UNOS DONATE LIFE
UNITED NETWORK FOR ORGAN SHARING

Since 1984 — sharing organs, sharing data, sharing life.

700 North 4th Street, Richmond, VA 23219
P.O. Box 2484, Richmond, VA 23218
tel: 804-782-4800
fax: 804-782-4817
www.unos.org

August 20, 2008

John P. Roberts, MD
ASTS President
2461 South Clark Street
Suite 640,
Arlington, VA 22202

Dear Dr. Roberts,

Attached to this letter, please find responses to the questions posed by the ASTS regarding possible revisions to the kidney allocation system (letter dated February 12, 2008). These responses were developed and reviewed by a subcommittee of the OPTN/UNOS Kidney Transplantation Committee with assistance from the Scientific Registry of Transplant Recipients (SRTR) and the Organ Procurement and Transplantation Network (OPTN) staff.

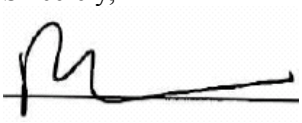
When your document was forwarded to the committee, the concept of a kidney allocation score (KAS) was still in development. Many of the queries were directed at a system driven by LYFT alone. In fact, the Kidney Allocation Score (KAS) will be calculated by 3 factors, including; 1) LYFT, 2) donor profile index (DPI), and 3) Time on dialysis. We hope to follow this up with a concept document as a Request for Information (RFI) (see below), which will provide updated comprehensive description of how the kidney allocation score is calculated.

As a next step in the policy development process, the Committee intends to release a Request for Information (RFI) later this fall. The purpose of the RFI is to continue to gather feedback from the transplantation community and general public. The responses to the RFI will be discussed at a public forum in the first quarter of 2009 and we hope the ASTS will be represented at this forum.

The Committee appreciates the time and effort expended by the ASTS to thoughtfully craft its questions. We believe that the attached responses will lead to a continued dialogue with the ASTS regarding the development of a proposal for a new kidney allocation system.

If you have any questions regarding the contents of the attached document, please do not hesitate to contact me at stockp@surgery.ucsf.edu. Ciara Gould, UNOS Staff Liaison to the Committee is also available to provide any additional information you may need. Ms. Gould may be reached at 804-782-4800 or gouldcij@unos.org.

Sincerely,



Peter G. Stock, MD, PhD
Chair, OPTN/UNOS Kidney Transplantation Committee

cc: Robert S. Higgins, MD, MSHA, President OPTN/UNOS
Walter K. Graham, Executive Director, UNOS
Cliff McClenney, ASTS Liaison, UNOS

Response to the ASTS Letter dated February

The following document is a point by point response to each of the queries posed by the ASTS. Many of the questions that were posed by the ASTS were based on a system driven principally by LYFT. Since that time, the concepts under consideration have evolved to include a mechanism for candidates to gain priority over time (time on dialysis [DT]) and a mechanism for better classification of donors (the donor profile index [DPI]). These three elements, (DT, DPI, and LYFT) along with a measure of candidate sensitization (CPRA) are combined into a kidney allocation score (KAS).

Figures 1 and 2 illustrate the proportional and relative impact of a candidate's LYFT score and the candidate's DT based on the DPI for a specific donor's kidney. Each candidate would receive a KAS that includes the LYFT calculation (the solid line) and DT (the dashed line). For the purposes of this illustration, these examples assume that the candidate has a sensitization level of 0%. As a donor kidney becomes available, it would receive a donor profile index (DPI) score. Each candidate would then receive a KAS based on a combination of his or her LYFT score and DT. The proportion of LYFT and DT is determined by the DPI score.

For example, when a donor kidney from the 39th DPI percentile becomes available, each candidate would receive an allocation weight comprised of 50% LYFT score, and 50% time on dialysis. For a donor kidney from the 20% percentile, each candidate would receive an allocation weight of 36% of time on dialysis and 64% of LYFT. For sensitized candidates, the appropriate points for sensitization would be added into the KAS.

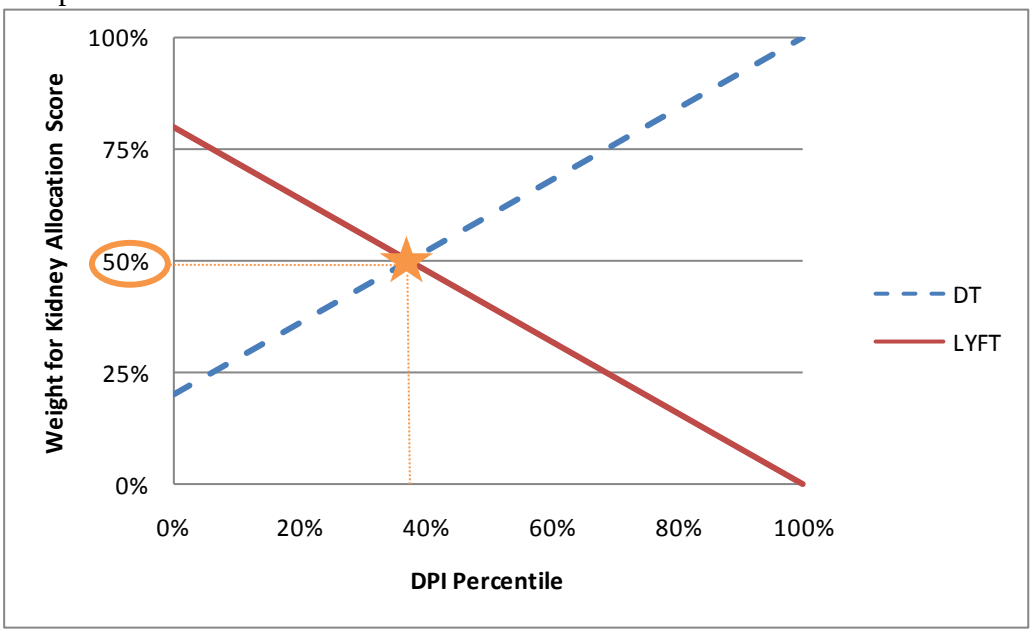


Figure 1: Interaction between time on dialysis (DT), life years from transplant (LYFT), and donor profile index (DPI) in the kidney allocation score (KAS). In this example, a donor from the 39^{h} percentile is available. Candidate KAS scores will be comprised of 50% LYFT and 50% time on dialysis.

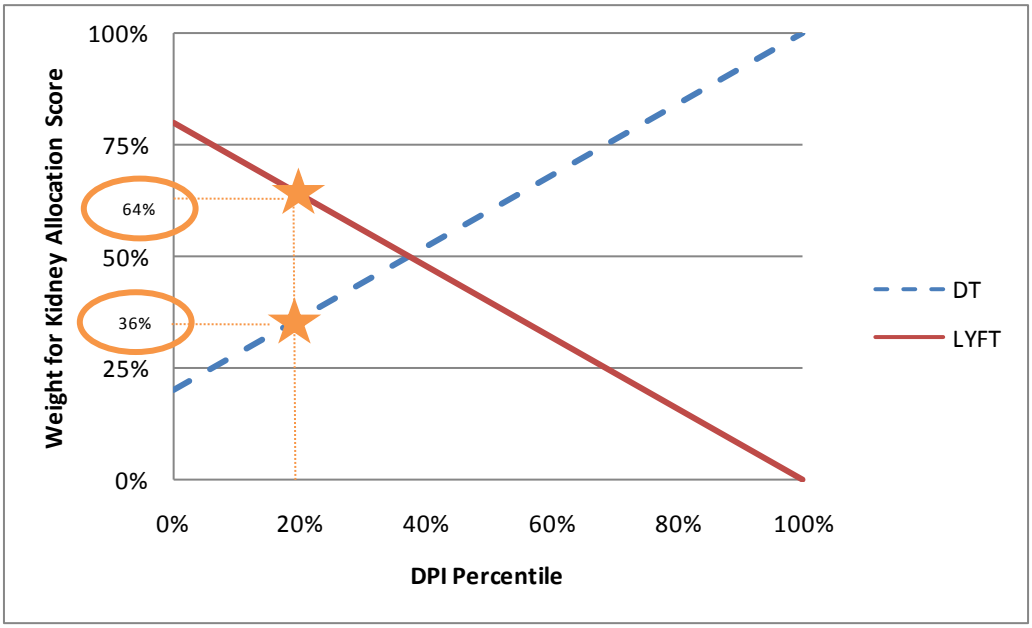


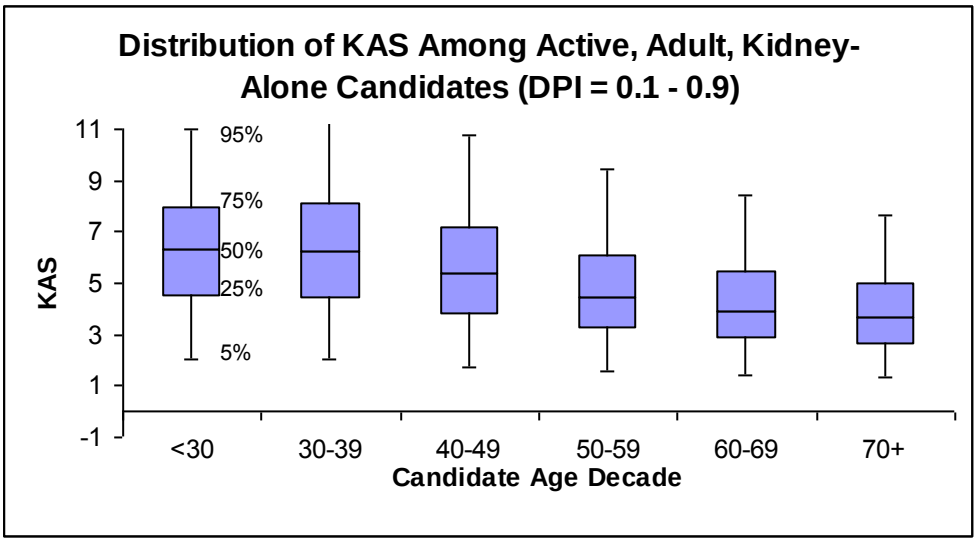
Figure 2: Interaction between time on dialysis (DT), life years from transplant (LYFT), and donor profile index (DPI) in the kidney allocation score (KAS) In this example, a donor from the 20th percentile is available. Candidate KAS scores will be comprised of 64% LYFT and 36% time on dialysis.

Responses to Specific Questions

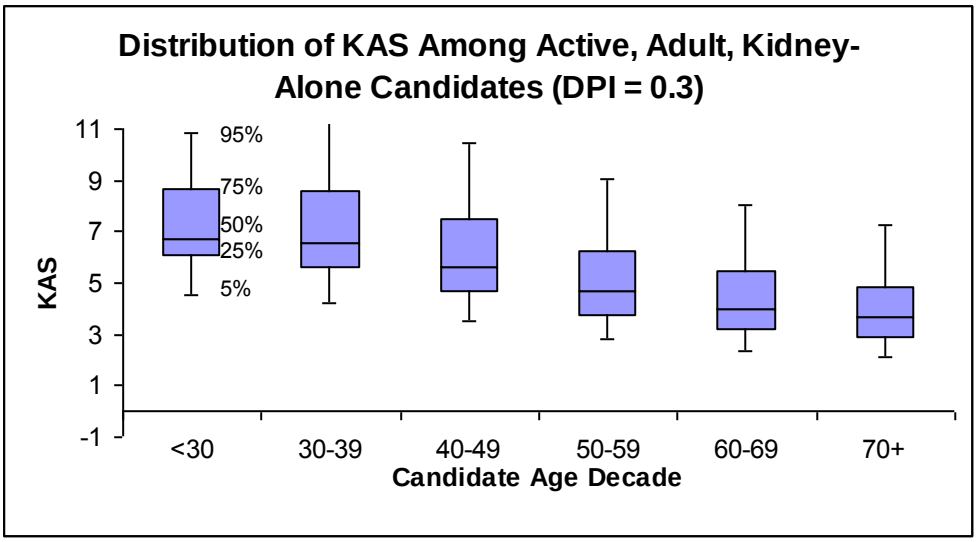
1. LYFT Score, Modeling:
    a. To provide interpretable information for the public, it would be helpful to provide some graphical representation of the overlap of the population by age decile (i.e. what percentage of the 60 year olds would get the same benefit as the population that is 5 years younger?).

Answer:

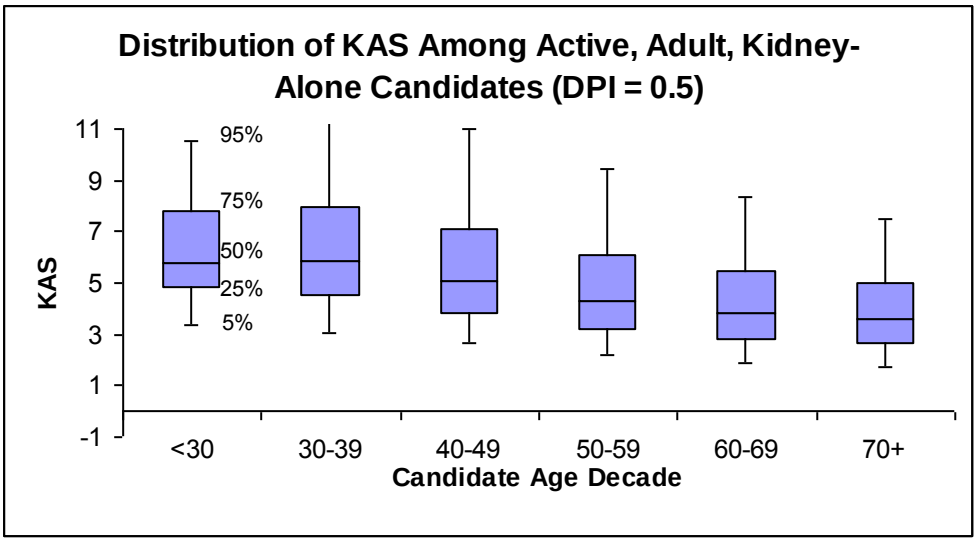
The Kidney Allocation Score (KAS) incorporates three recipient characteristics including LYFT, dialysis years, and CPRA; in addition, the potential lifetime of donor kidneys, as measured by DPI, is used to determine the relative contributions of these factors to the final score. For example, LYFT carries more weight for donor kidneys with longer potential lifetimes, and dialysis years carries more weight for donor kidneys with shorter potential lifetimes. The following box plots show, by decile of age, the distribution of each of these components, illustrating the degree of overlap and the trend (if any) in each by age.



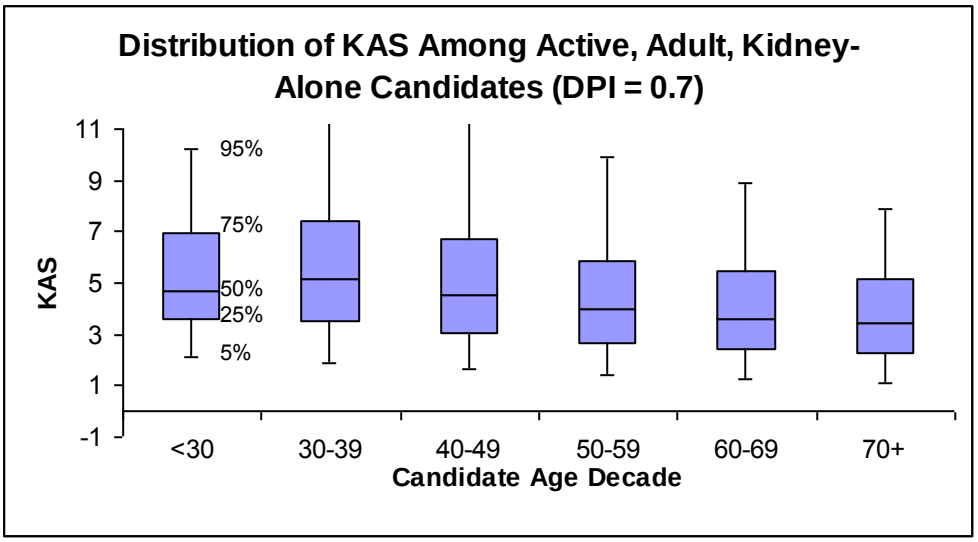
Using a wide range of DPI kidneys, the overall distribution of KAS by age-decade shows a trend downwards, but there is a great deal of overlap among kidney allocation scores of consecutive age groups and even between the <30 and the 70+ groups.



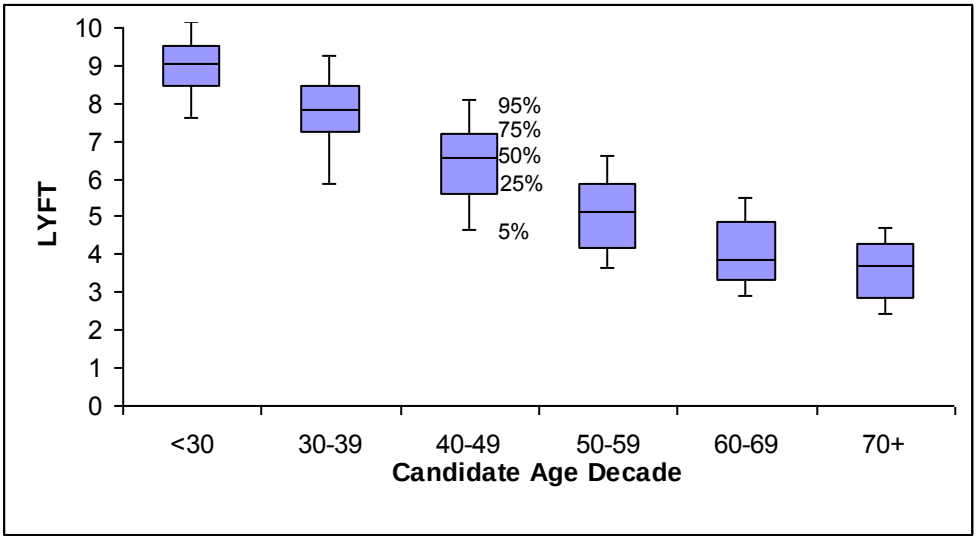
For DPI = 0.3 organs alone, there is more of a trend in KAS with age, although the overlap between adjacent decades remains large.



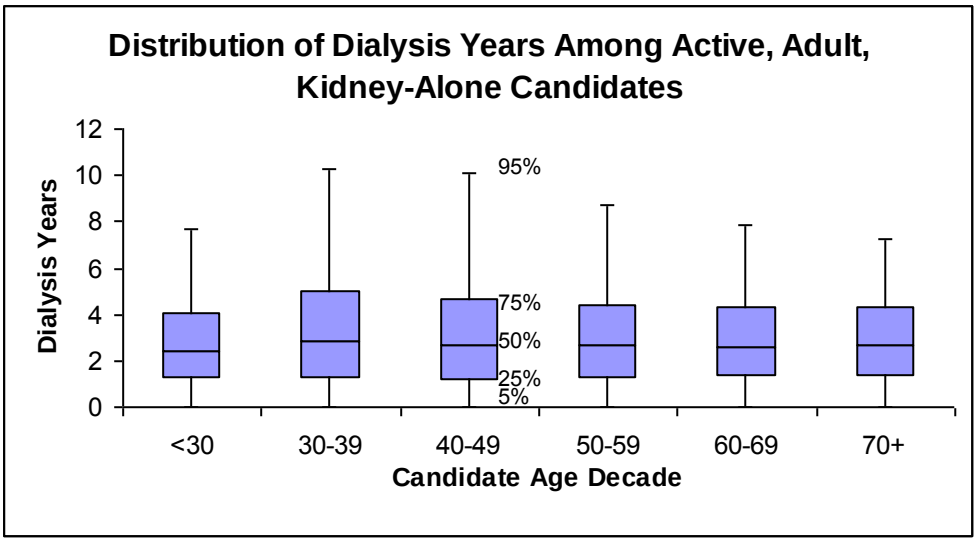
For the median kidney, there is still a trend in KAS with age, although there is substantial overlap even between candidates with extreme ages (<30 and 70+).



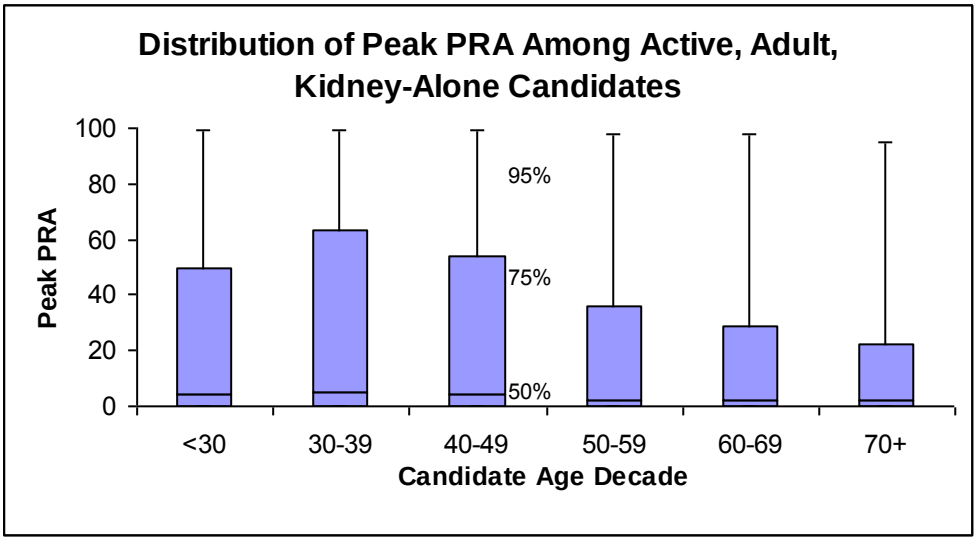
For kidneys with DPI = 0.7, there is very little trend in KAS with age. Over 25% of candidates over the age of 70 have KAS scores exceeding the median KAS among candidates <30.



The reason KAS has a trend with age is that the LYFT (a component of the KAS) varies with age.



The overall distributions of dialysis years (a component of the KAS) tends not to vary much with age.



The median PRA (a component of the KAS) is < 10 for all age groups.

b. The results of the simulations demonstrate that while life years are increased, graft years are not affected. How much of the life years are coming in patients who have returned to dialysis after a failed transplant? What would be the effect of discounting these life years in this population by 0.6 rather than 0.8 as the quality of life after return to dialysis is arguably much worse?

Answer:

The following table gives the additional years of life, including years spent with functioning graft, resulting from the current national allocation

system and systems that use LYFT to varying degrees. Results obtained from the Kidney Pancreas Simulated Allocation System (KPSAM).[1]

Years	Simulated Current National Allocation	LYFT for SCD, Wait Time for ECD	LYFT (0-100%) v. DY by Continuous DPI	LYFT (0-80%) v. DY by Continuous DPI
Number of candidates (on waitlist at start or joining during run)	80,549	80,549	80,549	80,549
Number of transplant recipients	9124	9616	9111	9035
Total lifespan after transplant	107865	143062	122140	118133
Total graft years of life	72814	86423	75600	73772
Total extra years	48187	59770	52947	51589
Increase in lifespan after transplant		35,018	14,275	10,268
Increase in graft years of life		13,105	2,786	958
Increase in extra years		11,346	4,760	3,402
Lifespan benefit per transplant	5.28	6.22	5.81	5.71

"Simulated Current National Allocation" simulates the national allocation rules, without any local variances or alternative allocation units. This is the baseline for all comparisons in the above tables.

"LYFT for SCD, Wait Time for ECD" uses LYFT to allocate SCD organs and waiting time to allocate ECD organs. Pediatric allocation is kept the same as in the current allocation system.

"LYFT (0-100%) v. DY by Continuous DPI" allocates all organs (SCD, DCD, and ECD) by a formula weighting LYFT and DY. For the lowest DPI kidneys, allocation is 100% LYFT and 0% DY. For the highest DPI kidneys, allocation is 100% DY and 0% LYFT. The weighting is linearly scaled between these two endpoints. Pediatric allocation is kept the same as in the current allocation system.

"LYFT (0-80%) v. DY by Continuous DPI" allocates all organs (SCD, DCD, and ECD) by a formula weighting LYFT and DY. For the lowest DPI kidneys, allocation is weighted 80% LYFT and 20% DY. For the highest DPI kidneys, allocation is 100% DY and 0% LYFT. The weighting is linearly scaled between these two endpoints. Pediatric allocation is kept the same as in the current allocation system.

Allocation using LYFT can contribute as much as 13,105 extra years with functioning graft or as few as 958, depending on how LYFT is incorporated into the allocation system. This estimate showed the simulated long term results of a single year of transplants, so the effects of repeat transplants after the simulated year of allocation were not included.

[1] Simulating the Allocation of Organs for Transplantation. David Thompson; Larry Waisanen; Robert Wolfe; Robert M. Merion; Keith McCullough; Ann Rodgers. Health Care Management Science 7(4), 331-338, November 2004.

Time trade-off quality of life research shows that years spent on dialysis after a failed transplant should be discounted by a factor of 0.8. The Hornberger et al.[2] paper states that the time-trade-off ratio for the first six months after a failed transplant is 0.77 (compared to life with a functioning transplant), and that after six months the time-trade-off ratio is 0.81.

c. What proportion of the LYFT score is derived from HLA DR typing? That is, all other things being equal, what effect on the probability of transplantation does 0, 1, 2 DR matching provide?

Answer:

The LYFT score is one component of the KAS. The proportion of the LYFT score due to HLA DR typing is not the same thing as the probability of transplantation. The Kidney Allocation Score depends on LYFT, DY, DPI, and CPRA. Two candidates with the same LYFT score can have widely varying Kidney Allocation Scores. The KAS that will lead to transplant will vary by geography and blood type.

All other things being equal, and excluding zero ABDR HLA recipients from the comparison, DR matching tends to affect recipients' LYFT scores in the following ways:

Mismatch level	0 DR	1 DR	2 DR
Survival with transplant	+0.6 years	0 (Reference)	-0.2 years
Graft survival	+0.5 years	0 (Reference)	-0.4 years
LYFT	+0.5	0 (Reference)	-0.2

The percentage of the LYFT score that is due to DR mismatch will vary according to the total LYFT score.

Overall, the range of estimated median patient survival with transplant is from 3.6 to 27.0, the range of estimated median graft survival is from 3.5 to 13.3 years, and the range of the LYFT score is from 3.2 to 11.4, among adult kidney-alone candidates active on the list.

d. We are concerned about the significant bias towards young people because of the significant number of years post-transplant that they will be alive when compared with older patients with competing risks for death. We would argue that life years after transplant should be discounted as is typical in most effectiveness research. Further, the discounting would decrease the uncertainty associated with the tails of the survival distribution and would provide for the ability to test other survival curves than the median survival curve in calculation of LYFT. We think it would inform the community more accurately if you would provide outcome models with standard discounting of life years along with test cases of 50% and 200% of the discount rate. Please discuss the effect discounting will have on outcomes and

[2] Hornberger, John C., Jennie H. Best, Louis P. Garrison, Jr. "Cost Effectiveness of Repeat Medical Procedures: Kidney Transplantation as an Example." Med Decis Making 1997; 17:363-372.

the ability to use other models. After discounting, please examine models using truncation of survival times at 5, 10, and 15 years.

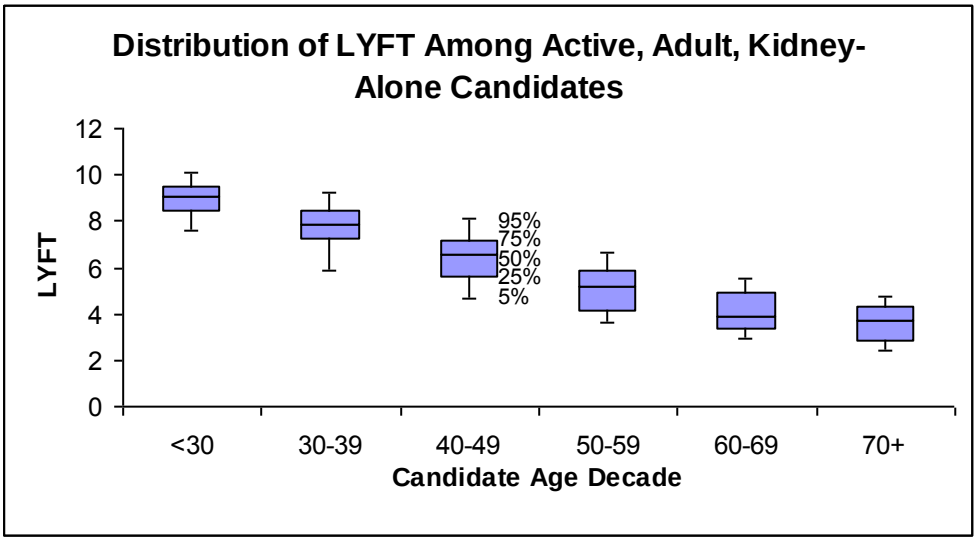
Answer:

The use of the median yields stable lifetime estimates based upon available data without extrapolation beyond the available data for most patients. The use of the median does represent the typical lifetime of candidates.

The OPTN Kidney Committee reviewed discounting, using the World Health Organization (WHO) standard of 3%/year, along with various truncations at 10, 20, and 30 years during the February 9, 2006 meeting. Discounting costs is rather non-controversial, however, discounting benefits (such as life-years gained) is not as straightforward. For cost-effectiveness and cost-benefit analyses, the WHO tends to discount costs and benefits at the same rate.[3]

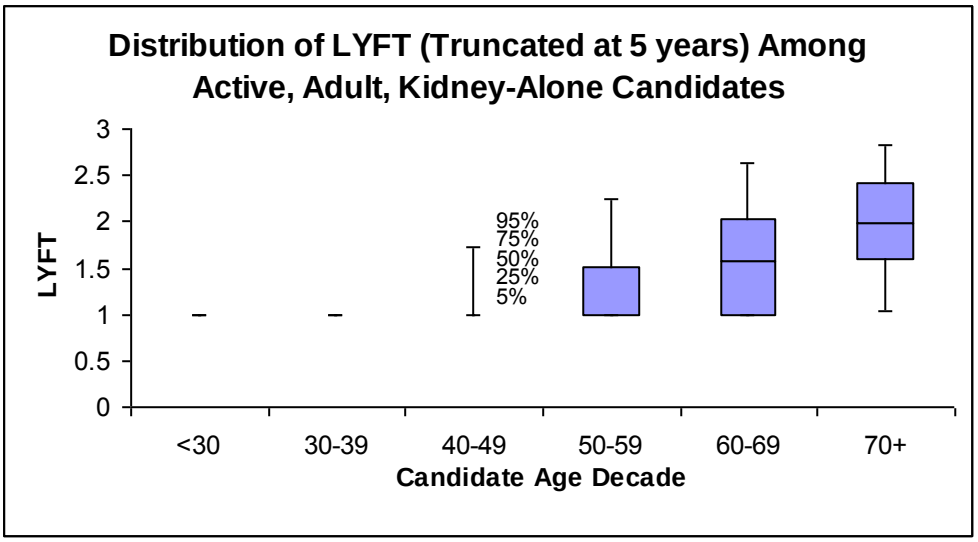
The consequence of truncation is that the resulting LYFT calculation would capture the entire transplant benefit for candidates with shorter lifespans, but would capture little if any of the benefit for candidates with longer lifespans.

Based on the current LYFT calculation, the distributions of LYFT by age decade using 5, 10, and 15-year truncation and using the standard WHO discounting of 3% per year are shown in the following box-and-whisker plots.

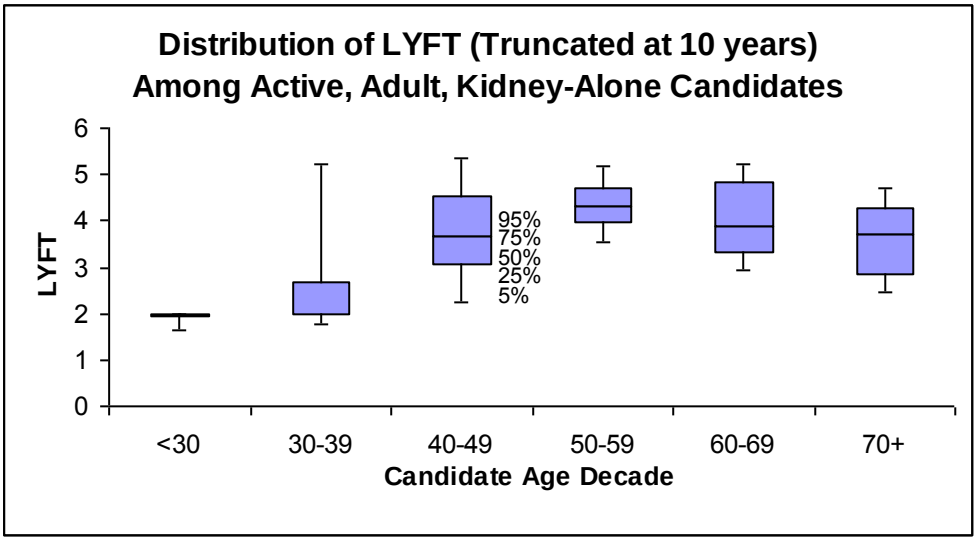


The figure above is based on median lifetimes (without discounting or truncation) and is shown for comparison purposes.

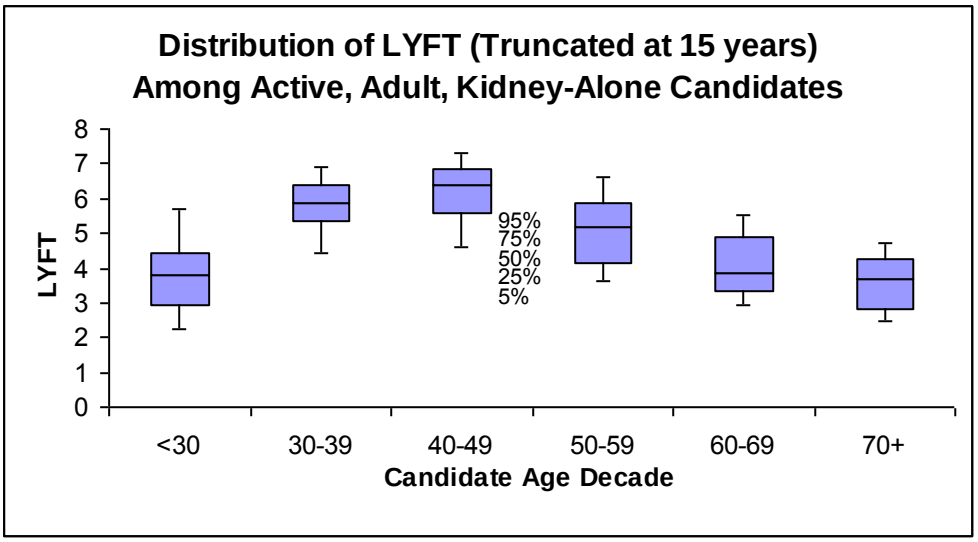
[3] The World Health Report. 2002. Chapter 5: Some Strategies to Reduce Risk. Accessed 11 August 2008.



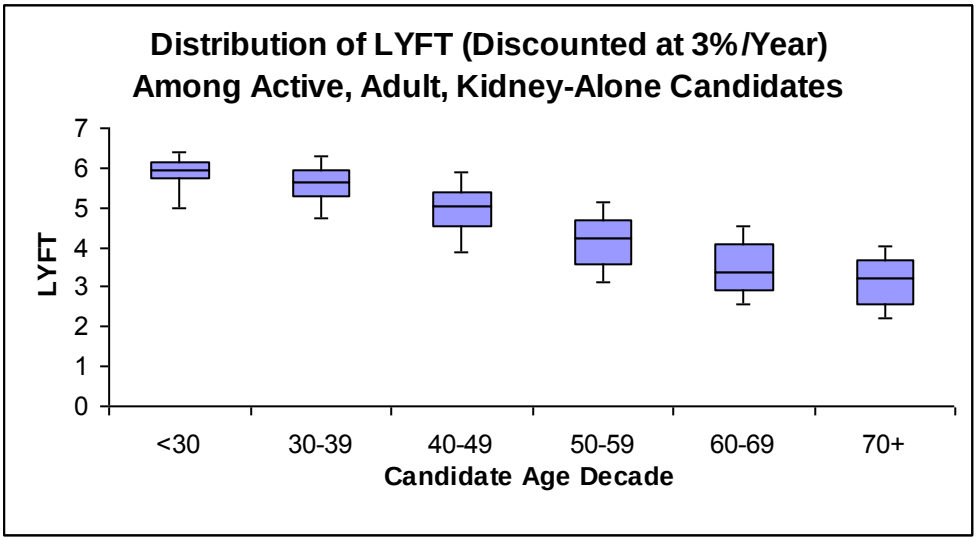
Truncated at 5 years, the increase in life among candidates under the age of 50 is largely excluded (see above figure).



Truncated at 10 years, the increase in life among candidates under the age of 30 is largely excluded (see above figure).



Truncated at 15 years, the increase in life among candidates under the age of 40 is only partially captured (see above figure).



Discounting affects longer lifespans more than shorter lifespans (see above figure).

e. What was the mean survival time of patients used in the survival models?

Answer:

Calculation of the mean survival time would involve extrapolation of the survival curve well beyond the available data. The median survival without a transplant was 6.2 years, and the median survival after transplant was 12.1 years.

f. Have results of re-transplants, as well as interactions between re-transplantation and other covariates in the LYFT model, been taking into account in the LYFT calculations?

Answer:

Yes, prior transplantation is a covariate in the LYFT calculation. There is an interaction between prior transplantation and diabetes, and that is accounted for in the LYFT calculations.

The interaction terms included in the LYFT calculation are as follows:

- Candidate age and candidate diabetes status
- Candidate albumin and candidate diabetes status
- Candidate BMI and candidate diabetes status
- Candidate not yet on dialysis and candidate diabetes status
- Candidate PRA > 10 and candidate diabetes status
- Candidate previous transplant and candidate diabetes status
- Candidate DSA different from Donor DSA
- Candidate/Donor HLA MM (0 ABDR MM, 0-2 DR MM)

g. Various contradictory reasons have been given for the exclusion of race from the LYFT calculation. In the past, the committee justified not using ethnicity in the algorithm – saying that there was now little difference in SRTR data. However, recently the rationale was presented that the committee excluded ethnicity as it was too politically charged and the groups were not well-divided. From SRTR data for live donors and non-ECD deceased donors, there is a difference in outcomes between African Americans and Caucasians. What exactly is driving the exclusion of this predictor from the LYFT score? What is the ethical justification of excluding some variables, race, and including others, diagnosis and age?

Answer:

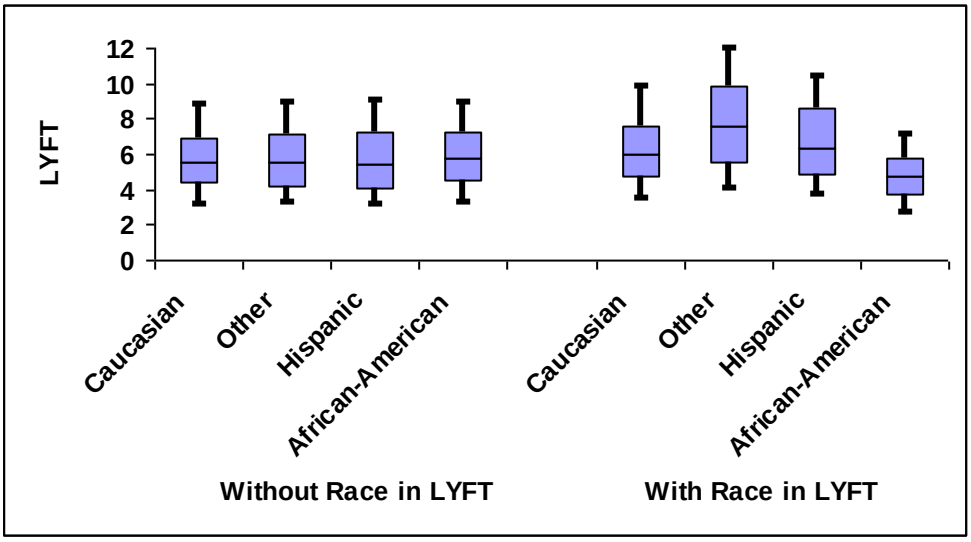
LYFT is a comparison of two lifetimes; patient survival after transplant is similar between African Americans and Caucasians in SRTR data for ECD, non-ECD, and living donor recipients:
http://www.ustransplant.org/annual_reports/current/512a_can-race_ki.htm
http://www.ustransplant.org/annual_reports/current/512b_can-race_ki.htm
http://www.ustransplant.org/annual_reports/current/512c_can-race_ki.htm

A recent analysis of data from the Dialysis outcomes and Practice Patterns Study by Robinson et al. reported that dialysis lifetimes are similar among non-Hispanic White and African American populations (HR=0.97).*

*Robinson BM. Joffe MM. Pisoni RL. Port FK. Feldman HI. Revisiting survival differences by race and ethnicity among hemodialysis patients: the Dialysis Outcomes and Practice Patterns Study. Journal of the American Society of Nephrology. 17(10):2910-8, 2006 Oct.

The SRTR was instructed to consider variables in the LYFT calculation that were A) collected by the OPTN (completeness and validity were considered), B) objective, and C) had statistical and clinical relevance.

The SRTR provided the distribution of LYFT scores by race with and without race in the model to the Kidney Committee. These distributions are as shown in the following figure.



The following criteria were used to select data elements for inclusion in the LYFT calculation.

- **Objectivity**
  The Committee intends to create a system that is based upon objective medical criteria (as required by the OPTN Final Rule). Therefore, the data elements selected must be consistently assessed across transplant centers. For example, the calculation for body mass index (BMI) can be replicated reliably for all transplant candidates. Since, there is not a validated, objective test of ethnicity/race, the determination of race or ethnicity, could not be reliably replicated for all transplant candidates.. Through the use of objective medical criteria, the Committee intends to build a system that is not subject to selective or biased reporting of candidate characteristics.

- **Statistical Significance or Clinical Importance**
  The Committee assessed each data element to determine if it had an effect on the LYFT calculation and excluded those that did not have a statistically significant effect. The Committee also excluded those elements which were not thought to be clinically relevant (e.g., type of dialysis treatment).

- **Data Quality**
  Since the simulation models rely on actual data collected from 1987-2002 on kidney transplant candidates and recipients, data elements selected for LYFT had to be available during this time period in order to be utilized. Since certain data elements were added to UNet[sm] after 1987, or were discontinued prior to 2002, there may not have been complete enough data to analyze for the LYFT calculation. The Committee excluded those data elements where the data quality or completeness was compromised.

The following table summarizes the variables initially considered, those recommended for exclusion, and those retained.

Candidate Variables Initially Assessed	Included	Excluded	Reason for exclusion
Time Exposed to ESRD	X		
Albumin	X		
Body Mass Index (BMI)	x		
Age	x		
Diagnosis:	x		
HTN	x		
Polycystic	x		
Diabetic	x		
Other/Missing	x		
Previous Transplant	x		
Peak PRA (= highest PRA recorded)	x		
Ethnicity/Race		x	Not objective
Angina		x	Not objective, not statistically sig
PVD		x	Not objective, not statistically sig
Calendar Year of Listing		x	Not statistically significant or clin important
Gender		x	Not statistically significant or clin important
NYHA Functional Class		x	Not objective; no longer collected
Primary Insurance Status		x	Not statistically significant
Drug-Treated Hypertension		x	Not objective
Type of Dialysis Treatment		x	Not clinically important
DSA (Surrogate for Geography)		x	Issue of geography deferred for ph consideration
Previous Malignancy		x	Not statistically significant

h. The variables included in the model include a number where the data was noted as missing. While these variables have significant effects, one wonders about the veracity of the data with inclusion of these variables. Please provide estimates of outcome and uncertainty with and without inclusion of these variables. The list of variables excluded has a number of variables that have significant effect on LYFT, but were excluded for apparently arbitrary reasons, such as "inappropriate for allocation."

Answer:

The most predictive covariates, including age, diabetes, and time on dialysis, have very small amounts of missing data due to the fact that the SRTR has used CMS data to supplement the OPTN information on candidates. Some fields such as prior transplant and shared v. local, while less influential in their effects on LYFT, also have no missing data. For albumin and BMI, missing data were largely a result of changes in data collection which would not bias the results. Models have been run with missingness indicators for missing data and with multiple imputation. The LYFT scores calculated using these techniques correlate with an $R^2 >$ 0.99. Variables with missing data were handled using missingness indicators because limiting survival calculations to those based only on historical candidates without any missing data could bias survival estimates.

The LYFT model incorporates the median estimated lifespan for three separate outcomes, with each outcome's estimate based on two separate models. For this reason, the variability in the LYFT estimates has been calculated using a bootstrap approach. This approach involves randomly re-sampling the data with replacement, calculating the models, and calculating the LYFT estimates to obtain a measure of the variability in these estimates. The data used to calculate each of the LYFT models (transplant recipients from 1987 – 2006 and active candidates sampled from five dates during this period) was randomly resampled 50 times with replacement.[4] The average standard deviation among the bootstrapped estimates of LYFT among active candidates offered an average kidney was 0.14. This standard deviation tended to be larger for candidates with larger LYFT estimates, and the 99th percentile of this standard deviation was 0.43. The average between-run Kendall's tau (a measure of the agreement between two rankings) was 0.95.

i. It is unclear how PRA was used in the model, in particular PRA greater than 80. Please explain how the effect of PRA was estimated.

Answer:

[4] Procedure as described in "Efron, Bradley, and Robert J. Tibshirani. "An Introduction to the Bootstrap." P. 47. 1993, Chapman & Hall, Inc.

In LYFT, PRA was modeled using categorical variables based on the peak PRA reported to the OPTN up to the time of transplant or offer, while in the KAS, PRA has a continuous contribution. For LYFT, PRA was allowed to have a non-linear effect by using indicator variables based on peak PRA < 10, peak PRA 10-79, peak PRA 80-100, and a separate missingness indicator for candidates whose PRA score was never reported.

j. The committee is looking at a model that decreases the benefit of LYFT in order to encourage living donation. Please provide estimates of the decrease in the number of living donor transplantations using estimates from transplant centers with very short waiting times and the pediatric population. Please estimate the outcome with 1.0, 0.8, and 0.6 LYFT.

Answer:

The SRTR did not provide direct data about this question because KPSAM cannot predict changes in behavior. The SRTR did provide information on the decline in living-donor transplantation following the pediatric rule change to prioritize pediatric candidates for kidneys from donors <35. This analysis is is posted on the OPTN Website.[5]

k. Apparently the LYFT modeling was performed using only SCD kidneys. Please explain what affect this would have on the models as compared to using both SCD and ECD kidneys. What is the effect of having donor factors in both LYFT and DRI?

Answer:

Initial models for LYFT used only SCD kidneys. The more recent models upon which the KAS is built used all adult kidney-alone candidates and recipients (SCD, ECD, and DCD). Inclusion of donor factors in the LYFT formula reflects the actual change in lifetime due to transplantation with each organ for each candidate. This captures the effect of donor/recipient matching. The use of the DPI in the KAS is for a different purpose. DPI is used to apportion the KAS between DY and LYFT.

2. LYFT Score, Statistics:
   a. We understand that LYFT score accurately predicts life years from transplantation. In other words, we understand that for given patient characteristics, the predicted LYFT would be comparable to the average life years gained among historical patients with those characteristics. However, it is likely that some patients with those characteristics had much shorter life years gained than would be predicted by LYFT, and some had much longer life years gained than would be predicted by LYFT. We would like to see this uncertainty about the variability in life years quantified. Please provide a c-statistic and an expression of the residual uncertainty of the outcome and estimates of this uncertainty vs. age. The examination of the variability should also examine the effect of varying

[5] OPTN Pediatric Committee data request #5051 "Evaluation of Share 35 Policy," presented to the Committee November 29, 2007.
http://www.unos.org/CommitteeReports/board_main_PediatricTransplantationCommittee_2_25_2008_15_24.pdf

the accounting of post transplant years from the median to shorter time intervals such as 25/75, 5 years and 10 years (see below).

Answer:

The c-statistics for the various components of the LYFT score (short- and long-term patient survival without a transplant and short- and long-term patient and graft survival with a transplant) are as follows:

Index of Concordance for:	0-4 years	4-15 years
Patient survival without transplant	0.66	0.60
Patient survival with transplant	0.67	0.68
Graft survival	0.59	0.57

The rationale for not truncating survival calculations are presented in response to question 1.d.

b. Patient survival was censored at the time of re-transplantation. This seems arbitrary. Please explain the advantages and disadvantages of this censoring.

Answer:

The SRTR considered several possible approaches to handling recipients who subsequently receive a re-transplant in the LYFT survival analyses. These include: censoring at re-transplant, excluding these patients from the analysis, treating a re-transplant as a failure (death), continuing follow-up beyond the re-transplant, or imputing the survival time that would have occurred with the original transplant that is missing due to the re-transplant event. Censoring assumes that the censored recipient's survival would be similar to that of other recipients who are still alive at that point in their post-transplant period if their other characteristics were also similar. Censoring may be considered a more reasonable assumption than assuming that patient survival ends at re-transplantation, and may also lead to less bias than would excluding these patients entirely from the analysis. Continuing follow-up in survival analyses until the patient eventually dies after their second transplant would assign additional benefit from a kidney transplant to patients who live long enough to receive a re-transplant. Finally, imputation, if based on the characteristics of the recipients used in the survival models, would on average result in the same answer as censoring.

c. The long term extrapolation of the data necessary appears to be fanciful and would have significant effects on the outcome of LYFT. The use of the latest results to correct the slope of earlier curves but with the use of the later data to create these slopes would argue against the use of long time frames after transplantation.

Answer:

Use of the median obviates the need for extrapolation for most patients. 1% of the lifespans without transplant and 26% of the lifespans with transplant required extrapolation. These extrapolations were shorter than 3 years among extrapolated lifespans without transplant and 8 years among extrapolated lifespans with transplant for 90% of the candidates who

required extrapolation. The methodology used for long-term extrapolation was evaluated by the SRTR Technical Advisory Committee, and various alternatives were suggested to test the sensitivity of the extrapolation to the methods used. The correlation between lifespans estimated through the population-based (i.e. using hazards proportional to those experienced by the general population as they age over time) and the log-linear extrapolation method (the method used in the LYFT scores in the current proposal) exceeded 0.99, indicating that the estimates were robust with regard to the specific extrapolation method used.

d. We are concerned that the LYFT score is based on a dataset that suffers considerably from unmeasured confounding, residual confounding, missing data, and possible misclassification bias. Examples include key missing recipient factors such as cardiovascular disease, and binary comorbidities instead of severity indicators. The appropriateness of risk adjustment based on SRTR data is already a point of controversy in terms of transplant center outcomes monitoring. We are concerned that these same flaws will be perpetuated to something that not only monitors our outcomes but decides which of our patients dies. We understand that LYFT (and the OPTN/SRTR database) is "the best that we can currently do with the data we have." However, we would point out that although the bar for validity and reliability is not as high for research studies, and not even as high for outcomes monitoring, it should be very high when we put our patients' lives on the line.

Answer:

Even with the limitations of the data available, substantial differences in lifetimes are predicted and have been validated using split halves, and the index of concordance remained virtually unchanged. Even with these limitations, the LYFT calculation demonstrates the potential to substantially increase years of life with a new allocation system. Use of improved data in the LYFT calculation will lead to even greater increases in patient lifetimes.

e. Has an expert panel examined the mathematical and statistical methods used in the development of this allocation policy? If so, was a report made? Were changes in the methods suggested?

Answer:

The methods used by the SRTR for survival calculations have been published.[6] The SRTR Technical Advisory Committee, and Technical Expert Panels convened by HRSA have reviewed the statistical modeling methods and simulation methods used by the SRTR in determining transplant benefit and in simulating the effects of changing the allocation system. Tests of the assumptions and the methodology were suggested and carried out, and advice on methodology was incorporated into the analyses used in the current kidney allocation proposal. A written report by the

[6] "2007 SRTR Report on the State of Transplantation," available on the American Journal of Transplantation (AJT) website (www.blackwell-synergy.com/toc/ajt/8/4p2)

Independent Expert Panel is titled "Independent Expert Panel on Statistical Methods for the Analysis of Organ Transplantation Data: Analytic Methods and Simulation Modeling." (Attachment A)

3. Donor Profile Index (DPI)
    a. This score comprises a significant portion of the LYFT-based simulations, yet nothing on the SRTR website details its methodology and validation. We would request that simulations from UNOS should be based on either a working paper on the website or a published manuscript. We would hope that the paper describing the DPI would describe the uncertainty of the score. The estimates of uncertainty and those of LYFT should be graphically represented in estimates of the benefit of the system and the age distribution of those receiving transplantation.

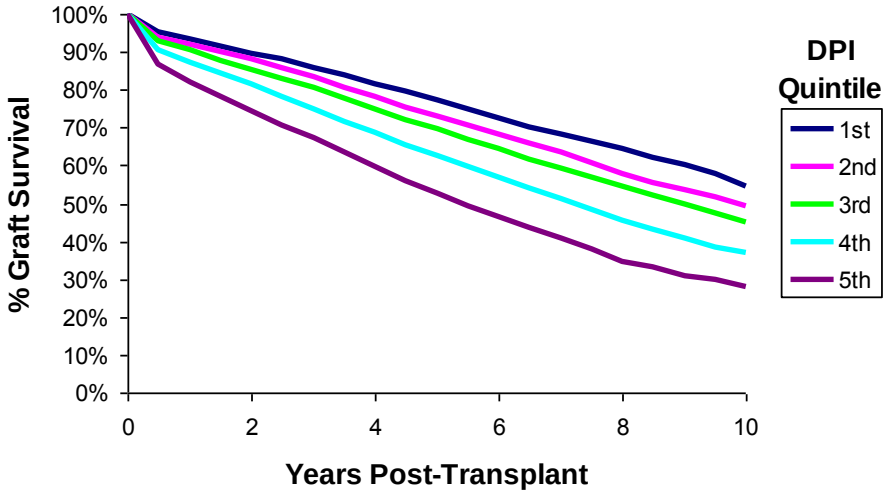
Answer:

Just to clarify- the DPI is separate from the LYFT calculation. DPI is used to determine the relative contributions of LYFT and Dialysis Years in the KAS calculation.

The authors of the paper on which the DPI is based are planning to submit their paper soon to a peer-review journal. The methodology was reviewed by the SRTR Technical Advisory Committee on April 10, 2008.

The DPI is highly predictive of transplant outcome. See table and figure below on allograft survival. Lower DPI kidneys have longer potential allograft survival; whereas, higher DPI kidneys have shorter potential allograft survival.

DPI Quintile	Median Lifetime (yrs)
0-20%	11.22*
20-40%	9.7
40-60%	8.9
60-80%	7.2
80-100%	5.4



b. We are concerned that the DPI might neglect effect modifiers (interaction terms). In other words, we know clinically that some kidneys will do better in some recipients than in others, and wonder if this kind of donor/recipient matching will be taken into account with the DPI score.

Answer:

The DPI is used to rank order organs based on organ quality alone in the current KAS calculation. Donor/recipient interaction terms are accounted for in the LYFT calculation which is a component of the KAS.

The overall Kidney Allocation Score (KAS) does take into account the fact that some kidneys will do better in some recipients than others. However, this is taken into account in the candidate LYFT score and not in the DPI. LYFT estimates survival based on receiving a kidney from a *specific* donor. The LYFT score includes several donor characteristics including donor age, weight, cause of death, hypertension, and CMV status. The donor profile index (DPI) is an assessment of donor characteristics, independent of recipient characteristics. Donor/recipient interaction terms, including donor/recipient HLA mismatch and donors coming from a different DSA than the recipient, are accounted for in the LYFT calculation. Donor age and recipient age was tested as a potential interaction, but was not a significant interaction in the post-transplant patient survival models.

4. Simulations:
   a. Since some patients will be much less likely to receive a transplant (i.e. older patients), it is likely that death on the waiting list will increase. We feel that death on the waiting list should also be a metric taken into consideration when evaluating simulation outcomes. What is the increase in death on the waiting list by age decile from patients left on the list at one and three years as compared to the current system?

Answer:

During the course of a three-year simulation run, the average 1-year Kaplan-Meier survival (censored at transplant) on the waitlist by age decade was as follows:

Age	Current national system	Proposed System
<30	97.80%	97.89%
30 - 39	96.68%	96.62%
40 - 49	95.53%	95.46%
50 - 59	94.13%	94.12%
60 - 69	91.86%	91.85%
70 - 79	88.85%	89.05%
80+	83.22%	83.55%

As can be seen in the table above, there are very slight shifts in waitlist survival.

b. After the run of the simulations, what is the resulting distribution of age by DPI by time on waiting list of those patients transplanted? What is the distribution of time on the list and age for patients left on the list?

Answer:

The distribution of age by DPI by dialysis years (DY) (DY are used in the Kidney Allocation Score in place of waiting time) among recipients is shown in the table below. Based on three-year simulations of the current national allocation system and the proposed system, lower DPI kidneys (that is kidneys with longer potential allograft survival) tend to shift towards candidates under the age of 50.

Age	DY	DPI	Recipients: Current National System	Recipients: Proposed System
0-17	<1	0-0.24	535	697
0-17	<1	0.25-0.49	368	431
0-17	<1	0.50-0.74	136	61
0-17	<1	0.75+	39	14
0-17	1-2	0-0.24	124	125
0-17	1-2	0.25-0.49	87	79
0-17	1-2	0.50-0.74	25	9
0-17	1-2	0.75+	10	3
0-17	3+	0-0.24	55	73
0-17	3+	0.25-0.49	50	46
0-17	3+	0.50-0.74	16	7
0-17	3+	0.75+	6	1
18-34	<1	0-0.24	337	862
18-34	<1	0.25-0.49	297	638
18-34	<1	0.50-0.74	243	433
18-34	<1	0.75+	165	206
18-34	1-2	0-0.24	182	444

Age	DY	DPI	Recipients: Current National System	Recipients: Proposed System
18-34	1-2	0.25-0.49	193	337
18-34	1-2	0.50-0.74	204	239
18-34	1-2	0.75+	183	125
18-34	3+	0-0.24	314	566
18-34	3+	0.25-0.49	375	586
18-34	3+	0.50-0.74	475	525
18-34	3+	0.75+	338	302
35-49	<1	0-0.24	651	822
35-49	<1	0.25-0.49	605	813
35-49	<1	0.50-0.74	547	757
35-49	<1	0.75+	445	580
35-49	1-2	0-0.24	369	443
35-49	1-2	0.25-0.49	414	489
35-49	1-2	0.50-0.74	504	486
35-49	1-2	0.75+	461	397
35-49	3+	0-0.24	698	722
35-49	3+	0.25-0.49	879	969
35-49	3+	0.50-0.74	1136	1195
35-49	3+	0.75+	973	903
50-64	<1	0-0.24	910	452
50-64	<1	0.25-0.49	818	578
50-64	<1	0.50-0.74	734	769
50-64	<1	0.75+	908	1048
50-64	1-2	0-0.24	510	271
50-64	1-2	0.25-0.49	616	396
50-64	1-2	0.50-0.74	701	533
50-64	1-2	0.75+	943	807
50-64	3+	0-0.24	744	350
50-64	3+	0.25-0.49	971	660
50-64	3+	0.50-0.74	1277	1186
50-64	3+	0.75+	1385	1442
65+	<1	0-0.24	323	94
65+	<1	0.25-0.49	282	130
65+	<1	0.50-0.74	267	195
65+	<1	0.75+	431	469
65+	1-2	0-0.24	190	58
65+	1-2	0.25-0.49	203	88
65+	1-2	0.50-0.74	270	148
65+	1-2	0.75+	459	390
65+	3+	0-0.24	224	62
65+	3+	0.25-0.49	301	118
65+	3+	0.50-0.74	395	293
65+	3+	0.75+	534	541

The distribution of time on the list and age for candidates left on the list varies depending on the time point used to decide that a candidate has been "left on the list." Later time periods (e.g. 5 years after listing) will have different distributions of age than earlier time periods (immediately after listing) due to DSA waiting times, transplantation, removal from the waitlist, and death.

c. How would this system affect likelihood and outcomes after re-transplantation?

Answer:

With regard to outcomes, the results from the simulations indicating increases in post-transplant lifespan and graft survival include both recipients who have and have not had prior transplants, but outcomes for patients who have had prior transplants have not yet been separately analyzed.

With regard to the likelihood of re-transplantation, there are three aspects of the Kidney Allocation Score that might be affected by a prior transplant: the LYFT score, dialysis years, and PRA. The LYFT score of candidates who had prior kidney transplants is, all else equal, typically 0.2 lower than those candidates who have not had a prior kidney transplant, reflecting the combined effect of reduced patient survival without a transplant (0.3 years), reduced patient survival following transplant (0.4) and reduced graft survival (0.6). Dialysis years are calculated from the most recent initiation of maintenance dialysis. On average, PRA is higher among prior recipients.

d. Decreasing the weighting of LYFT to 80% appears to shift more kidneys to older patients. As the purported benefit of this change was to increase living donation, has the committee examined delaying the effect of LYFT directly, for example one or two years after listing, rather than decrementing the LYFT weighting?

Answer:

These analyses have not been performed. Some options for addressing this unintended consequence in the new allocation system included mandating a certain period of waiting time for all candidates, but determining a length of time was not possible due to a lack of evidence for the period of time on dialysis that would be necessary to motivate pursuing a living donor. Additionally, since waiting times are so variable across DSAs, a mandatory period of waiting time would have differential impact across the country. Instead, the Committee decided that the contribution of LYFT to the KAS should be capped at 80%. This means, that even for the highest quality kidneys, that 20% of the KAS would be candidate time on dialysis. This approach does not deny candidates access to a deceased donor transplant for an arbitrary period of time. The Committee investigated the effect that this 80% cap would have on post-transplant outcomes and found that it would not substantially reduce the

average lifetime, graft lifetime, or extra years of life experienced by transplant recipients.

e. Looking at the older patient, with long waiting times and the highest risk kidneys, what is the benefit of transplantation and the variability around this benefit? Please explain how using LYFT, which looks at survival using SCD kidneys, can be used to examine survival following higher DPI index kidneys.

Answer:

After the Forum in Dallas (February 2007), the Committee reconsidered its initial plan to only revise SCD kidney allocation. Based on the feedback from participants, the Committee determined that there is a need to replace the current SCD and ECD categories for kidneys with a more continuous measure of donor quality (donor profile index or DPI). In more current iterations of the simulation modeling.

LYFT used both SCD and ECD kidneys in the KAS models, and thus can be used for both low and high DPI kidneys. Some older patients may die soon after transplant and others may live as long as twice the median lifetime, but the typical, or median lifetime is known very accurately. See the response to question 1a for a discussion of the ranges of LYFT and KAS by age.

f. We do not have the outcome of all the simulations available. Given the relatively low bar for the increase in LYFT being considered with the 0.8 LYFT model, have the previous simulations been examined in light of this bar (i.e. has the committee re-reviewed the previous simulations)? Please make available (in a public place such as www.ustransplant.org) the outcome of the previous simulations and the reasons why the committee decided not to use the allocation scheme.

Answer:

All of the simulations and resulting Committee discussion are available in the Committee reports posted at http://www.optn.org/kars.asp. As the system under consideration today was built based on a series of incremental changes, comparison to early models is not possible. For example, early simulation runs only included SCD kidneys. Based on public feedback in February 2007, the Committee decided to include all kidneys in the proposed system and to replace the SCD/ECD distinction with a donor profile index (DPI). A table of the simulation runs and a brief description of each is provided below.

System	Run #	Description	Ba Ru
Current	1	Current national allocation system as baseline	

	2	No interleave – separate allocation for adult & pediatric candidates	1
SCD-LYFT	3	LYFT in place of kidney points for adult candidates of SCD organs; KP Priority	2
	4	No paybacks; KP Priority	3
	5	No 0MM sharing; KP Priority	4
	6[3]	Eliminate 0MM priority locally for adult candidates of SCD organs; KP Priority	5
	7[4]	No KP Priority – KP and KI candidates compete by LYFT for adult candidates	6
	8	A_2-> B; KP Priority	7
	9	0MM sharing for PRA 80%+ adult candidates; KP Priority	8
	11	National allocation, no geographic boundaries; KP Priority	9
	15	LYFT (No HLA A, B); KP Priority.	9
	16	LYFT (No HLA A, B; with PKD, DM); KP Priority.	9
	16a	LYFT (No HLA A, B; with PKD, DM); KI follows PA.	9
	17	LYFT (No HLA A, B; with DM only); KP Priority.	16
SCD- LYFT modified by ESRD years	10	LYFT + X*(DY) (X=1,2) (Note: This run superceded by run 12a-12d)	9 9
	12a	LYFT + 0.1*(DY)	
	12b	LYFT + 0.2*(DY)	9
	12c	LYFT + 0.5*(DY)	9
	12d	LYFT + 1*(DY)	9
	24	LYFT + 0.5*(DY) (LYFT with no HLA A,B; with PKD, DM; KI follows PA)	16
SCD-LYFT modified by waitlist time	13a	LYFT – x*Waitlist lifetime. X=0.2	9
SCD-LYFT modified by PRA	14a	LYFT + 0.01*PRA	9

All Donors-LYFT	18a	LYFT * (1-DPI) + DY * DPI. KP Priority.	17
modified by ESRD time	18b	LYFT * (1-DPI_2) + DY * DPI_2 . KP Priority.	17
and continuous	18c	LYFT * (1-DPI) + DY * DPI + PRA*4. KI follows PA.	16
DPI	18d	LYFT * (1-DPI_2) + DY * DPI_2+ PRA*4. KI follows PA.	16
	18e	LYFT * (1-DPI_3) + DY * DPI_3+ PRA*4. KI follows PA.	16
		Committee Decision Point—18c, 18d, or 18e	
All donors - LYFT modified by continuous DPI and Lifetime matching	19	LYFT * (1-DPI_X) + DY * DPI_X+ PRA*4 + Lifetime matching (PL – GL) Note: X (power of DPI) needs to be selected by committee after reviewing results of 18c, 18d, 18e	Ch on or or
All donors - Discrete Categories	21	LYFT quintiles, DPI quintiles, rank by DY within matching quintiles, allocate to nearest quintile, KP Priority.	17
	21a	LYFT quintiles, DPI quintiles, rank by DY within matching quintiles, allocate to nearest quintile, KI follows PA (redo 21)	16
	21b	Post-transplant lifetime (PL) quintiles, DPI quintiles, rank by DY within matching quintiles, allocate to nearest quintile, KI follows PA	16
	21c	21b + Allow patients with >80% PRA to receive kidneys from 1 quintile above and all quintiles of DPI below[6]	21
	21d	21c + Absolute trump for 0ABDR HLA mismatch in quintile #1, 1 point for 1 DR mismatch, 2 points for 2 DR mismatches[7]	21
	21e	Instead of quintiles of DPI and post-transplant lifetime, use deciles	21
All donors – Age Matching	22	Continuous age matching, KP priority	17
	22a	Continuous age matching, KI follows PA	16
All donors – Waiting Time	23	Current points system, with A_2 -> B, no paybacks, 0MM share only for PRA 80%+	1

Variations on 18c	25	LYFT* (1-DPI) +DY*DPI+PRA*8 (instead of 4)	18

g. Did the simulation examine the effect of policies that would force patients on the waiting list to wait for the best kidney, such as waiting for a better matched kidney or a younger kidney?

Answer:

Some simulations that prioritized quintiles of candidates by LYFT to the corresponding quintile of donor kidneys by DPI were run by the SRTR and evaluated by the OPTN Kidney Committee. Simulations with age matching were also run. Results and synopses of committee discussions on these are available at the OPTN website (http://www.optn.org/kars.asp, "August 14, 2007 Committee Meeting Summary and Simulation Results").

h. Comparing the different simulations, what is the degree of matching of recipient age and DPI/donor age?

Answer:

Donor and recipient age are correlated under the current national allocation system with a Pearson correlation coefficient[7] (R) of 0.10. Using the KAS with DPI, the simulations resulted in a correlation coefficient of 0.31.

5. Logistics:

a. We are concerned about the cost of HLA typing for continuous PRA. Has this been modeled?

Answer:

In December 2006, the OPTN/UNOS Board of Directors approved modifications to its PRA policy (3.5.11.3) to use calculated PRA (CPRA) to determine the degree of a candidate's sensitization against the donor pool. This policy change was recommended by the Histocompatibility Committee and remains separate from the work to develop a kidney allocation proposal. Questions regarding the cost of HLA typing to determine CPRA would be better answered by the Histocompatibility Committee.

b. It seems counter-intuitive to decrement the LYFT score for patients with high PRA and then provide a factor to increase transplantation in this population. What are the results of simulation if both factors are removed from the model? What is

[7] Note: Correlations can range from -1 to +1. To obtain the R^2, also known as the proportion of explained variation, square the above numbers (i.e. an R of 0.10 would correspond to an R^2 of 0.01).

the sensitivity of the simulations to the factor given to increase transplantation in this population? Looking at the SRTR data on outcome in this population, what is the LYFT tradeoff of providing transplant to this population? What is the justification of provision of priority to this population rather than to another underserved population?

Answer:

The LYFT metric is a measure of transplant lifetime, and the accuracy of this measure is improved when PRA is taken into account. Simulations were performed that varied the number of points added to the KAS score by PRA. These results, including the number of transplants by PRA group and total numbers of life-years lived, are available on the OPTN website (http://www.optn.org/kars.asp, "December 3, 2007 Committee Meeting Summary and Simulation Results")

LYFT is a calculation to determine the incremental gain in life years that a candidate may attain with a kidney from a specific donor. PRA met the criteria for inclusion in the LYFT calculation as it is objective, statistically significant, clinically important and of sufficient data quality to justify inclusion. However, allocation policies must also be balanced to achieve certain objectives, goals, or to meet certain requirements. The National Organ Transplant Act of 1984 (NOTA) specifically requires that the OPTN establish "a national system, through the use of computers and in accordance with established medical criteria, to match organs and individuals included in the list, especially individuals whose immune system makes it difficult for them to receive organs".[8] Therefore, the Committee instituted a sliding scale based on CPRA to provide both moderately and highly sensitized candidates assured access to transplantation.

c. What is the plan for transitioning from the current system to a LYFT-based system? If this is not a smooth transition, we are concerned that patients who have been waiting for many years, but who have low LYFT scores, will be angered and demoralized. One possibility would be to gradually phase in the weight of the LYFT (start with 0.8 waiting time and 0.2 LYFT and then slowly transition to the goal of 0.2 waiting time and 0.8 LYFT).

Answer:

Candidates who have been waiting on dialysis under the current system will retain those years in the KAS calculation.
The Kidney Committee has not yet asked the SRTR to simulate any transition plans.

When assessing whether to have a transition period, the Committee considered whether to phase in the new system gradually and determined that the system, in most donor service areas, will favor those candidates

[8] NOTA

with longer waiting times. While the system would be implemented for all candidates at the same time, there will be pre-implementation phases for transplant centers to ensure that they have the necessary data for all of their candidates.

d. What will happen to paybacks as this system transitions in, with particular attention to centers that are owed a number of paybacks?

Answer:

The KAS system simulations eliminated paybacks and payback debts.

Current policy states that when an OPO accepts a zero-antigen mismatch offer from another OPO, it is required to "pay back" the kidney with a kidney from the same blood group after at least two debts have been accrued. Kidneys are offered as paybacks after being offered to zero-antigen mismatch candidates, prior living organ donors, highly sensitized candidates who are listed in the same donation service area (DSA) as the donor, and children (if the kidney is from a donor younger than 35 years old). The current payback policies were instituted as a mechanism to address the imbalances created by the zero antigen mismatch sharing policies. Since zero-antigen mismatches are more frequent among those with common antigens, the payback policy was intended to rebalance the allocation system so that no patient population benefited from the zero antigen mismatch sharing rules to the harm of any other patient population.

Some OPOs have reported difficulty in placing payback kidneys for several reasons including the fact that that OPOs are not required to accept payback kidneys. Therefore, the OPO may choose to turn down several payback offers, and wait for a kidney from a donor with specific characteristics to be offered, before the OPO accepts a payback offer. Since the pediatric priority policy for donors under the age of 35 went into effect in 2005, some OPOs have observed a decrease in the percent of kidneys from donors <35 available for payback offers. The Kidney Transplantation Committee has reviewed acceptance rates for kidneys offered as paybacks and has found that few of the kidneys offered for paybacks are actually accepted, and acceptance rates vary widely based on OPO.

The penalties for exceeding the stated payback debt levels have not served as an effective governor for the zero-antigen mismatch and payback policies. One penalty reprioritizes unsensitized candidates from OPOs with payback debt levels greater than 9 to the bottom of the match run category for zero-antigen mismatches. Additionally, when a donor service area (DSA) reaches a debt level of 9 kidneys across all blood groups, it may no longer retain a kidney for local simultaneous pancreas-kidney (SPK) transplantation. Instead, the kidney must be offered as a payback. The result has been a decrease in SPK transplantation in some of OPOs with consistently moderate to high debt levels. In 2006 and 2007, the

Pancreas Transplantation Committee reviewed reports from several pancreas transplant programs that were unable to perform SPK transplants due to kidney payback debt levels. Thus, SPK patients within the DSA are penalized. Informed by the actual data on this issue, the Kidney Transplantation Committee and Pancreas Transplantation Committee both believe that this situation is disproportionately affecting candidates listed for kidney-pancreas transplantation.

Given these problems, the kidney payback system will be eliminated in the proposed system. Eliminating the kidney payback system, while still ensuring equitable access for highly sensitized candidates to zero-antigen mismatched kidneys, is expected to improve the efficiency of the kidney allocation system.

Due to the complexities associated with the payback accounting system, all existing payback debts and credits need to be settled prior to implementation of a new kidney allocation system. Any debts/credits remaining at the time of implementation would be eliminated.

Considering that repayment of debt and fulfillment of credits may take some time, organ procurement organizations with either high debt or credit levels should now consider ways to reduce these levels.

e. Although the use of dialysis time (instead of waiting time) makes up for nephrologists who fail to refer patients to transplantation in a timely manner (which we acknowledge has been reported many times and amplifies racial, socio-economic, and gender disparities), we are concerned that patients who were offered transplantation but initially refused (and eventually consider it when they get into trouble) will be treated with the same level of priority (or higher) as those who "did everything the right way from the start."

Answer:

During the development of this proposal, the Committee heard from many stakeholders that any proposed kidney allocation system should include a mechanism for candidates to gain priority over time. Some reported that without a way to improve the kidney allocation score over time, candidates (especially those with lower LYFT scores) may lose hope that they could receive a transplant, resulting in negative health-related consequences or the discontinuation of dialysis.

The current kidney allocation system gives considerable priority based on waiting time for adult candidates. Waiting time in the current system is defined as the amount of time since the candidate was placed on the waiting list. A candidate's waiting time in the current system may be affected by many factors including geographic region,[9] blood type, referral patterns, geographic location (e.g., urban versus rural), and proximity to a

[9] Ashby VB, Lin M, Kalbfleisch JD, Port FK, Wolfe RA, Leichtman AB. Geographic Variability in Access to Kidney Transplantation in the United States, 1996-2005. Am J Transplant 7(5):1412-1423, 2007.

transplant center. Initiation of dialysis, however, is determined based on candidate medical factors.

The intent of including dialysis time in the allocation system is to provide hope to candidates that their opportunity for a transplant can improve with time. The intent is also to ensure that the allocation system is based on objective medical criteria, as required by NOTA and the OPTN Final Rule, and rate of disease progression.

f. There currently are extreme variations in waiting time until transplantation around the country by DSA. In the LYFT/DPI model, it is likely that these variations in waiting time will translate into variations in age and DPI scores of patients transplanted. Please examine the variation in age at transplantation probability of transplantation by age, and DPI scores across DSA. While the model maybe unstable at the individual DSA level, some estimation of the overall variation can likely be made.

Answer:

The average KAS score among recipients in DSAs with long, medium, and short wait times and discrepancies in access are projected to persist. These results are available on the OPTN website (http://www.optn.org/kars.asp). As was seen with MELD, LYFT and KAS may provide metrics by which differences in access can be quantified,

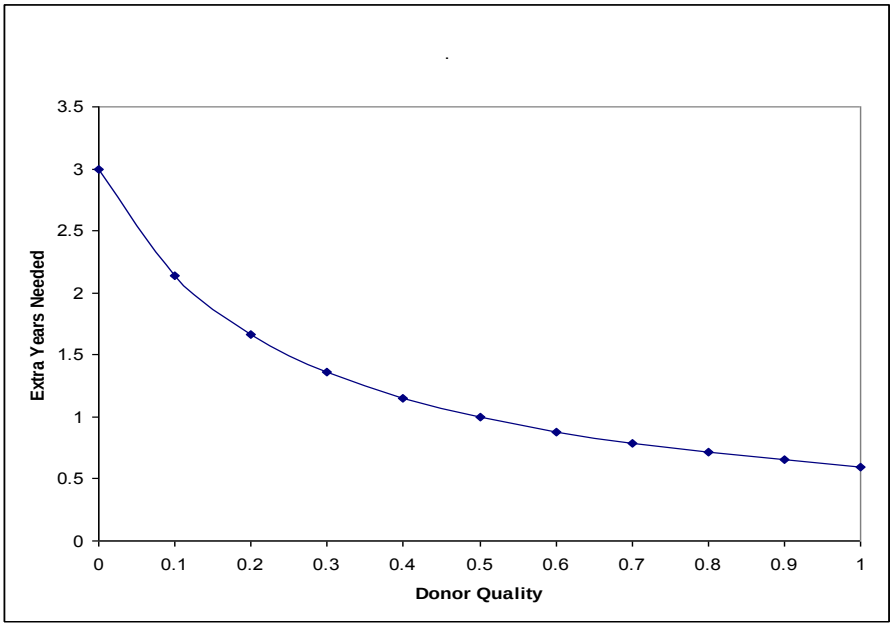
g. Current policy regarding ECD kidneys has a hard cutoff for the DPI that constitutes a higher risk kidney. While we recognize this as an artificial boundary, it does provide for the ability to consent patients up front for the issues regarding outcome for these kidneys. The recent JAMA paper by Merion et al suggests patient characteristics associated with benefit from these kidneys. Given that the OPTN has created a standard of care for kidneys with a DPI of greater than 1.5, how does the committee propose the consent process would work in the future state?

Answer:

Information on the survival differences associated with kidneys of different DPI is given in the table and figure provided in the response to question 3.a above. This survival information can be combined with the expected waiting time in different DSAs for a kidney of a given Blood Type and DPI. Such information can be used by the physician and patient to plan the best strategy for care for each transplant candidate.

Information on the KAS at which transplants occur could be provided to each transplant center on a DSA and blood type specific basis. An example of such a tool is shown below. This tool would allow the center to estimate waiting times for individual patients for kidneys of differing qualities. This hypothetical patient with a given blood type and DSA could receive a high-DPI kidney by waiting just a few months and a medium-

DPI kidney by waiting a year, but would need to wait almost three years to receive the lowest DPI kidney.



Inputs: KAS at which transplants occur within given DSA/blood type, LYFT variables, PRA, Dialysis years

Output: Curve displaying expected future waiting time to receive a kidney of any given DPI

Attachment A

Independent Expert Panel on Statistical Methods for the Analysis of Organ Transplantation Data: Analytic Methods and Simulation Modeling

September 19 through 21, 2004
Hilton Chicago O'Hare Airport

CLOSEOUT REPORT

Submitted by:
Gregory Fant, PhD
Health Statistician
Division of Transplantation (OAB)
Healthcare Systems Bureau
Health Resources and Services Administration
U.S. Department of Health & Human Services
Rockville, Maryland

Healthcare Systems Bureau
Division of Transplantation
FY 2005

James Burdick, MD, Director
Richard Durbin, MBA, Deputy Director

Lorah Tidwell, MPA, Chief, Operations and Analysis Branch (OAB)
CAPT Michael Dreis, PharmD, MPH, Deputy Chief

Branch Members:
Renee Dupee, JD, Public Health Analyst
Gregory Fant, PhD, Health Statistician
CAPT Henry Krakauer, MD, PhD, Medical Officer
CAPT Richard Laeng (ret.), MPH, Public Health Analyst
Monica Lin, PhD, Biostatistician
Ginny McBride, RN, MPH, CPTC, Public Health Analyst
Jade Perdue, MPA, Public Health Analyst
Sherry Whipple, Public Health Analyst
LCDR Hui-Hsing Wong, MD, JD, Medical Officer

Memorandum

December 30, 2004

To: Lorah Tidwell
Chief, OAB

Fr: Gregory Fant
Health Statistician

Re: Closeout Report

Please accept this document as the Closeout Report for the Independent Expert Panel (IEP) project which you assigned to me earlier this year.

This memo briefly presents the background of the project. The attachments to this memo serve not only to document the major features of the project but, also, to present the findings of the statistical experts. The work and recommendations of the statistical experts need no additional comment from me.

Background

The need for assembling a panel of expert statisticians to examine the statistical methods used in the analysis of organ transplantation data was discussed in June 2004. After the July 2004 SAC Meeting, I was directed to convene a panel of experts so that we/HSB/DOT could report to the Scientific Advisory Committee (SAC) on our findings at their October 2004 meeting.

In an accelerated manner, I completed several important activities concerning the planning and execution of the panels that were held in September 2004. The panels came to be known by their short title, Independent Expert Panels (IEPs) on Statistical Methods, and had two sections–a section dealing with analytic methods and another section dealing with simulation modeling. I worked with Sherry Whipple on the logistical details of this project; the logistics contractor was PSA. Michael Dreis, PharmD, MPH, provided me with the overall direction and the mentorship necessary to complete this project.

Excluding the logistical support details (that were, both numerous and essential to the success of this project), I am providing a list of the major steps which I followed in order to convene the IEP: Analytic Methods and IEP: Simulation Modeling in Chicago from September 19 through 21, 2004:

- Identify the content of the sessions
- Identify the dates of and location for panel sessions
- Identify the potential panel members
- Invite the potential panel members
- Invite panel presenters
- Write formal confirmation letter

- E-mail instructions to panel members and panel presenters
- Coordinate the collection of panel presentations
- Work with logistics contractor to arrange conference site, travel, hotel accommodations, pre-conference mail-out, site details and requirements, etc.
- Develop draft panel report suitable for publication
- Brief SAC on results of the Independent Expert Panels

Points of Personal Learning:

- I needed to have a better sense of "due dates" for major tasks and discuss these with contractor in advance. We did not have early, face-to-face meetings and these would have helped. In fact, two, early, face-to-face meetings would have helped.
- I could have distributed to the presenters a template for presentations and required that this template be used for formal presentations. Originally, I wanted the presenters to have as much flexibility as they needed to present their evidence to the panel experts. But the flexibility, in practice, added complications that made the formatting of materials, assembling pre-conference materials, and pre-loading of presentations on the laptop computer unnecessarily time-consuming.
- I, again, realized how essential it was to work with others in a cooperative manner. Certainly, working with Sherry Whipple (as the general supervisor of contract and the point-of-contact for the contractor, PSA), Michael Dreis (as the person who understood the process), Bryan Slattery (the PSA Contractor) were critically important: Without them, it would have been more difficult for me to carry-out this assignment.

Participants

Several people participated in the panels. I have attached a list of all participants along with the biographical statements of the experts panel members for the record (see Attachment A).

ACOT Presentation

I was asked to brief the DHHS Secretary's Advisory Committee on Organ Transplantation (ACOT) on the work of the IEPs during the November 2004 ACOT meeting in Rockville, Maryland. This presentation was intended to provide an overview of the IEPs and sketch the tentative findings which were available at that time. I am attaching a copy of that presentation for the record (see Attachment B).

Consensus Statements

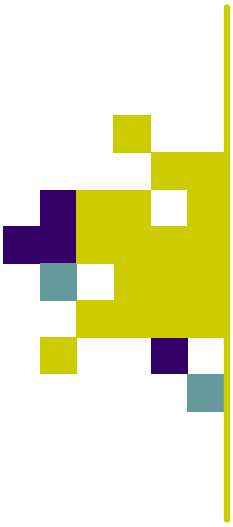
Both IEP Chairs submitted draft consensus documents in October 2004. I received their final consensus statements by mid-December 2004. I am attaching copies of the documents for the record (see Attachment C and D).

Closing Comments

Additional information can be found in the LAN at the following location: H:\Shared\DT\General\IEP Sept 2004

I want to acknowledge my colleague, Kelley Weld, Program Analyst, HAB/DSP, who helped me format this document.

Please let me know how I may be of further assistance in this matter.



Attachment A: List of Participants and Biographical Sketches

Participant Directory

Analytic Methods Panel , September 19, 2004

James Burdick, MD

Director, Division of Transplantation
U.S. Department of Health and Human Services
Health Resources and Services Administration
4350 East-West Highway, 10th Floor
Bethesda, MD 20814
P: (304) 443-7577
Email: jim.burdick@hrsa.hhs.gov

Tom Greene, PhD

Department of Biostatistics and Epidemiology
The Cleveland Clinic Foundation
9500 Euclid Avenue
Cleveland, OH 44195
P: (216) 444-9933
Email: tgreene@bio.ri.ccf.org

Jack D. Kalbfleisch, PhD

Professor and Chair, Department of Biostatistics
University of Michigan - Biostatistics Department
1420 Washington Heights
Ann Arbor, MI 48109
P: (734) 615-7067
F: (734) 615-7068
Email: jdkalbfl@umich.edu

Felicia B. LeClere, PhD

Associate Research Professor
University of Notre Dame
913 Flanner Hall
Notre Dame, IN 46530
P: (574) 631-4528
F: (574) 631-8700
Email: leclere.1@nd.edu

Danyu Lin, PhD

Dennis Gillings Distinguished Professor
University of North Carolina
Department of Biostatistics, CB 7420
Chapel Hill, NC 27599-7420
P: (919) 843-5134
F: (919) 966-3804
Email: lin@bios.unc.edu

David C. Naftel, PhD

Professor of Surgery, Chair
University of Alabama at Birmingham
790 Lyons - Harrison Research Building
Birmingham, AL 35294
P: (205) 934-4240
F: (205) 975-0085
Email: dnaftel@uab.edu

John P. Roberts, MD

Chief, Transplant Services
University of California, San Francisco
Transplantation Program
505 Parnassus Avenue, M896, Box 0780
San Francisco, CA 94143-0780
P: (415) 353-9321
F: (415) 353-8709
Email: robertsj@surgery.ucsf.edu

R. Clifton Bailey, PhD

Statistical Consultant
6507 Divine Street
McLean, VA 22101
P: (703) 893-8719
Email: rcbailey@mac.com

Michael W. Dreis, PhD, MPH

Deputy Chief, Operations and Analysis Branch
U.S. Department of Health and Human Services
Health Resources and Services Administration
Department of Transplantation
4350 East-West Highway
Bethesda, MD 20814
P: (301) 443-8850
F: (301) 594-6095
Email: mdreis@hrsa.gov

Karen A. Eddleman, BS

Scientific Writer/Editor
Logos Scientific Communications
601 North Oak Street
Falls Church, VA 22046
P: (703) 533-3470
Email: keddleman@starpower.net

Gregory V. Fant, PhD

Health Statistician
U.S. Department of Health and Human Services
Health Resources and Services Administration
Department of Transplantation
4350 East-West Highway, EWT Rm 10-1B2
Bethesda, MD 20814
P: (301) 443-8728
F: (301) 594-6095
Email: gregory.fant@hrsa.hhs.gov

Ann M. Harper, BA

Policy Analyst
United Network for Organ Sharing
700 North 4th Street
Richmond, VA 23219
P: (747) 489-8660
F: (757) 489-2029
Email: harperam@unos.org

Henry Krakauer, MD, PhD

Medical Officer
U.S. Department of Health and Human Services
Health Resources and Services Administration
Department of Transplantation
4350 East-West Highway, Room 103D4
Bethesda, MD 20814
P: (301) 594-4374
F: (301) 594-6095
Email: henry.krakauer@hrsa.hhs.gov

Alan B. Leichtman, MD

University of Michigan
Scientific Registry of Transplant Recipients
315 West Huron, Suite 260
Ann Arbor, MI 48103
P: (734) 665-4108
F: (734) 665-2103
Email: aleicht@umich.edu

Monica Lin, PhD

Biostatistician
U.S. Department of Health and Human Services
Health Resources and Services Administration
Department of Transplantation
4350 East-West Highway, 10th Floor
Bethesda, MD 20814
P: (301) 443-2776
Email: mlin@hrsa.gov

Keith P. McCullough, MS

University of Michigan
Scientific Registry of Transplant Recipients
315 West Huron, Suite 260
Ann Arbor, MI 48103
P: (734) 665-4108
F: (734) 665-2103
Email: keithm@urrea.org

Robert M. Merion, MD

University of Michigan
Scientific Registry of Transplant Recipients
2926 Taubman, Box 0331
1500 E. Medical Center Drive
Ann Arbor, MI 48109-0331
P: (734) 936-7336
F: (734) 998-6620
Email: merionb@umich.edu

Friedrich K. Port, MD, MS

President
Univ. Renal Research and Education Association
Scientific Registry of Transplant Recipients
315 West Huron, Suite 260
Ann Arbor, MI 48103
P: (734) 665-4108
F: (734) 665-2103
Email: fport@urrea.org

Douglas E. Schaubel, PhD

University of Michigan
Scientific Registry of Transplant Recipients
315 West Huron, Suite 240
Ann Arbor, MI 48103
P: (734) 998-6611
F: (734) 998-6620
Email: deschau@umich.edu

Bryan D. Slattery, MA

Facilitator and Senior Meeting Planner
Professional and Scientific Associates
2100 Reston Parkway, Suite 300
Reston, VA 20191
P: (703) 234-1734
F: (703) 234-1701
Email: b_slattery@psava.com

Robert A. Wolfe, PhD

Deputy Project Director
University of Michigan
Scientific Registry of Transplant Recipients
315 West Huron, Suite 240
Ann Arbor, MI 48103
P: (734) 998-6611
F: (734) 998-6620
Email: bobwolfe@umich.edu

Simulation Modeling Panel, September 20, 2004

James Burdick, MD

Director
Division of Transplantation
U.S. Department of Health and Human Services
Health Resources and Services Administration
4350 East-West Highway, 10th Floor
Bethesda, MD 20814
P: (304) 443-7577
Email: jim.burdick@hrsa.hhs.gov

Joseph E. Cavanaugh, PhD

Associate Professor
Department of Biostatistics
The University of Iowa
Department of Biostatistics
200 Hawkins Drive, C22 GH
Iowa City, IA 52242-1009
P: (319) 384-5024
F: (319) 384-5018
Email: joe-cavanaugh@uiowa.edu

Gabriel M. Danovitch, MD

Director
Renal Transplant Service
University of California
Los Angeles Medical Center
10833 Le Conte Avenue
Los Angeles, CA 90024
P: (310) 206-6741
Email: gdanovitch@mednet.ucla.edu

Donald T. Gantz, PhD

Professor
Department of Applied and Engineering Statistics
George Mason University
Applied and Engineering Statistics Department
4400 University Drive, MS4A7
Fairfax, VA 22030
P: (703) 993-1695
F: (703) 993-1700
Email: dgantz@gmu.edu

Douglas P. Landsittel, PhD

Assistant Professor Statistics
Duquesne University
Department of Mathematics and Computer Science
600 Forbes Avenue, 419 College Hall
Pittsburgh, PA 15282
P: (412) 396-1419
F: (412) 396-5197
Email: landsittlel@mathcs.duq.edu

Elena Losina, PhD

Assistant Professor of Biostatistics and Research
Massachusetts General Hospital
50 Staniford St., 9th Floor
Boston, MA 02114
P: (617) 724-3341
F: (617) 726-2691
Email: elosina@partners.org

Mark S. Roberts, MD, MPP

Chief
Section of Decision Sciences and Clinical Systems Modeling
University of Pittsburgh School of Medicine
Suite 600, 230 McKee Place
Pittsburgh, PA 15213
P: (412) 692-4826
F: (412) 692-4838
Email: robertsm@upmc.edu

Charles A. Rohde, PhD

Professor
Department of Biostatistics
John Hopkins University
615 North Wolfe Street
Baltimore, MD 21205
P: (410) 955-3539
F: (410) 955-0958
Email: crohde@jhsph.edu

Bruce W. Schmeiser, PhD

Professor
Industrial Engineering
Purdue University
School of Industrial Engineering
Grissom Hall Rm. 228
West Lafayette, IN 47907
P: (765) 491-8665
Email: bruce@purdue.edu

R. Clifton Bailey, PhD

Statistical Consultant
6507 Divine Street
McLean, VA 22101
P: (703) 893-8719
Email: rcbailey@mac.com

Michael W. Dreis, PhD, MPH

Deputy Chief
Operations and Analysis Branch
U.S. Department of Health and Human Services
Health Resources and Services Administration
Department of Transplantation
4350 East-West Highway
Bethesda, MD 20814
P: (301) 443-8850
F: (301) 594-6095
Email: mdreis@hrsa.gov

Karen A. Eddleman, BS

Scientific Writer/Editor
Logos Scientific Communications
601 North Oak Street
Falls Church, VA 22046
P: (703) 533-3470
Email: keddleman@starpower.net

Erick B. Edwards, Jr., PhD

Assistant Director of Research
United Network for Organ Sharing
700 North 4th Street
Richmond, VA 23219
P: (804) 782-4832
F: (804) 782-4835
Email: edwardeb@unos.org

Gregory V. Fant, PhD

Health Statistician
U.S. Department of Health and Human Services
Health Resources and Services Administration
Department of Transplantation

4350 East-West Highway, EWT Rm 10-1B2
Bethesda, MD 20814
P: (301) 443-8728
F: (301) 594-6095
Email: gregory.fant@hrsa.hhs.gov

Ann M. Harper, BA

Policy Analyst
United Network for Organ Sharing
700 North 4th Street
Richmond, VA 23219
P: (747) 489-8660
F: (757) 489-2029
Email: harperam@unos.org

Henry Krakauer, MD, PhD

Medical Officer
U.S. Department of Health and Human Services
Health Resources and Services Administration
Department of Transplantation
4350 East-West Highway, Room 103D4
Bethesda, MD 20814
P: (301) 594-4374
F: (301) 594-6095
Email: henry.krakauer@hrsa.hhs.gov

Alan B. Leichtman, MD

University of Michigan
Scientific Registry of Transplant Recipients
315 West Huron, Suite 260
Ann Arbor, MI 48103
P: (734) 665-4108
F: (734) 665-2103
Email: aleicht@umich.edu

Monica Lin, PhD

Biostatistician
U.S. Department of Health and Human Services
Health Resources and Services Administration
Department of Transplantation
4350 East-West Highway, 10th Floor
Bethesda, MD 20814
P: (301) 443-2776
Email: mlin@hrsa.gov

Keith P. McCullough, MS

University of Michigan
Scientific Registry of Transplant Recipients
315 West Huron, Suite 260
Ann Arbor, MI 48103
P: (734) 665-4108
F: (734) 665-2103
Email: keithm@urrea.org

Robert A. Wolfe, PhD

Deputy Project Director
University of Michigan
Scientific Registry of Transplant Recipients
315 West Huron, Suite 240
Ann Arbor, MI 48103
P: (734) 998-6611
F: (734) 998-6620
Email: bobwolfe@umich.edu

Friedrich K. Port, MD, MS

President
Univ. Renal Research and Education Association
Scientific Registry of Transplant Recipients
315 West Huron, Suite 260
Ann Arbor, MI 48103
P: (734) 665-4108
F: (734) 665-2103
Email: fport@urrea.org

Bryan D. Slattery, MA

Facilitator and Senior Meeting Planner
Professional and Scientific Associates
2100 Reston Parkway, Suite 300
Reston, VA 20191
P: (703) 234-1734
F: (703) 234-1701
Email: b_slattery@psava.com

David E. Thompson

Senior Analyst
Altarum Institute
3520 Green Court, Suite 300
Ann Arbor, MI 48105
P: (734) 302-4675
F: (734) 302-4991
Email: david.thompson@altarum.org

Larry A. Waisanen

Altarum Institute
3520 Green Court, Suite 300
Ann Arbor, MI 48105
P: (734) 302-4603
Email: larry.waisanen@altarum.org

Biographical Sketches

Analytic Methods

Ex-Officio Panel Member

James Burdick, MD

James Burdick received a medical degree in 1968 at Harvard University Medical School and completed a surgical residency at the Massachusetts General Hospital where he became interested in vascular surgery. In addition, Dr. Burdick completed a clinical research position for 2-1/2 years in NCI/Surg in the U.S. Public Health Service which resulted in his interest in immunology. These things led to a Fellowship in Transplantation at the Massachusetts General, followed by an appointment at Johns Hopkins Medical Institutions, where he had a career doing vascular surgery and abdominal organ transplants for 25 years. His academic work has included laboratory and clinical research in controlling the allograft immune response. Dr. Burdick is Professor of Surgery at Johns Hopkins School of Medicine, was on the Board of the Mid-Atlantic Renal Coalition (Network 5), the Council of the American Society of Transplant Surgeons and is a Past President of the South Eastern Organ Procurement Foundation and Past President of UNOS, the OPTN contractor. In July, 2003 he accepted the position of Director, Division of Transplantation in the Special Programs Bureau, Health Resources and Services Administration, U.S. Department of Health and Human Services.

Jack D. Kalbfleisch, PhD

Jack Kalbfleisch is Professor of Biostatistics and Chair of the Department of Biostatistics at the School of Public Health, University of Michigan. He received a Ph.D. (statistics) in 1969 from the University of Waterloo. Dr. Kalbfleisch's primary research interests include the development of models and methods for analyzing failure time or event history data. Applications of this work arise in many areas including epidemiology, medicine, demography and engineering. Additionally, Dr. Kalbfleisch is interested in modeling and analyzing mixture models and the use of resampling or bootstrapping techniques when an estimating function or equation forms the basis for inference. Dr. Kalbfleisch is a Fellow of the American Statistical Association.

Felicia B. LeClere, PhD

Felicia LeClere is the Director of the Laboratory of Social Research and Associate Professor of Sociology at the University of Notre Dame. She received a Ph.D. (rural sociology and demography) in 1990 from the Pennsylvania State University. Dr. LeClere's research interests include healthrelated issues in sociology, including ethnic neighborhoods, immigrant mortality and income inequality in the United States and differential mortality. Dr. LeClere's areas of teaching expertise include primary data collection and survey methodology, event history analysis, and regression analysis for categorical, censored, and truncated data. From 1991 to 1997, Dr. LeClere was a Federal Civil Servant and, specifically, served as a Health Statistician for the National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services.

Danyu Lin, PhD

Danyu Lin is the Dennis Gillings Distinguished Professor of Biostatisitics at the School of Public Health, University of North Carolina. He received a Ph.D. in 1989 from the University of Michigan. Dr. Lin's primary research interests include the development of statistical models and methods for the analysis of failure time data, with applications to medical studies. He is also interested in other aspects of the designs and analysis of medical studies, including sequential designs, two-stage studies, longitudinal data analysis and health economics. Recently, he is interested in the development of statistical methods for genetic/genomic data. Dr. Lin's areas of expertise include survival analysis methods, longitudinal data analysis, designs and analysis of medical studies and health economics.

Dr. Lin is a Fellow of the American Statistical Association and a Fellow of the Institute of Mathematical Statistics.

David C. Naftel, PhD (Chair)

David Naftel is Professor of Surgery and Professor of Biostatistics at the University of Alabama, School of Medicine. He received a Ph.D. (biostatistics) from the University of Alabama Birmingham in 1979. Dr. Naftel leads the UAB Health Transplantation Team in analyzing databases focused on predictors of early and late outcomes after transplantation. Dr. Naftel provides vital expertise necessary for sophisticated statistical analyses of two, large, multi-center databases. Dr. Naftel is the chapter author of "The Basis for Decision Making in Coronary Artery Disease" In: Ischemic Heart Disease: Surgical Management, Ed. Brian Buxton, et. al. (London: Mosby-Wolfe International, 1999). Dr. Naftel is a member of the Data Advisory Committee of the United Network of Organ Sharing.

John P. Roberts, MD

John Roberts is Chief of the Transplant Service and Professor of Surgery at the University of California San Francisco Medical Center. He received his medical degree for the University of California, San Diego in 1990. Dr. Roberts completed his residency in surgery at the University of Washington in Seattle and a fellowship in transplant surgery at the University of Minnesota. Dr. Roberts is an expert in liver transplant surgery for adults and children and has spearheaded an effort to expand a procedure called "living donor" transplants, one of many efforts to make more donor organ available. Dr. Roberts is the past chair of the Scientific Advisory Committee, a committee established by the SRTR to advise the SRTR staff, the joint OPTN/SRTR Data Working Group, the OPTN Board of Directors and various OPTN subcommittees on issues that. affect clinical and scientific data. Dr. Roberts is a Fellow of the American Surgical Association.

Tom Greene, PhD

Tom Greene is an Associate Staff Biostatistician in the Department of Biostatistics and Epidemiology at the Cleveland Clinic Foundation. He received a Ph.D. (statistics) from the Cornell University in 1985. Dr. Greene works on National Institutes of Health (NIH) clinical trials and is interested in epidemiological methods and multivariate data analysis. Dr. Greene is a member of the American Society of Nephrology, American Statistical Association, and Biometric Society. Dr. Greene is an Associate Editor of the Journal of the American Society of Nephrology. Dr. Greene is a member of the Scientific Advisory Committee of the SRTR.

Simulation Modeling

Ex-Officio Panel Member

James Burdick, MD

James Burdick received a medical degree in 1968 at Harvard University Medical School and completed a surgical residency at the Massachusetts General Hospital where he became interested in vascular surgery. In addition, Dr. Burdick completed a clinical research position for 2-1/2 years in NCI/Surg in the U.S. Public Health Service which resulted in his interest in immunology. These things led to a Fellowship in Transplantation at the Massachusetts General, followed by an appointment at Johns Hopkins Medical Institutions, where he had a career doing vascular surgery and abdominal organ transplants for 25 years. His academic work has included laboratory and clinical research in controlling the allograft immune response. Dr. Burdick is Professor of Surgery at Johns Hopkins School of Medicine, was on the Board of the Mid-Atlantic Renal Coalition (Network 5), the Council of the American Society of Transplant Surgeons and is a Past President of the South Eastern Organ Procurement Foundation and Past President of UNOS, the OPTN contractor. In July, 2003 he accepted the position of Director, Division of Transplantation in the Special Programs Bureau, Health Resources and Services Administration, U.S. Department of Health and Human Services.

Joseph Cavanaugh, PhD

Joseph Cavanaugh is an Associate Professor of Biostatistics at the College of Public Health, University of Iowa. He received a Ph.D. (statistics) from the University of California, Davis in 1993. Dr. Cavanaugh's primary research interests include model selection, time series analysis (discrimination and classification, local self-similarity), and modeling diagnostics (assessing predictive influence). Dr. Cavanaugh co-authored "A diagnostic for assessing the influence of cases on the prediction of missing data" which appeared in the Journal of the Royal Statistical Society, Series D, 50: 427-440, 2001; and "Criteria for linear model selection based on Kullback's symmetric divergence" which appeared in the Australian and New Zealand Journal of Statistics (2003). Dr. Cavanaugh has completed research involving simulation and resampling (bootstrapping, jackknifing, cross validation). Dr. Cavanaugh is a Member of the American Statistical Association.

Gabriel Danovitch, MD

Gabriel Danovitch is a nephrologist, Professor of Medicine, and Director of the Renal Transplant Service at the University of California, Los Angeles Medical Center. He received a medical degree from the University of London in 1968. Dr. Danovitch completed a residency in medicinenephrology at the Soroka Medical Center and a fellowship in medicine-nephrology at the Albert Einstein College of Medicine. Among his ongoing research activities, he is a co-principal investigator on an NIDDK protocol entitled "Genetics of Diabetic Nephropathy in Hispanics" (UO1-DK-057303). The goal of this study is to identify the genes that confer susceptibility to diabetic nephropathy among Hispanics. Dr. Danovitch was co-author of the paper "Prospective, randomized trial of the effect of antibody induction in simultaneous pancreas and kidney transplantation: three-year results" in Transplantation, 77(8): 1269-75, 2004. Dr. Danovitch is board certified in both internal medicine and nephrology.

Donald Gantz, PhD

Donald Gantz is Professor of Statistics at George Mason University. He received a Ph.D. (mathematics) from the University of Rochester in 1974. Dr. Gantz has taught undergraduate and graduate courses in basic statistics, probability, stochastic systems, computer simulation, case studies in applied statistics, and use of the SAS system for statistical analysis, among others. Dr. Gantz's research interests are mathematical economics, applied statistics, flight test analysis, computer performance engineering and capacity planning, computer simulation, and management decision systems. He is an active researcher and practitioner in modeling systems and decision support systems with considerable experience in the development of management decision systems and in litigation related analyses. He has done research, published papers, and made presentations about the relationship between tuberculosis, demographic, and socioeconomic factors in Northern Virginia. This work has involved both statistical and geographic systems analyses. He has also lectured on statistical methods for disease surveillance systems. Dr. Gantz has extensive experience in working with academic biologists, epidemiologists, and educators in public policy analysis.

Douglas Landsittel, PhD

Douglas Landsittel is an Assistant Professor of Statistics at Duquesne University. He received a Ph.D. (statistics) from the University of Pittsburgh in 1997. Prior to his appointment at Duquesne University, Dr. Landsittel was a Research Assistant Professor in Biostatistics at the University of Pittsburgh Cancer Institute. Dr. Landsittel's research interests include inference with neural networks, high dimensional data analysis, simulation and computational methods, early phase clinical trials, and analysis of cancer biomarkers. His methodological research interests include quantifying model complexity and conducting likelihood inference with neural networks. This includes characterizing the null distribution of the likelihood ratio statistic and deriving a generalized measure for degrees of freedom. A recent commentary in the American Statistician (Landsittel, et al., 2002, 56: 337-338) suggested using the recently derived "range of influence statistic" by Fay (The American Statistician, 2002 56: 5-9) as a dichotomous analogy to Ye's Generalized Degrees of Freedom (JASA, 1998, 93:120-131). Past areas of applied research (at CDC/NIOSH) have included evaluation of intervention effectiveness in workplace settings, dose-response relationships between occupational exposures and pulmonary toxicity, modeling quantitative structure-activity relationships in the area of skin permeability, and risk assessment methods for occupational disease.

Elena Losina, PhD

Elena Losina is Assistant Professor of Biostatistics and Research Assistant Professor of Medicine at Boston University, School of Public Health. She received a Ph.D. (biostatistics) from Boston University in 1999. Dr. Losina teaches Generalized Linear Models class for graduate students. Dr. Losina's research activities focus in two methodological areas: 1) application of decision analysis models to the analysis of clinical and economic long term implications of specific treatment or prevention strategies and 2) statistical analysis of data collected longitudinally or via complex hierarchical surveys design. She has been awarded a Career Development Award from NIH to undertake studies in model based evaluation of optimal sequential treatment strategies for HIV infected patients who have failed initial treatment regimen. Dr. Losina also received Investigator Award from American College of Rheumatology to examine long-term impact of regionalization of total knee replacement surgery. She is actively working with fellows and junior faculty specializing in cost-effectiveness analyses, health services and outcome research in rheumatology, orthopedics and infectious diseases (HIV/AIDS). Dr. Losina is a Co-Principal Investigator and lead biostatistician in multidisciplinary Musculoskeletal Research Unit and a Principal Statistician of the CEPAC (Cost-Effectiveness of Preventing AIDS complications) research team.

Mark Roberts, MD, MPP (Chair)

Mark Roberts is the Associate Professor of Medicine and Chief of the Section of Decision Sciences and Clinical Systems Modeling, Division of General Medicine at the University of Pittsburgh. He holds a secondary appointment as an Associate Professor of Health Policy and Management and Industrial Engineering. He received a medical degree from Tufts University School of Medicine and the degree Master of Public Policy from the Kennedy School of Government at Harvard University. Dr. Roberts' primary research interest is in quantitative methods and decision sciences, and how they are used to answer questions that are difficult to structure in a standard randomized controlled trial format. His work has been a combination of methodology development and the application of those methods to solve real problems. He applied Monte Carlo simulation methods to solve large Markov processes, and has investigated the effect of various statistical estimating techniques on the stability of the solution to decision analysis models. His current work involves the mathematical modeling of diseases, and the use of those models to understand the optimal timing of technologic intervention in chronic disease. He is currently involved in the development of models of HIV disease that incorporate the development of viral resistance, Pelvic Inflammatory disease, and is using discrete event simulation coupled with quantitative disease progression models to understand the effect of organ allocation schemes on liver transplantation outcomes. Dr. Roberts is a Fellow of the American College of Physicians.

Charles Rohde, PhD

Charles Rohde is a Professor of Biostatistics at the Johns Hopkins University Bloomberg School of Public Health; Dr. Rohde served as chair of the department from 1981 to 1996. He has a Ph.D. (statistics) from North Carolina State University in 1964. Dr. Rohde's teaching expertise is in the area of advanced generalized linear models where he reviews linear algebra and develops the least squares approach to linear models through projections. Topics include linear estimability, the Gauss Markov theorem, and distribution theory under normality assumptions. He develops principle of conditional error for testing linear hypothesis and discusses connections with maximum likelihood and applies those theories to standard experimental designs. Presents random and mixed effects models, and best linear unbiased prediction. Introduces methods of statistical learning, including multivariate adaptive regression splines, classification and regression trees, boosting and bagging. Dr. Rohde's areas of expertise include biostatistics, environmental statistics, generalized linear models, linear models, and multivariate analysis. Dr. Rohde is currently researching the interactions present in generalized linear models that arise in the context of analysis of variance, log-linear models, and logistic models; he is investigating a unifying formulation. Dr. Rohde is a Fellow of the American Statistical Association.

Bruce Schmeiser, PhD

Bruce Schmeiser is a Professor of Industrial Engineering at Purdue University. He received a Ph.D. (industrial and systems engineering) from the Georgia Institute of Technology in 1975. His research interests include probabilistic and statistical aspects of digital-computer simulation: random-variate generation, input modeling, output analysis, variance reduction, root finding, and optimization; stochastic models; and applied operations research. Dr. Schmeiser has been a consultant to At&T Bell Laboratories, Pritsker Cor-

poration, United Nations Industrial Development Organization, Medical Decision Making, Inc., Symix System, and United Network for Organ Sharing. Dr. Schmeiser is a long-time participant in the Winter Simulation Conference, a member of the Institute of Operations Research and the Management Sciences (INFORMS), and a Fellow of the Institute of Industrial Engineers.

Attachment B: ACOT Presentation

Independent Expert Panel on Statistical Methods for the Analysis of Organ Transplantation Data: Analytic Methods and Simulation Modeling

Healthcare Systems Bureau, Division of Transplantation
Health Resources and Services Administration
U.S. Department of Health and Human Services

Presentation for the
Advisory Committee on Organ Transplantation

Gregory Fant, Ph.D.
Health Statistician

Rockville, Maryland
November 4, 2004

Outline

- Background
- Independent Expert Panel (IEP) : Description of the process
- Highlights of the draft IEP Comments
- Next Steps: HRSA

Background

- For ACOT to forward recommendations to the HHS Secretary on matters of…
  - Organ donation
  - Organ allocation
  - Effective and equitable operation of a national, transplantation system
  - Increasing public confidence in the integrity and effectiveness of the national transplantation system

  …requires current and reliable analytic data from which decisions can be made.

Background

- The natural part of providing analytic data is to carry-out actions to assure that the analytic results are current, correct, and reliable.

Background

- Assuring that the analytic results are current, correct and reliable depend on several parties:
  - ACOT
  - OPTN
  - HRSA
  - SRTR

Background

- HRSA is committed to assuring that the analytic results relied upon in the decision-making process are current, correct, and reliable.
- This commitment is on-going and one that HRSA shares with the entire transplant community.

IEP: Description of the process

- As part of our on-going commitment to assuring that analytic results are current, correct, and reliable, HRSA convened a panel of independent statistical experts to assess salient statistical issues inherent in analytic results used in decision-making.

IEP: Description of the process

- The agenda for each panel outlines the major topics discussed (please see the agenda for details).
- The members of each panel are experts in advanced statistical data analysis and simulation modeling (please see the biographical statements for details).

Highlights of the draft IEP Comments
IEP: Analytic Methods—current themes

- IEP: The panel noted that semi-parametric and fully parametric survival analysis techniques may be appropriate for a given research question when used in the hands of statistical experts.

Highlights of the draft IEP Comments
IEP: Analytic Methods—current themes

- IEP: The Cox proportional hazard model (and its extensions), as a semi-parametric survival analysis technique, is a standard statistical method which has the flexibility and the advantage of being able to address over 95% of survival analysis needs that include the consideration of risk factors.

Highlights of the draft IEP Comments
IEP: Analytic Methods—current themes

- IEP: The panel agrees with the SRTR and HRSA of the need to better define concepts such as "benefit of transplantation," candidate ranking, etc. before selecting a statistical method and noted that more than one statistical technique may be appropriate.

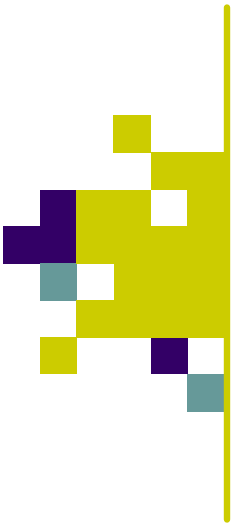
Highlights of the draft IEP Comments
IEP: Simulation Modeling—current themes

- Topic: Should changes in a patient's state (e.g. died, transplanted, removed from list) be drawn from the histories of existing patients? or by probabilistic models?

- IEP: The recommendation is that the disease progression component of a simulation model should try to model the underlying biology of the process which is the path followed by the SRTR.

HRSA: Conclusion

- The assurance of the quality and validity of data analysis is a continuous process that involves:
- ACOT
- OPTN Committees
- HRSA
- SRTR



Attachment C: IEP - Analytic Methods Consensus Document

Analytic Methods: Final Consensus Statement

In the summer of 2004, HRSA created the **Independent Expert Panel on Statistical Methods Data: Analytic Methods.** The members consisted of John D. Kalbfleisch, Felicia B. LeClare, David C. Naftel (Chair), John P. Roberts, Tom Greene, James Burdick (Ex-Officio). The charge of the Panel, as described by Dr. Burdick is: 1) We seek your advice n statistical techniques and analytic options. 2) We want you to weigh-in on various analytic approaches employed: merits, recommendations, suggestions. 3) If a given approach is inappropriate, we want you to tell us. If there is latitude in terms of the approaches used, we want you to tell us that, too. 4) Results will be reported to the SAC and the OPTN. 5) Results will also be published.

On September 19th, 2004 the Panel met in Chicago with representatives from SRTR and HRSA. The statisticians from each group presented their approaches to a number of transplant related analyses. This report is the Panel's evaluation of the methods and contains recommendations to SRTR.

I. Key Objectives

A. How to measure the survival benefit associated with transplantation?

The broad question "What is the survival benefit of transplant" has been formulated in the following three ways:

1. What is the survival benefit of transplant compared to remaining indefinitely on the wait list at the current level of comorbidity?
2. What is the survival benefit of transplant compared to remaining indefinitely on the wait list with the expected progression of comorbidity?
3. What is the survival benefit of transplant compared to remaining on the wait list until a future transplant is offered according to the allocation rule which is currently in effect?

Each of the three more precisely worded questions is associated with different "parameters" which can be estimated statistically. In the view of the Expert Panel, all three formulations may be relevant for different perspectives, and it does not appear useful to designate one of the three formulations as the correct one to the exclusion of others.

It is the Panel's understanding that SRTR has primarily used the first formulation to guide statistical analyses of survival benefit, but that SRTR's newly developed sequential stratification methodology will allow them to use the second formulation once this technique receives statistical peer review. The panel feels that this strategy is appropriate, with the proviso that the formulation that is used in a particular publication or presentation should be explicitly stated.

Some members of the Panel noted that the comparison to "remaining indefinitely on the wait list at the current level of comorbidity" may be artificial because in fact the comorbidity level does change over time. Thus, assuming peer review is positive, the anticipated ability of SRTR to address transplant benefit relative to remaining on the waitlist with progression should be a major advance.

It is the understanding of the Expert Panel that there is disagreement over which of the three formulations should be used to rank wait-listed candidates in order to maximize the total patient-years of survival for the population. Our understanding is that this question cannot be decided statistically, and probably can only be adequately addressed by simulation modeling for the dynamic real-world situation where organs become available sequentially over time. Moreover, it is likely that none of the three statistical formulations actually provides the optimal ranking from the perspective of maximizing total survival time.

While the question of which formulation leads to a better ranking of candidates must be addressed by modeling, the Panel noted that it appears awkward to rank candidates based on Formulation (3) due to the endogenous nature of the parameter being estimated. Specifically, once Formulation (3) is used to rank candidates, the allocation rule will have changed, so the allocation rule used to define the ranking will no longer be in effect. This would necessitate a new ranking, leading again to a new allocation rule, and so on. Underlying this is a dynamic programming problem that would be very difficult to formulate or to solve. It should also be noted that use of Formulation (3) would seem to require consideration and modeling of decisions of patients as to whether or not to accept an offered organ. The SRTR presentation raised some excellent questions regarding the appropriateness of allowing organ allocation to depend on the perceived probability of an individual's acceptance.

B. How to measure other components of the burden of disease and how to combine them into an overall measure to permit the estimation of the overall benefit afforded by transplantation?

The SRTR statisticians have focused on separate analyses of various components. The advantage of this approach lies in the simplicity of analysis and interpretation. One may gain statistical efficiency and biological insights by considering several components simultaneously. There are several methods for simultaneous inference in the statistical literature, some of which are available in standard software. The HRSA statisticians have proposed a joint model for the simultaneous inference. This model is highly parametric, and the robustness of the corresponding inference procedures to the violation of the parametric assumptions is unclear. In order for such a model to be recommended for use by SRTR, it would be imperative for a paper (describing the model, theory and applications) to be reviewed and published in a mainstream statistical journal.

C. How to measure the progression of the condition of the patient on wait list, both with respect to survival and with respect to the other components of the burden of disease?

The panel concluded that not enough information was presented on this complex question to enable the panel to make a judgment on which approach was most appropriate. Some aspects of this question are discussed in A above.

D. How to accommodate changes in allocation rules in predicting the benefit due to transplantation?

This question perhaps would be better directed to the Panel on Simulation. However, there is not a clear line between methods for statistical analysis of data and methods for predicting future outcome in a group of patients. Predictive statistical models can be used in simulation predictions where the value of an independent variable (i.e. risk factor) is modified to estimate the end result. For example, suppose a proposed change to the allocation system is to mandate that patients who are listed for transplant and have blood type O will have the first chance at receiving a donor with blood type O. If recipient blood type O is a risk factor

for delayed transplant, then the magnitude of the delay can be estimated by statistically removing the effect of this risk factor. These predictions can be applied to a group of blood type O patients along with an estimate of the increased availability of blood type O donors to estimate one effect of an allocation policy change.

Neither the SRTR statisticians nor the HRSA statisticians gave a full explanation of their approach for this question. Nevertheless, the methods employed by both groups could be employed to answer the questions.

E. How to reconcile the need for objective rules in the development and application of policy decisions on allocation with the subjective patient preferences that determine acceptance of offers?

The panel voted to table this question.

II. Specific Statistical Issues

A. Estimated allocation scores probably do not lead to perfect ordering of candidates. Should allocation scores be rounded to cause ties among candidates in order to account for uncertainty?

We agree that estimated allocation scores do not lead to perfect ordering of the candidates. This is true for at least two reasons:

- The uncertainty of the model (expressed by both the variance associated with the regression coefficients and the unexplained variance).
- The entire strategy for selecting possible risk factors for the allocation score may have "missed" some key clinical variables.

To quantify (some of) the uncertainty of the estimated allocation scores, one could calculate the confidence interval for an individual. While the width of these confidence limits will depend on the values of the risk factors for an individual, it may still be true that they tend to have a relativity consistent width. One could use this as a guide for a rather gross rounding effort.

Alternatively one could not round at all and say the estimated allocated score is the best statistical guess at a proper ranking of the patients. It was the view of the panel that this was more sensible than an arbitrary grouping to force ties.

B. Are SRTR models flawed because they look at groups of patients rather than at individual patients? In general, are patient-specific predictions appropriate or useful?

The panel was uncertain as to how to interpret this question. The SRTR approach is to incorporate differ-

ences between patients by using measured covariates to modulate the failure rates, either through stratification or through regression modeling. This approach is consistent with approaches generally used in scientific studies, and it does yield the same predictions and decisions for all individuals who share the same covariate values. There are always unmeasured covariates that will affect patient outcomes, but because these covariates are unknown it is not possible to use them systematically in a statistical analysis. The general approach of identifying measured covariates that are informative and using them in the statistical analysis seems appropriate.

C. How should goodness of fit of a survival analysis model be measured? Can the bootstrap identify the impact on estimated coefficients or death rates of future changes in the population?

There exist a number of methods for assessing the goodness of fit of a survival model. The recent release of SAS 9.1 includes many of these methods. Cross-validation is a useful approach to assessing the predictability of the model. The bootstrap is suitable for assessing the precision of parameter estimates under a particular model, but does not address the goodness-of-fit of a model.

D. Parametric versus Semi-Parametric

A number of questions posed to the panel were directed toward the relative merits of parametric versus semi-parametric or non-parametric methods of analysis. The parametric analyses were advocated strongly by the HSRA staff and concentrated on the use of the Bailey-Makeham model. On the other hand, SRTR has been primarily utilizing semi-parametric analyses based on the Cox model.

It is the view of the panel that appropriate analyses could be based on either parametric or semi-parametric techniques. Well utilized, either would give a valid approach to inference and to addressing the principal issues in transplantation, and one would not expect to see substantial differences in the conclusion they would support. With the parametric approach, it is important to test whether the assumed class of failure rates in the model provides an adequate description of the data whereas in the semi-parametric approach, there are no assumptions made about these failure rates, and in this regard the model is very flexible. In either model, there are assumptions about the dependence of the rates on covariates and the appropriateness of the regression models need to be assessed. For parametric analyses, there would be many different highly flexible models that could be used and the Bailey-Makeham model would be only one possibility. Other highly flexible models include for example the Generalized F models (see Kalbfleisch and Prentice, 2002), which also allow for parametric description of a variety of baseline rates. In general, parametric models when appropriate lead to some increases in efficiency in estimates, though the increases are generally relatively modest, especially for the estimation of regression coefficients.

It is the view of the panel that there was no convincing evidence that the parametric analyses improved upon the semi-parametric analyses or allowed the development of useful predictions or results beyond those available from the semi-parametric approach. We see no reason for SRTR to change the fundamental basis of its analyses. Nowadays, the Cox model is very widely used in many areas of investigation. The model itself is highly flexible and allows for relatively simple investigation of many different types of covariate effects. This flexibility is achieved through the use of strata to avoid proportionality assumptions or of time dependent strata or covariates to allow for covariate effects that vary over time. There is extremely good and flexible software to fit these models. The approach of SRTR makes use of this flexibility and expands upon it. It seems well in keeping with current practice in the discipline.

The parametric approach to mortality proposed by the HSRA staff, the Bailey-Makeham model, is a highly flexible model which in its full scope allows for covariate modification of three separate parameters. The

interpretation of these separate regressions was not fully developed or considered in the presentations. One might expect high correlations and/or redundancies among the regression coefficients of these various regression components. The strength of a proposal to use this model for the basis of analyses here or in other contexts would be greatly enhanced by the publication of these methods. The proposed regression framework in three separate directions is somewhat novel and a clear evaluation of properties and applications of these methods in the peer reviewed statistical literature could be very valuable.

III. Discussion

Due to the short duration and the format of the September 2004 meeting, the Panel does not have sufficient information to address the appropriateness of the specific statistical analyses conducted by SRTR. However, the general methodological approaches employed by SRTR appear to be appropriate. It is the Panel's opinion that the HRSA presentations did not document any specific shortcomings in SRTR's analyses that could be better addressed by other techniques.

IV. Recommendations

Due to the many complexities and intricacies of analyzing transplant data, the recommendations from the panel are fairly general in nature. It was the Panel's view that the SRTR is doing very good work, and our recommendations are aimed at further developing their expertise and contributions.

The methodologies used by SRTR statisticians are well within the mainstream of current statistical methodology and they should continue to develop and use these methods as the basis of many analyses and approaches.

Statistical methodologies for analyzing time related event data are constantly evolving and improving. SRTR statisticians have been an important part of this evolution. We encourage SRTR to keep abreast of and maintain their involvement in these developments. It is important that this developmental and methodological activity be viewed as an integral part of the mandate funded by the HRSA.

It is apparent that SRTR has devoted considerable efforts to the difficult and important task of making its analyses accessible to clinicians and when possible to patients. We encourage the SRTR statisticians to continue to strive to develop state of the art statistical analyses while at the same time continuing to collaborate closely with clinical colleagues to help in defining important questions and in providing easily interpreted results.

Expert Panel Members: Analytic Methods

Jack Kalbfleisch, PhD

Professor of Biostatistics and Chair, Department of Biostatistics
School of Public Health
University of Michigan

Felicia LeClere, PhD

Associate Professor of Sociology and Director of the Laboratory of Social Research
University of Notre Dame

Danyu Lin, PhD

Dennis Gillings Distinguished Professor of Biostatistics
School of Public Health
University of North Carolina

David Naftel, PhD (chair)

Professor of Surgery and Professor of Biostatistics
School of Medicine
University of Alabama

John Roberts, MD

Chief of the Transplant Service and Professor of Surgery
University of California San Francisco Medical Center

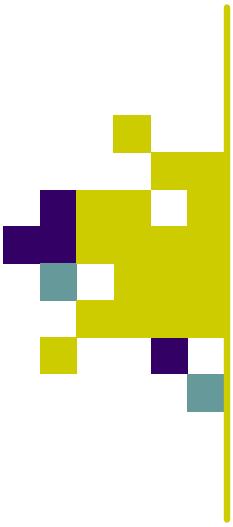
Tom Greene, PhD

Associate Staff Biostatistician
Department of Biostatistics and Epidemiology, Cleveland Clinic Foundation

Ex-Officio Panel Member

James Burdick, MD

Director, Division of Transplantation
Special Programs Bureau
Health Resources and Services Administration
U.S. Department of Health and Human Services



Attachment D: IEP -Simulation Modeling Consensus Document

Simulation Modeling: Final Consensus Statement

The IEP on Simulation Modeling met with members of the SRTR on September 21, 2004, and was asked to comment on several questions related to the design, development and evaluation of the mathematical simulation model used to inform decisions regarding policy changes in the organ allocation system. The following recommendations and discussion represent the panel's opinions on the specific questions that were asked concerning the methods chosen by the SRTR for the simulation of the organ allocation process. We did not evaluate the specific details of all aspects of the Simulation Allocation Models (SAMs).

Inputs to Simulation Modeling

Parametric modeling of patient inputs (recipients, donor organs) versus sampling from actual data

The issue: Simulation models require the capability of generating representations of patients and organs that will be processed through the model. These "virtual" patients need to faithfully represent the characteristics of the real population across a wide array of clinically important variables. There are two basic methods for generating these patients. The first is to randomly sample the actual existing patient arrivals and use the specific data elements from that patient to initialize the patient in the simulation model. The second method is to use the existing data to generate distributions (either empirical or parametric) and have the model generate pseudorandom multivariate samples from those fitted models..

Assessment and Recommendation: Both methods are commonly used in simulations across a wide array of applications. One advantage of direct sampling is that the model is highly likely to represent the inherent associations that exist in the data (which may or may not represent true dependencies). A disadvantage is that the model can be initialized only with examples of patients who have already been listed, when this clearly does not represent the diversity of patients who could be listed. Parametric sampling has the advantage of being able to generate examples of patients outside of the patients who have actually presented, but it has the disadvantage of requiring large amounts of data to assure that the associations in the data are accurately modeled. Currently, the SRTR SAMs sample actual baseline characteristics, which the IEP believes is an appropriate but less flexible method of initialization of the model. However, if the SRTR chooses to move to parametric modeling, then because the complexity of transplant data does not allow for a complete statistical description of the associations and interactions in the data, the specific dependencies that are modeled need to be chosen a priori with a combination of statistical and clinical oversight.

Tracking the Progress of the Patient

Should changes in state (e.g., death on the wait list, changes in MELD score) be drawn from the histories of existing patients (with weighted sampling) or represented by models of progression?

The issue: Similar to requiring a method for initializing patients, simulation models need a mechanism for "aging" patients over time in the event that they are not transplanted in any given timeframe. As the model seeks to predict changes in outcomes across different allocation and prioritization schemes, patients in the model will be transplanted at different times, some patients will become sicker, some may improve. The

model must be able to determine the level of severity of illness and track the predictors of transplant success and failure over time under different policies.

The previous ULAM model developed by the Pritzker Corporation determined natural history solely by the rate of progression through the existing prioritization scheme. Initially this meant determining the probability of moving from a given priority status (1, 2a, 2b, 3, etc) to any other in a specified time. More recently, with the move to the MELD score, disease progression could be determined by the probability of moving from one MELD score to another in a given time period. One problem with this method is that the probabilities of moving between MELD score are not fully independent from the allocation rule. As the allocation rules are changed, different people will be transplanted at different times, and perhaps different groups of people with different distributions of diseases will arrive at transplant. This may directly alter the rate of progression of disease as the distribution of diseases changes, making the probability of transitioning between MELD scores less reliable. The SRTR SAMs now use actual individual histories of transitions between MELD scores to represent the progression of disease if not transplanted. Furthermore, by using actual individual histories, the current SRTR methodology can incorporate changes in the prioritization scheme; provided the scheme uses variables that are collected in the evaluation process.

The final problem with static transition probabilities is that they do not allow for the incorporation of history. For example, in the case of liver transplantation, a static transition probability matrix would provide the probability of moving from each MELD score to different MELD score in the next time period. However, the likely future MELD states for a current MELD state of 25 are very different if the previous MELD was 20 vs 30. Using sampled natural histories implicitly incorporates the prior history whereas static transition probability matrices do not.

The solution of using actual patient histories does have a limitation, in that an actual history may become inconsistent with the history that occurs in the model. For example, if a patient's real history is being used for a modeled patient, it is possible (likely) that the virtual representation of that real patient will not be transplanted at the same time as the actual patient was transplanted. In fact, the virtual patient may not even receive a transplant when the real patient did, and visa versa. The current version of the SRTR allocation models then searches for a similar patient to fill out the remaining history, which is appropriate.

An alternative is to develop models of the underlying clinical and biological characteristics that would allow the model to calculate membership in any prioritization scheme. For example, if the model had a mechanism for predicting the changes in the patient's bilirubin, creatinine, International Normalized Ratio (INR) and whether the patient was on dialysis, the MELD score could be calculated at any time. Empirically, the current SRTR models do that by tracking actual patient histories. If a reasonable number of variables were tracked, this would allow the model to make predictions across prioritization schemes as well and allocation rules. Because it is more complicated, however, the actual modeling of the biological and clinical characteristics is less prevalent in simulation. The Freedberg *et al* CEPAC HIV model is an excellent example of this type of biological modeling in a clinical situation.

Assessment and Recommendation: The IEP was unanimous in their recommendation that to the extent possible, the disease progression component of a simulation model should try to model the underlying biology of the processes. Sampling of actual patient histories is a reasonable method to accomplish this, and avoids the problem of assuming that transition probabilities are Markovian, and do not depend upon history. However, the IEP also notes that as long as the current system used by UNOS to assign prioritization is stable, the transition probabilities felt to be history independent, and as long as the transition matrices for moving between prioritization (such as MELD or PELD for liver transplantation) are updated often, the simulation model should be able to predict effects of allocation changes for certain allocation questions for short periods of time in the future. We also recommend the use of sensitivity analysis to evaluate the effects of different modeling methods on outcomes.

How should events post transplant (graft failure, death with functioning graft, re-transplant) be generated?

The issue: Post transplant events are important to simulation models for several reasons. The most pressing of these is that characteristics of the allocation scheme can alter the level of illness of patients who are transplanted, which (more for some organs than others) may change the graft failure rate. As the graft failure rate increases, the need for retransplantation grows, which places additional demands on the supply of organs. Modeling post transplant events is more complicated than traditional survival analysis, as they involve "competing risks"; the patient may survive, die or experience organ failure (rejection) and retransplantation. The modeling of these events is also complicated by the fact that some of those events are actually the result of the workings of the model (termed "emergent properties" in the simulation literature). For example, while the *need* for retransplantation (graft failure or disease recurrence) can be statistically modeled, as can the probability of death, the actual *event* of retransplantation must be determined by the current allocation mechanism being evaluated by the model. A trivial example illustrates this point. Suppose the rates of re-transplantation are estimated with regression equations, but the allocation mechanism being evaluated by the model is testing the effect of dramatically lowering the priority of re-transplantations on the list. The statistical prediction methodology would provide the wrong estimate of the effect of such a policy change. Calculating the re-transplants through the actual execution of the model would correctly incorporate that policy change.

Assessment and Recommendation: The IEP agrees that post-transplant events are important to model, and that simultaneously predicting patient and graft survival is crucial. Patient survival is important for predicting the overall benefit to transplantation and graft survival is important because it feeds back into the overall allocation system as a need for retransplantation. However, we would argue that it is incorrect to specifically model actual post-transplant retransplantation rates: the retransplantation rate should be determined by the execution of the simulation model (as it is in the current SRTR model). Statistical prediction is appropriate for the need for retransplantation (graft failure) but not the event of transplantation itself. Retransplantation should be an "emergent property" of the model that is calculated from the execution of the model itself. Finally, the IEP recommends that the SRTR consider alternative prediction methods such as neural networks and/or classification trees to evaluate the effect of different prediction methods on model performance.

Assessing the Impacts of Change

Providing for and assessing the impacts of changes in OPTN rules. How should national rules be represented? How should variances (local rules) be accounted for?

The issue: The IEP had some difficulty understanding the exact question being posed; the most obvious answer is that a simulation model that represents the US organ allocation system needs to faithfully represent the actual system. If there are local rules that effect who obtains organs, these need to be instantiated in the model. We do understand that there is a tradeoff between granularity and effort. If there are a large number of local variances that only make a difference in very rare cases, one can infer that they may not need to be incorporated, as that local variance has little ability to affect the overall results of an allocation policy change.

Assessment and Recommendation: The IEP argues that any model of the allocation system should faithfully represent the actual rules under which the current (or any proposed system) operates. The current SRTR model has a flexible structure and appears to faithfully represent many of the local variations in transplantation allocation that exist across regions. Any deviation from fully representing an allocation rule should be evaluated by the clinical oversight body to assure that the level of potential outcome differences

implied by representing a less-than-exact version of the allocation system would not be of a clinically concerning magnitude. This is another situation in which sensitivity analysis is important. One can test whether a full, detailed representation of local variances produces different overall outcomes than a less detailed representation.

How should behavioral changes in response to the changed rules be factored into the models?

The issue: We know that the actual "letter" of the allocation rules as they exist is not always followed directly at the local OPO or center level. We know that patients multi-list, and the rates of multi-listing likely respond to changes in allocation rules. These behaviors are real and do affect the numbers of people waiting on lists. However, who does and who does not change their behavior to alter the likelihood of receiving an organ is hard to predict when allocation rules change.

Assessment and Recommendation: The IEP recommends that the base case on the simulation model run should usually assess the impact of an allocation rule change assuming that the rule is applied as designed. However, because idiosyncratic behaviors by both patient and centers are clearly important, we would recommend that the SRTR develop the capability of incorporating hypothesized behavioral reactions to policy changes into its SAM models to be used as a sensitivity analysis tool. We clearly see an advantage to being able to make statements about the magnitude of possible patient and center reactions to allocation policy changes.

Testing and Validation of the Models

What is the relative rate of agreement with historical data and of logical consistency with expectations of the effects of allocation?

The issue: Simulation model verification and validation are critically important to the overall usefulness of a modeling system. However, in any complex simulation, the model will not exactly reproduce the behavior of the real system: there will be differences. Determining how large the magnitude of these differences can be without calling into question the accuracy of the model is not trivial, and is as much dependent upon clinical expertise as it is on modeling or statistical expertise.

Assessment and Recommendation: The IEP agrees that the ability of a model to replicate historical behavior is an important first step in the overall validation of a model. However, a far stronger test is whether the model predicts the observed behavior of a system when the system is changed. For example, the liver allocation SAM should be able to predict the outcomes (waiting times, size of list, etc.) that occurred when the allocation model was changed from the prior status to the MELD, or when the regional preference order was changed with the Final Rule. The ability to predict the effect of an actual change in system behavior increases the belief in the model's ability to predict hypothetical changes in allocation rules. However, the actual magnitude of what constitutes "agreement" is a complex process that must be determined with input from clinicians and policy makers.

How should the precision of estimates of the predictions be estimated (relative weight to variance of models based on historical data and to standard deviations of replicates)?

The issue: This issue produced a substantial amount of discussion, both within the IEP and between the IEP and the presenters. The fundamental issue is that a model (if run on sufficiently many replications) is essentially a deterministic system. It will always asymptotically arrive at the same answer if started with the same initial conditions. Therefore, it is always possible to state that a particular output of the model is "statistically" different between models run under different allocation policies simply by running the simulations for a long enough period of time, and driving the standard errors of the sample average to an arbitrarily small number.

Assessment and Recommendation: The IEP recommends that the overall purpose of comparing model runs be kept in mind when comparing output results. Just as the appropriate design of a randomized clinical trial requires the a priori knowledge of what constitutes a clinically important difference in outcome, simulation model tests should first have a priori knowledge of what level of outcome change would be considered clinically relevant or policy relevant. For example, if changing allocation induced an increase in the retransplantation rate by 1%, would that be considered important? Similarly, alternative allocation policies may change the size of the waiting list and the average time patients remain on that list. The magnitude of change that one would care about producing is important to assess first.

The IEP recommended that although the standard error of the mean is not a useful parameter in evaluating simulation models because its magnitude is so dependent upon the number iterations, the standard deviation of the model's estimate of the random variable of interest does provide a sense of how the model outputs vary with changes in the model inputs. In a simulation model, the standard deviation of the model output represents the variability in the input data. The standard error of the mean (calculated as a function of the number of observations) will tend toward zero as the number of iterations increases. Therefore, the standard error assesses variation due to simulation, which can be controlled by the modeler; the standard deviation assesses variation due to natural dynamics governing the phenomenon, which cannot be controlled. In addition, for some model outcomes, medians mare be more representative of central tendency and interquartile ranges may be more representative of the spread of the model-based result.

We also recommend that the SRTR be more precise in using terminology, and that accuracy, precision and bias have specific statistical meaning that should be carefully adhered to.

Making the Models Useful

How should the models and their results be structured and presented to the "consumers", the OPTN policy-making committees to make them useful and usable (intelligible scenarios and outputs)?

The issue: Simulation models are complex, and their results have the largest possibility for impact if the models themselves are believed by decision makers to represent the actual behavior of a real system.

Assessment and Recommendation: The IEP feels that the current display of model output data from the SRTR makes sense to us. Although there are other outputs that are not currently considered that we feel are important (costs and quality of life), these are not apparently in the purview of the mandate to the SRTR. However, we also would argue that we are not the appropriate group to decide if the model outputs and presentation style are appropriate. The real consumer of the model outputs is the SRTR Scientific Advisory Committee.

Expert Panel Members: Simulation Modeling

Joseph Cavanaugh, PhD

Associate Professor of Biostatistics
College of Public Health, University of Iowa

Gabriel Danovitch, MD

Professor of Medicine, and Director of the Renal Transplant Service
University of California, Los Angeles Medical Center

Donald Gantz, PhD

Professor of Statistics
George Mason University

Douglas Landsittel, PhD

Assistant Professor of Statistics
Duquesne University

Elena Losina, PhD

Assistant Professor of Biostatistics and Research Assistant Professor of Medicine
Boston University, School of Public Health

Mark Roberts, MD, MPP (Chair)

Associate Professor of Medicine, Health Policy and Management and Industrial Engineering
University of Pittsburgh School of Medicine

Charles Rohde, PhD

Professor of Biostatistics
Johns Hopkins University Bloomberg School of Public Health

Bruce Schmeiser, PhD

Professor of Industrial Engineering
Purdue University

Ex-Officio Panel Member

James Burdick, MD

Director, Division of Transplantation
Special Programs Bureau
Health Resources and Services Administration
U.S. Department of Health and Human Services