasing Duty Hour Limitations. Am J Surg. Letter to the Editor. Accepted for Publication

14. **Cannon RM**, et al. BK Viral Disease in Renal Transplantation. Curr Opin Organ Transplant. 2011; 16(6): 576

15. **Cannon RM**, et al. Surgical Resection for Hilar Cholangiocarcinoma: Experience Improves Resectability. HPB. 2012; 14(2):142

16. **Cannon RM**, et al. Management of Diffuse Hepatocellular Cancer (≥ 10 Lesions) with Doxorubicin Loaded DC Bead is Safe and Effective Treatment Option. Onkologie. Accepted for Publication

17. **Cannon RM**, et al. A Multi-Institutional Analysis of Pancreatic Adenocarcinoma Demonstrating the Effect of Diabetes Status on Survival After Resection. HPB. Accepted for Publication.

18. **Cannon RM**, et al. The Contribution of Donor Quality to Differential Graft Survival in African American and Caucasian Renal Transplant Recipients. American Journal of Transplantation. Accepted for Publication.

19. **Cannon RM, et al**. Traditional Open Repair of Blunt Thoracic Aortic Injury Remains Relevant in the Endovascular Era. Journal of the American College of Surgeons. Accepted for Publication.

20. Paramesh AS, Davis JY, Mallikarjun Chaitanya, Zhang R, **Cannon R**, et al. Kidney Transplant Alone in Hepatitis C Cirrhotic ESRD Patients. Transplantation. Accepted for Publication.

21. Bozeman MC, **Cannon RM**, et al. Utility of CT Findings and Contrast Extravasation in Predicting the Need for Embolization With Pelvic Fractures. The American Surgeon. Accepted for Publication.

<u>Under Review</u>

1. **Cannon RM**, et al. The Impact of Recipient Obesity on Outcomes Following Renal Transplantation. Annals of Surgery

2. **Cannon RM**, et al. Negative Effects of Transfused Blood Components Following Hepatectomy for Metastatic Colorectal Cancer. The American Surgeon

3. **Cannon RM**, et al. Laparoscopic Versus Open Resection of Hepatic Colorectal Metastases. Surgery

4. Killackey M, **Cannon RM**, et al. Safety and Feasibility of Laparoscopic Caudate Resection. Surgery

5. St Hill CR, Mbah NA, **Cannon RM**, et al. Laparoscopic versus Open Hepatic Resection for Hepatocellular Carcinoma: Improvement in Outcomes and Similar Cost. Annals of Surgical Oncology

6. Buell JF, Gayet B, Ho-Seong H, Wakabayashi G, Kim KH, Belli G, **Cannon RM**, et al. Evaluation of Stapler Hepatectomy During Laparoscopic Liver Resection. HPB

7. **Cannon RM**, et al. Machine Perfusion Reduces Delayed Graft Function After Deceased Donor Liver Transplantation. NEJM

8. **Cannon RM**, et al. Right Versus Left Kidneys for Live Donor Transplantation: Is there Really a Difference? AJT

9. **Cannon RM**, et al. A Review of the United States Experience with Combined Heart-Liver Transplantation. Transplant International

10. **Cannon RM**, et al. Safety and Efficacy of Irreversible Electroporation for Hepatic Tumors in Proximity to Vital Structures. HPB

In Preparation

1. **Cannon RM**, et al. Clinical and Financial Analysis of Laparoscopic Versus Open Hepatic Resection.

2. **Cannon RM**, et al. Advances In Laparoscopic Living Kidney Donation Surgery. Invited Review for European Nephrology

Book Chapters

1. Buell JF, Gayet B, Lei H, Tzanis D, **Cannon RM**, Dagher I. Principles of Laparoscopic Liver Resection. In *Principles of Hepatic Surgery*. Vicinius Gava, ed. In Press.

2. **Cannon RM**, Buell JF. Laparoscopic Liver Resection. In *Updated Topics in Minimally Invasive Abdominal Surgery*. Ahmed AbdelRaouf ElGeide ed. 2011 Intech, Rijeka, Croatia.

Abstracts and Presentations

1. Siler D., Wynn J., Mulloy L., Jagadeesan M., Merchen T., **Cannon R.**, Vasil D., Williams D. Two Dose Daclizumab Produces Rejection Rates and Short Term Graft Survival Equivalent to Basiliximab in Kidney Transplantation. Poster Presentation. 2008 American Transplant Congress

2. **Cannon R**, et al. Surgical Resection for Hilar Cholangiocarcinoma (Klatskin's Tumor): Palliative or Curative. Oral Presentation. 2010 SSO Cancer Symposium

3. **Cannon R**, et al. Laparoscopic Liver Resection: Evolution of a Procedure in 300 Patients. Oral Presentation. 2010 SSO Annual Cancer Symposium

4. Hibbs J, Hamdallah O, **Cannon R**, et al. Alemtuzumab Induction with Rapid Corticosteroid Elimination is Associated with Comparable Outcomes in High Immunologic Risk and Non High Risk Renal Transplant Recipients. Poster Presentation. 2010 American Transplant Congress

5. Buell J, **Cannon RM**, et al. Alemtuzumab Provides Superior Outcomes Under Rapid Steroid Elimination. Poster Presentation. 2010 American Transplant Congress

6. **Cannon RM**, et al. Analysis of BK Virus Infection After Alemtuzumab Induction for Renal Transplant. Poster Presentation. 2010 American Transplant Congress

7. **Cannon R**, et al. Laparoscopic Hepatic Lobar Resection: Standard of Care? Oral Presentation. 2010 International Hepatopancreatobiliary Association

8. **Cannon RM**, Eng M, Marvin MR, Buell JF. Living Kidney Donation at a Single Center: An Examination of Donor Outcomes with Increasing Experience. 2011 Southeastern Surgical Congress Residents Forum

9. **Cannon RM**, et al. Continuous Flow Ventricular Assist Devices Are Associated With Increased Gastrointestinal Hemorrhage. Quick Shot Presentation, 2011 Southwestern Surgical Congress

10. **Cannon RM**, et al. Safety and Efficacy of Hepatectomy for Colorectal Metastases in the Elderly. Poster Presentation, 2011 SSAT Annual Meeting

11. **Cannon RM**, et al. Traditional Open Repair of Blunt Aortic Transection Remains Relevant in the Endovascular Era. ACS Surgical Forum 2011

12. **Cannon RM**, et al. Upregulation of mir-122 is a Significant anti-HCC Mechanism of Curcumin. Research!Louisville 2011

13. **Cannon RM,** et al. Negative Effects of Transfused Blood Components Following Hepatectomy for Metastatic Colorectal Cancer. Poster Presentation. 2012 Southeastern Surgical Congress Annual Meeting

14. Bozeman MC, **Cannon RM**, et al. Utility of CT Findings and Contrast Extravasation in Predicting the Need for Embolization With Pelvic Fractures. Resident Forum, 2012 Southeastern Surgical Congress Annual Meeting

15. **Cannon RM,** et al. A Review of the United States Experience with Combined Heart-Liver Transplantation. Quick Shot Oral Presentation. 2012 Academic Surgical Congress.

16. **Cannon RM**, et al. Safety and Feasibility of Laparoscopic Caudate Resection. Oral Presentation. 2012 Central Surgical Association Annual Meeting.

17. **Cannon RM**, et al. Laparoscopic Versus Open Resection of Hepatic Colorectal Metastases. Oral Presentation. 2012 Central Surgical Association Annual Meeting.

18. **Cannon RM**, et al. Financial Analysis of Laparoscopic Versus Open Liver Resection. Oral Presentation. 2012 Society for Surgical Oncology Annual Cancer Symposium.

19. **Cannon RM**, et al. Safety and Efficacy of Irreversible Electroporation for Hepatic Tumors in Proximity to Vital Structures. Oral Presentation. 2012 Americas Hepatopancreatobiliary Association (AHPBA) Annual Meeting.

20. **Cannon RM**, et al. Pancreatic Fistula After Resection of Pancreatic Ductal Adenocarcinoma Does Not Adversely Affect Disease Free Survival. Poster Presentation. 2012 Americas Hepatopancreatobiliary Association (AHPBA) Annual Meeting.

Journals Reviewed For

-Annals of Transplantation

-Transplant International

Teaching

1. ACLS Instructor 2010-

2. Senior Medical Student Honors Surgery Anatomy Course July 15,2010

Society Memberships

1. Southeastern Surgical Society

2. Association for Academic Surgery

3. Americas Hepatopancreatobiliary Association