UNOS DONATE LIFE
UNITED NETWORK FOR ORGAN SHARING

Working together. Saving lives.

700 North 4th Street, Richmond, VA 23219
P.O. Box 2484, Richmond, VA 23218
tel: 804-782-4800
fax: 804-782-4816
www.unos.org

Brian M. Shepard
Executive Director & CEO

TO: UNOS STAR File Requestor

FROM: UNOS Research Department

RE: Standard STAR Files and Documentation

Date: September 2014

In response to your request, enclosed is a copy of the UNOS Standard Transplant Analysis and Research (STAR) File. We hope you will find this a useful tool in your research endeavors.

Please note that the format of the STAR files changed effective 04/01/2013. Please read this letter for important details about the new format.

****On the DVD, please look for a file called STAR File Documentation.xls located within the "IMPORTANT DOCUMENTATION" folder. This is your guide to using the STAR files and includes extensive documentation about each dataset and each variable.**

The following is provided within the STAR File Documentation.xls file:

- **a detailed description of the dataset and variable names and labels**
- **where the data has been collected, unless calculated (FORM) – see README DATA COLLECTION.pdf**
- **the time period during which the variable was collected (VAR START DATE, VAR END DATE)**
- **the format that belongs with the coded variable (SAS Analysis Format)**
- **any comments or additional information for particular variable (COMMENT)**

We recommend that you start by reviewing the documentation to familiarize yourself with the variables and formats that are most pertinent to your analysis. Keep in mind that many variables have been collected only during a portion of the entire period. As such, it is important to include only those variables collected for the entire period of your analysis.

The information contained in the DVD incorporates a 2 month lag in reporting to improve the completeness and accuracy of the data. Therefore, data files created in June 2014 will include data for registrations added and/or transplants occurring through March 2014.

Each main organ-specific file (i.e. KIDPAN_DATA, LIVER_DATA, etc.) contains information on all waiting list registrations and transplants of that organ type that have been listed or performed in the U.S. and reported to the OPTN since October 1, 1987. It includes both deceased and living donor transplants. There is one record per waiting list registration/transplant event, and each record includes the most recent follow-up information (including patient and graft survival) reported to the OPTN as of the date the file was created. If a patient was listed for a transplant, but was removed prior to transplant or is still waiting, all the transplant information for that

patient is null (i.e. there are waiting list records in the dataset with no transplant information). Similarly, if a patient received a living donor transplant, and was never on the waiting list, all of the waiting list specific information for that patient is null (i.e. there are transplants in the dataset with no wait list registration info). Waiting list registrations can be selected by choosing records where WL_ID_CODE is not null, and transplants performed can be selected by choosing records where TRR_ID_CODE is not null. If there was a waiting list registration that resulted in a transplant event, neither WL_ID_CODE nor TRR_ID_CODE will be null.

The follow-up portion of the STAR File (i.e. KIDPAN_FOLLOWUP_DATA, LIVER_FOLLOWUP_DATA, etc.) contains one record per follow-up per transplant event. Therefore, in most cases you will find multiple records per transplant. For instance, if a patient received a transplant in January 2001, the graft has not failed, and the patient has not been reported lost to follow up, you should find multiple follow-up records (a record at 6 months, and one at each year after 2001) with the same transplant identification number (TRR_ID_CODE). In the same example, if the patient received a simultaneous kidney-pancreas transplant, you should find follow-up records recorded for kidney alone and follow-up records recorded for pancreas alone until January 2003, and then a single kidney-pancreas follow-up record for each year onward. On follow-ups generated since January, 2003, a single kidney-pancreas follow-up is generated in place of the kidney alone and pancreas alone forms. These more recent data are found in the kidney-pancreas follow-up file. Similar to the lag for reporting registration and transplant information, there is a 2 month lag in reporting follow-up data based on the follow-up form expected date. Therefore, data files created in June 2014 will include follow-up data where the transplant anniversary (six month or annual) was prior to March 2014. The variable for linking all the follow-up data to the main datasets is TRR_ID_CODE.

Similar to the follow-up information explained above, additional waiting list information can be linked by WL_ID_CODE to the main dataset, additional information at the time of transplant can be linked by TRR_ID_CODE to the main dataset, and additional follow-up information can be linked by TRR_ID_CODE to the main dataset and by TRR_FOL_ID_CODE to the follow-up dataset.

These files do not include any patient or transplant hospital identifiers. However, there is an encrypted patient identification number (PT_CODE) – unique to each patient – that allows you to track the patient through multiple waiting list and transplant events.

The organ-specific transplant files include most of the pertinent donor information collected on the corresponding organ donor (living or deceased) such as donor age, donor gender, donor history of hypertension, etc. Additionally, the dataset “DECEASED_DONOR_DATA” contains more detailed information on all deceased donors in the U.S (one record per donor) where at least one organ was transplanted or recovered for transplant and reported to the OPTN since October 1, 1987. This allows the researcher to obtain additional information about each transplanted organ, and also to examine other issues, such as deceased donor organ recovery and utilization. The data can be linked to the corresponding transplants using the field DONOR_ID.

Similarly, the file “LIVING_DONOR_DATA” contains information on all living donors recovered in the U.S. since October 1, 1987. For living donors recovered prior to October 25, 1999, there are limited medical and demographic fields available. The data can be linked to the corresponding transplants using the field DONOR_ID.

The file “LIVING_DONOR_FOLLOWUP_DATA” contains one record per follow-up visit for living donors recovered since October 25, 1999. Prior to that date living donor follow-up information was not collected. For living donors recovered since then, there will be the potential for a six month and one year follow-up on donors recovered October 25, 1999 to February 29, 2008 with the additional two year follow-up form added for donors recovered since March 1, 2008.

The end-user assumes all responsibility for analyses performed with these data. We encourage you to consider the time period during which each field was collected when performing any analysis. Deaths and graft failures may be reported on interim follow-ups, but patients who are alive are more likely to be reported at follow-ups submitted on the anniversary of the transplant. Therefore, performing an analysis too early following the end of the follow-up period may lead to artificially low estimates of survival as deaths may have already been reported but living patients may not have reached their anniversary yet. For instance, if your analysis time point is survival at 1 year for transplants performed between January 2011 and December 2012, we would recommend that the analysis be performed only after the beginning of March 2014 (December 2012 + 1 year to reach analysis end point + 2 month lag).

The data are provided in four formats: SAS datasets, SAS transport files, delimited text files, and SAS export to STATA format. If you have access to SAS, the SAS transport file will be the easiest version for you to use. It contains within it all the necessary formats already created for your use. There is also some sample SAS code provided in a text file named *sample_code.txt* in the “IMPORTANT DOCUMENTATION” folder that will guide you through the process of using the transport file. If you are using SPSS, we suggest using the folder named “SAS Datasets”, as you should be able to import those datasets into SPSS. The format files can be found in the folder named “CODE DICTIONARY – FORMATS”.

When first using the data, you will note that the folders (one for each of the four formats described in the paragraph above) have been compressed (zipped) to fit onto a single DVD. If you are using WINZIP or PKZIP, you can access the data directly in the appropriate folder structure. If you are using SECUREZIP, you will need to “extract” the data prior to using it in order to maintain the desired folder structure. The SECUREZIP wizard can walk you through that process.

Included in your dataset are the results of linkage between the OPTN database and the Social Security Death Master File (SSDMF). If a patient in the OPTN dataset was found in the SSDMF file, the field SSDMF_DEATH_DATE will be populated with the date of death as provided from the SSDMF. It is up to the researcher to determine when and how that information should be incorporated into their analysis. This additional information is NOT included in any of the other patient or graft status fields in the datasets. For instance, the field PX_STAT provides the most recent patient status as provided by the transplant center that is following the patient. If the last patient status from the center was on January 1, 2012 and the patient was alive, but the SSDMF shows the patient as having died on June 30, 2012, all the fields in the database, except for SSDMF_DEATH_DATE, will reflect the patient status of alive on January 1, 2012. Please note that as of November 2011, the percentage of deaths accessible in the SSDMF data decreased significantly due to data release issues with the Social Security Administration.

If you have any questions, please email STARFile@unos.org or call 804-782-2970 and leave a message. (Email is preferred method of contact as the mailbox is monitored daily by Research staff.)

Special Considerations When Using OPTN Data in a Manuscript (including Abstracts) or Presentation

1. **Citing Data Obtained Through the Data Request System.** Please cite in the text or graphic presentation, the date and source of the data. This information can be found on the cover of DVD, and on the DVD label itself. For example, "Based on OPTN data as of October 12, 2012." Additionally, please add the following statement as an acknowledgment: "This work was supported in part by Health Resources and Services Administration contract 234-2005-370011C. The content is the responsibility of the authors alone and does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government."

2. **IRB Approval.** If these data are to be used as part of a presentation or manuscript (including abstracts), you must check with your institution to see whether IRB approval (expedited or otherwise) is required.

Please note: The UNOS Research Department requests that any UNOS staff member who makes a significant intellectual contribution to a study abstract, presentation, or manuscript be offered the opportunity to be included as an author. UNOS staff may not be listed as study authors without obtaining written permission from the appropriate staff member(s).