

CHAPTER II

Transplant Data: Sources, Collection, and Caveats

INTRODUCTION

Collecting, organizing, and disseminating data for research on transplantation involves tremendous effort from professionals at transplant centers and organ procurement organizations (OPOs), as well as other data collection specialists. To put this wealth of data to its best use, researchers and readers of analyses, whether practitioners or patients, should understand the full process of collecting and organizing these data. Familiarity with the data's structures and sources — as well as their limitations — can help ensure that readers and researchers use data effectively and accurately.

We hope to enable better interpretation of research results, sharper awareness of data limitations, and clearer concepts of how new analyses might proceed. By examining the sources, quality, and organization of the different types of transplant data available, we hope to improve the understanding of existing results, help researchers with study design, and stimulate new exploratory initiatives.

The data described here are the source of the figures and tables in the *2003 OPTN/SRTR Annual Report*. These data are used by the SRTR, the OPTN, and a wide variety of other researchers as the basis for reporting on and answering questions about the state of transplantation in the United States. Such topics and questions include the following:

As the basis for reporting on both OPTN and SRTR web sites, helping medical professionals, patients, and families investigate the best options of treatment: *Does this center have a high rate of transplant complications? How quickly might I be allocated an organ if I register at a different center, and are my prospects for survival after transplant there as good?*

As the source for analyses in support of policy-setting activities by the Secretary's Advisory Committee on Transplantation (ACOT), OPTN/UNOS Committees, and other government and nongovernment requesters: *Is a transplant candidate better off accepting an organ from a less-than-ideal candidate or staying on a waiting list? Has the new MELD-based liver allocation system affected waiting list mortality, and has it affected wait-listing behavior of transplant centers? What are the effects of allowing patients to be put on waiting lists at more than one transplant center?*

As a resource for private researchers in search of better treatments for end-stage organ failure: *How does a choice of immunosuppressive therapy affect patient outcomes?*

A researcher must first know which data are available to choose from in answering these questions. The beginning section of this chapter describes the **Scope of Data Available** about transplantation, organizing it into areas of interest such as waiting lists, transplants and posttransplant outcomes, and organ donation. Next, the researcher should know how these data are collected. In last year's *Annual Report* and *Report on the State of Transplantation* we detailed the tremendous evolution in transplant data since the late 1980s (1). This year the second section, **Primary Data Collection**, focuses on current technology and recent developments that yield ever more timely and reliable data.

The third section provides a survey of several **Secondary Data Sources**. These data are used by the SRTR and OPTN in determining data completeness and limitations, as well as augmenting the primary data for new research purposes. In many cases, these data may be available to other researchers.

Finally, the section on **Caveats for Researchers** examines how various types of data may require different methods of statistical analyses. We examine patterns in data submission and consider the effect of secondary data sources on measuring outcomes, such as waiting list and posttransplant mortality.

THE SCOPE OF DATA AVAILABLE

Data Structure and Units of Analysis

This section describes the scope of transplant data by organizing it into tables according to "unit of analysis." Though the examples here are taken directly from the SRTR, they are generic in application and might resemble data organized for similar purposes by the OPTN or any other researcher.

In assembling a data structure for transplantation research, data may be transformed from their original format into one more conducive to analysis. While data are collected originally in a format to facilitate organ matching and waiting list maintenance, Medicare billing, or management of Social Security benefits, these data often need to be summarized or transformed to formats that facilitate survival analysis, description of

the waiting list, or summarization of immunosuppressive medications over time. Extensive “parent-child” organization is useful for maintaining data integrity in applications that keep track of constantly changing values, such as OPTN organ allocation procedures. However, it may make research with these data computationally intensive, and researchers like the OPTN and SRTR reorganize these data extensively to a schema more suited to easily answering research questions (1).

In a data structure geared more toward research, consideration is given to the “unit of analysis” that may be of interest to the researcher when preparing analysis files (or “tables”). Different tables are organized for different research questions, using different units of analysis as rows in each table. Data from many sources and related tables may be summarized and attached to the record of interest. For example, many researchers want to examine transplants (unit of analysis) and posttransplant survival, as in Tables X.9 in each organ-specific section of the data tables in the *Annual Report*. A table in which each row represents a transplant may be augmented with data summarized from the related tables of follow-up sources, such as each recipient’s latest status as alive or dead and the date of that status, or the last time of tracking before being lost to follow-up.

Figure II-1 shows a useful method of organizing these data into “units of analysis,” also showing the breadth of commonly used records of interest and the relationships between them. The table entities in Figure II-1 relate to a specific subject of interest for research: candidacies, donors, transplants, and the components thereof. Also shown are some of the more specialized tables, ones from which the researcher might analyze organ turn-downs, immunosuppression medications used, or changes in status history.

In addition, this figure documents some of the primary and secondary data sources which may contribute to each table. Further detail regarding the specific data collection instruments, before the information is aggregated to records of interest, is shown in Figure II-2.

Candidate Analysis Tables

The **candidates** table shows where the “demand” side of the transplant process starts. This table might be used to answer analysis questions describing the types of patients placed on waiting lists, which of those are successful in receiving a transplant, and how long they remain on a waiting list before receiving an organ.

The candidates table includes patients placed on the waiting list as well as those who receive an organ from a

living donor without having been placed on the waiting list. The information in this table is taken from the waiting list maintenance and the Transplant Candidate Registration (TCR) record completed soon after registration and augmented with secondary data sources that may be of interest to the researcher. For example, a center’s reporting duties for a transplant candidate end upon the candidate’s removal from the waiting list, but events occurring in the months following removal — such as death or transplant at another center — might be interesting outcomes to the researcher. Therefore, a candidate file may incorporate information from additional mortality sources; or waiting list, transplant, and follow-up information reported by other centers for the same person. More about these additional sources is discussed later.

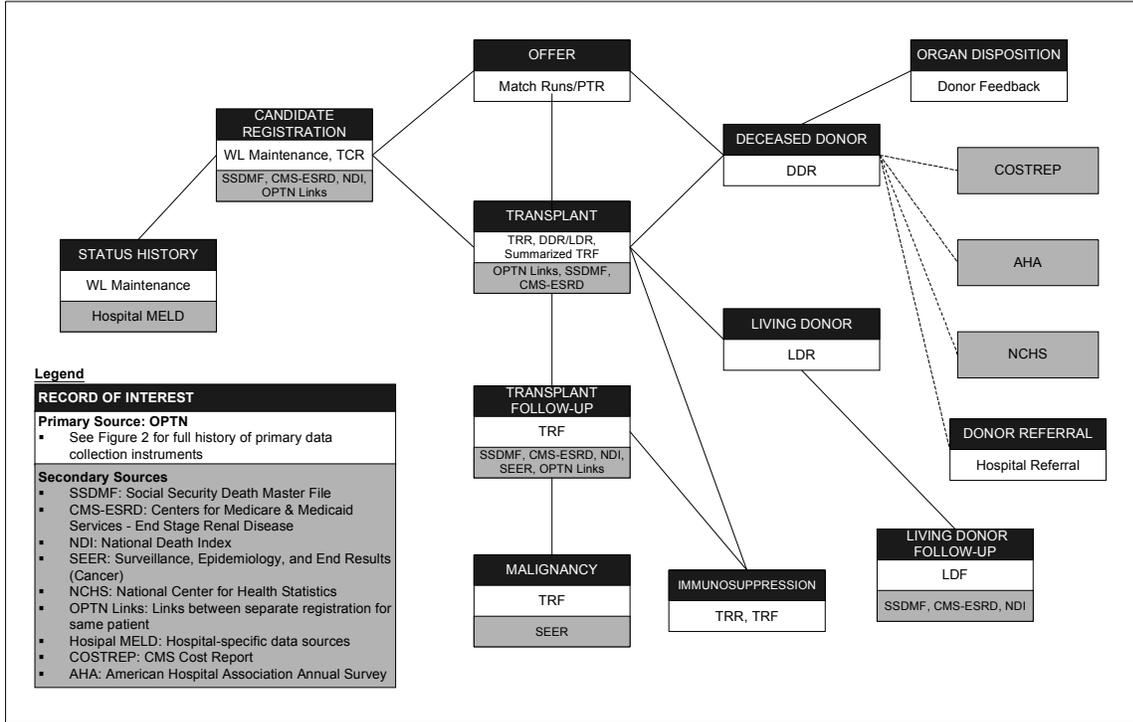
One sub-table, or “child” table, of the candidates file is the **status history** table. Built from an examination of the history of changes to the operational waiting list, this file helps the researcher track the progress of disease during the patient’s stay on the waiting list, along with any other characteristics that change over time. With one record for each patient covering a time period on the waiting list, this table describes such characteristics as waiting list urgency status or MELD at each given time. It allows the researcher to summarize the waiting list at a point in historical time, counting the number of people at a given urgency status in a given region; or to summarize information for a given patient, calculating how quickly her MELD is rising or falling, or how much time has been accumulated in a given waiting list status.

Transplant Analysis Tables

The **transplants** table contains one record for each transplant, including those from both living and deceased donors. These tables include a wide range of data pertinent at the time of transplant, including information about the recipient, the donor, and the transplant operation. This file is used by analysts wishing to characterize trends in the volume and characteristics of patients receiving transplants, as well as analyses examining transplant outcomes.

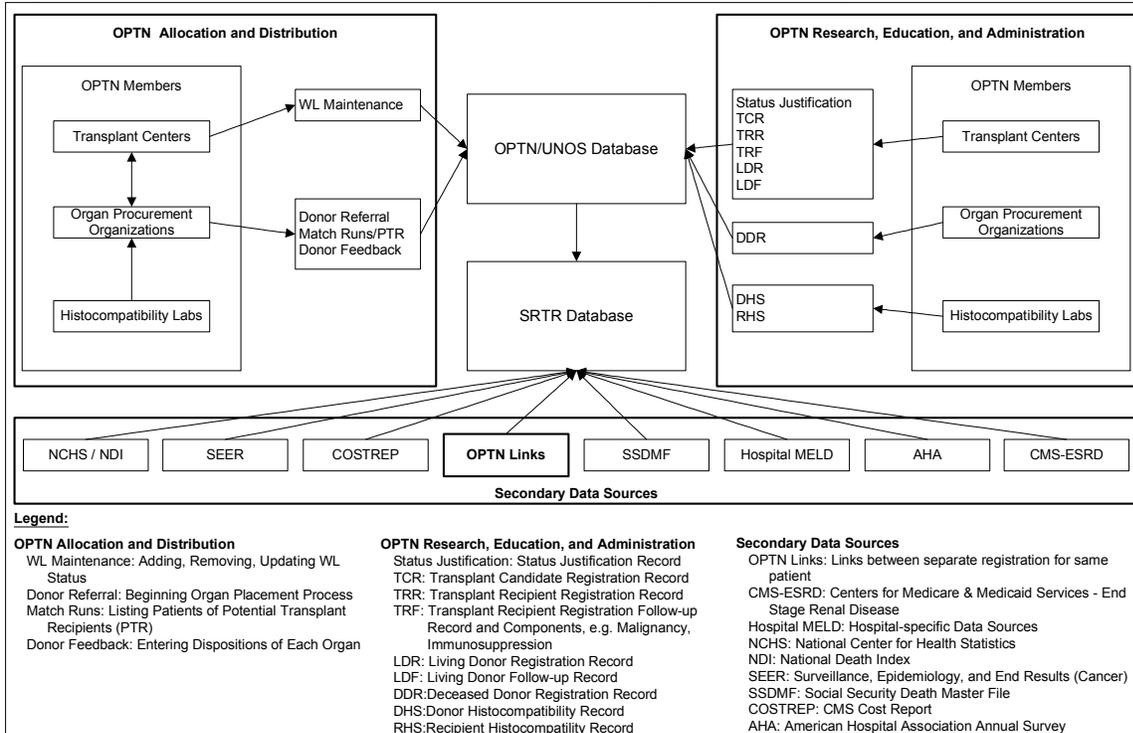
The data for the transplant tables are primarily taken from records collected by the OPTN, discussed in detail below. Additionally, characteristics taken from the donor and candidate files are added for ease of analysis, as are aspects of the interaction between donor and recipient characteristics. Examples include calculated HLA mismatch scores, ABO blood type compatibility, and whether the organ was shared, based on the relationship between the OPO recovering the organ and the transplanting center.

Figure II-1. Transplantation Research Data Organization, Primary and Secondary Sources



Source: SRTR.

Figure II-2. Data Submission and Data Flow, Primary and Secondary Sources



Source: SRTR and OPTN.

Two of the tables shown as linked to the transplant table may also be summarized in the transplant table for ease of analysis. The organ **offers** table is the operational table used for the matching process, recording offers and reasons for organ refusal. On its own, this table may be useful in analyses such as those to find what characteristics are associated with donor-organ-recipient combinations that are accepted more readily than others. In addition, the number of centers or patients that have turned down an organ may be summarized and added to the transplant record as a measure of donor quality.

The **transplant follow-up** data, collected primarily from the Transplant Recipient Follow-Up (TRF) record, may be summarized to the transplant level, creating indicators of death, graft failure, and time of follow-up. They may also be useful on their own — or in conjunction with their own sub-tables for **immunosuppression** or **malignancies** — for analysis of specific events that occur during follow-up. Many external sources are useful in augmenting follow-up data on mortality, graft failure, and tumor incidence, and are discussed below.

Donor Analysis Tables

While much donor-related data pertinent to transplant outcomes is summarized and attached to a transplant analysis file, these data may be used on their own to examine the donation mechanism, or the “supply side” of the transplant process. These data are collected and stored separately for living and deceased donors, not only because of their different scopes but also because the types of analyses, and therefore the secondary data tables, are different for each.

The **deceased donor** table contains one record for each deceased donor with at least one organ recovered for the purpose of transplant. In conjunction with the **donor disposition** table, which stores information about the placement or nonuse of each of the 11 organ types that might be recovered from each donor, analysts might look at the reasons for nonuse of each individual organ, or the number of organs recovered from donors from whom at least one organ was found suitable.

An additional source of donor-related data newly collected by the OPTN is the **donor referral** table. Though not collected for individual donors, these data are reported by each OPO to reflect the number of eligible donors referred by each hospital within the OPO’s Donation Service Area (DSA). These data, combined with actual deceased donor data, are particularly relevant in light of current interest in

increasing the conversion rate among eligible deceased donors. When combined with external sources for hospital characteristics or death record review — such as those from the National Center for Health Statistics (NCHS), American Hospital Association (AHA), the Association of Organ Procurement Organizations (AOPO), or CMS cost reports — these data may shed light on the potential for organ donation and practices that might help target potential donors.

The **living donor** analysis file includes one record per living donor transplant and reflects information collected at time of transplant. As with deceased donors, much of the relevant data is included in the transplant analysis files. The OPTN collects **living donor follow-up** information to facilitate analyses of post-donation outcomes. However, even though many centers submit these follow-ups as required (98% for donors from 2000), the fact that 46% of these records indicated “lost to follow-up” reflects the fact that, unlike the transplant recipients themselves who are well-tracked by the center, donors are often healthy and may not live near the transplanting center, minimizing their contact with the center.

PRIMARY DATA: THE OPTN DATA COLLECTION PROCESS

Most of the data described above are originally collected by the OPTN. This section describes the data collection process implemented by the United Network for Organ Sharing (UNOS) as part of the OPTN contract. UNOS has collected data on all organ transplants and organ donations since 1987, developing extensive expertise in the technology of both organ allocation and data collection. Since the introduction of the UNetsm system in 1999, the data collection process has evolved into an online, user-based data entry, verification, and reporting system. Figure II-3 shows a brief overview of that evolution (1).

UNetsm

UNet is the OPTN’s primary instrument for transplant data collection and verification. Upon its implementation on October 25, 1999, UNet represented a two-and-one-half year, 30,000 person-hour effort by UNOS to update the OPTN information system. Prior to UNet, UNOS fulfilled OPTN functions using a legacy mainframe system that could be accessed by OPTN members via dial-up connection to manage their waiting lists and run the donor/recipient matching process. UNOS undertook the UNet development project in 1997 to meet the following goals:

Figure II-3. OPTN/UNOS Data System Evolution

	1986-90	1990-94	1994-96	1996-99	1999 to Present
	Pre-OTIS		OTIS	OTIS + Tiedi®	UNet SM
Waiting List Management					
Communication	Phone to Organ Center with paper back up and validation. Some facilities use terminal emulation via modem.			Member online (Web-based).	
Donor-Recipient Matching					
Communication	Terminal emulation and modem or phone to organ center and faxed to OPO.			OPO generates online (Web-based).	
Data Collection Forms					
Mode of submission	Paper. Manual data entry at UNOS. Line prompt entry.		Electronic forms added.		Web-based submission. Paper forms phased out.
Submission prompting	Member- initiated.	Electronic events prompt form generation. Forms mailed by UNOS.			Electronic events prompt blank Web-form generation.
Edit checks	Few.		Checks added over this period, data verification reports by mail.		All fields validated electronically. Verification reports by mail.
System					
Storage system	VMS flat files.	VMS relational database	VMS relational database, Lotus Notes.		Microsoft SQLServer Relational Database.
Component integration	None.		Match and forms linked. WL addition initiates TCR.		All systems completely integrated.
Security	One password per center. No encryption during transmission.			User-specific passwords. Full 128-bit encryption.	

Source: OPTN.

Resolve Year 2000 issues with the legacy mainframe system

Increase integration of the allocation and research data collection systems, eliminating parallel systems

Increase member access and functionality in the system

Allow for faster implementation of system changes

Increase system security

Increase the OPTN's ability to utilize emerging technologies

UNOS worked closely with the OPTN membership, including transplant centers, OPOs, and histocompatibility labs, to plan and develop the UNet system. UNOS established an Information Technology User's Advisory Group to meet with UNOS technical staff and provide advice and feedback on the practical considerations for the UNet system from a member perspective.

Description

Accessed over the Internet, UNet is available to OPTN members at all times. The application provides member transplant centers, OPOs, and histocompatibility laboratories with the ability to perform the following OPTN tasks electronically, using a personal computer, Internet connection, and commercially available browser:

Manage the transplant center's list of waiting transplant candidates

Access, complete, submit, and validate OPTN transplant data records

Add donor information and run donor-recipient matching lists

Attach and distribute donor information to facilitate the organ placement process

Access transplant-related data reports, UNOS/OPTN policies, and various other resources

Verify and maintain all data currently and previously submitted to the OPTN

UNet maintains the list of patients waiting for organ transplant (over 82,000 people as of August 20, 2003) for the entire US, performs computer matches for every deceased donor organ to every transplant candidate, and maintains data on every organ donor and transplant event since 1987. Data collected before the implementation of UNet have been integrated into the UNet system for continuity of record keeping.

UNetsm Modules

UNet is made up of several modules, each dedicated to providing a particular service to the OPTN community: Waitlist, DonorNetsm, Tiedi®, Reports, and Resources.

Waitlist. Candidates who are approved for transplantation by a transplant center are added to the national transplant waiting list through the Waitlist section of UNet. This section allows the transplant center to modify information or remove listed candidates. Transplant centers may also maintain their center's organ acceptance information and create various waiting list reports.

Two important data collection processes that occur in the Waitlist section are the capture of initial listing information for an added transplant candidate and the recording of disposition information when a transplant candidate is removed from the waiting list. Upon listing a transplant candidate, key information points are recorded by the UNet system, such as name, Social Security Number (SSN), organ type, age/date of birth, etc., to create a record in the system database for this candidate. This process also generates the Transplant Candidate Registration (TCR) record in the Tiedi section. Many of these data automatically cascade to the TCR from the initial entry of the candidate on to the waiting list.

When removing a transplant candidate from the waiting list, the user must designate the reason for removal in the "recipient feedback" process. If transplant is the reason for removal, the user must enter the Donor ID, organ type and transplant date, to be cross-referenced against the donor information entered in DonorNet by the donor OPO. Additionally, when transplant is the reason for removal, Transplant Recipient Registration (TRR) and Recipient Histocompatibility (RHS) records are generated in the Tiedi module. OPTN Policy calls for patients receiving a transplant to be removed from the national waiting list within 24 hours of the transplant procedure.

Recent emphasis has been placed on including immediate data validation checks for key data points that are considered in the matching process. Several checks were added for the transplant candidate addition and donor entry process in early 2003. In addition, key data points such as ABO were highlighted on the on-screen and printed match results. The OPTN contractor and committees continue to implement real-time verification of the key data that affects organ availability and distribution for waiting transplant candidates.

The majority of the data collected in the Waitlist and DonorNet sections are considered to be "operational" in nature. These data represent point-in-time information that allows the system to match the transplant candidate at their current medical status with donors at the point they are entered and matches are run. Since the medical status of waiting candidates, and even organ donors, can change at any time, transplant centers and OPOs are encouraged to maintain the most current information in UNet for their waiting candidates and donors, thereby allowing the system to produce the most accurate matches between the waiting candidates and an entered donor.

DonorNetsm. UNOS incorporated DonorNet into the UNet system on July 16, 2003. This new system adds several significant enhancements to the primary features previously included in the UNet Placement section. This section of the application allows OPOs to add or modify information on donors and donor organs, initiate the donor-recipient matching process, and record organ placement information. The donor-recipient match process ranks all acceptable, active candidates with the specific information entered for a given donor. The resulting match list is the guideline by which all organs are offered to transplant centers for waiting transplant candidates.

New features added in this section allow OPOs to post donor information in an electronic file format for review by transplant personnel, thus increasing the efficiency of the organ placement process. Such files may include the OPO's donor information form, ancillary confirmatory information such as ABO confirmation documents or serology results, digital images of X-rays of the donor organ, and short video images of echocardiograms, angiograms or bronchoscopies. By viewing these posted source documents (on UNet or via fax), transplant center personnel can reach an informed decision of whether to accept the organ for their transplant candidate.

By adding these new features, the goal of the DonorNet system is to increase the efficiency and accuracy of the organ placement process. In its first month of implementation, 509 donor-related files were attached

on 275 donor records from nearly all of the active OPOs. These files were viewed a total of 1,370 times by personnel from 77 OPTN member organizations. Consideration is being given to enhance the utilization of DonorNet through wireless and mobile technologies that are becoming available.

DonorNet provides transplant center personnel with the following utilities designed to assist in the organ offer and acceptance process:

A section to access and view posted donor information. This section includes a utility to find the sequence number of a candidate on the specified donor match and a link to that candidate's waiting list record.

A Center-Specific Match Results utility that allows the transplant center to view a list of all their candidates by the sequence that they appear on the match. Also included is a Test Match utility that allows the center to see how their candidate would rank given a set of entered donor variables.

A Kidney Acceptance Criteria utility and Organ Offers utility that are currently also located in the Waitlist section.

The posting of donor-related information in DonorNet is not a data collection process. However, OPTN data — such as all organ offer acceptance and refusal information — are collected in the DonorNet section. This information is entered by the OPO or UNOS Organ Center staff, either at the time of organ offer or soon thereafter. OPTN policy directs that this information should be complete within 30 days of the execution of the donor-recipient match run.

Other data collected in DonorNet includes payback accounting, maintained by the Organ Center, as well as the donor referral information described above. Additionally, donor and organ disposition information is entered directly by OPO staff into DonorNet. These data, known as “donor feedback,” records whether organs were recovered from a donor that is entered in the system, the reasons why specific organs may not have been recovered, and if an organ was recovered, which transplant center received the organ. The information entered in the donor feedback section generates the appropriate data collection records. OPTN policy calls for donor feedback to be completed in the system within five working days of the organ procurement date.

Tiedi®. The Transplant Information Electronic Data Interchange (Tiedi®) was originally implemented in 1996 as a software-based transplant data collection alternative to the system of paper-based data forms. In

1999, Tiedi was integrated into UNet and became web-based. Tiedi collects transplant data from every transplant program, histocompatibility lab, and OPO in the country. It is the primary collection system for OPTN data. In June 2002, the OPTN Board of Directors approved mandatory submission of all data electronically through UNet beginning in January 1, 2003.

Tiedi integrates patient-specific information from the time a patient is entered on the national waiting list through the transplant event and follow-up processes, until graft loss or death. The system allows members to electronically report information on transplant candidates, recipients, and donors to the OPTN. Tiedi also allows each member direct access to all data that has been submitted by the member in the past. The goals of the system are to reduce the cost and time necessary on the part of the member to complete OPTN data requirements, to increase required data compliance, and to improve the quality and completeness of the data.

OPTN data records are event-driven: A record is generated when a significant event is reported or attained. The following describes the records generation process.

Transplant Candidate Registrations (TCR) are generated and available after a patient is listed on the national transplant waiting list. A TCR is also generated for a living donor transplant, where the recipient was not added to the waiting list, as reported through the living donor feedback process.

Transplant Recipient Registrations (TRR) are generated and available immediately after a transplant event is reported through the recipient feedback process in Waitlist. As above, a TRR is also generated when a recipient is added through the living donor feedback process.

Transplant Recipient Follow-ups (TRF) are generated in Tiedi at six months, one year, and annually thereafter following transplantation, until either graft failure, recipient death, or loss to follow-up is reported. Exceptions include those for thoracic organs, not requested at six months, and those for kidney and pancreas, which continue for two years after graft failure. A Post-Transplant Malignancies record is generated if one or more malignancies have been reported on the TRF.

Living Donor Registrations (LDR) are generated as soon as the living donor feedback process is completed by the transplant center.

Living Donor Follow-ups (LDF) are generated at six months and one year following living donor transplantation.

Deceased Donor Registrations (DDR) are generated and available as soon as the donor feedback process is completed in DonorNet.

Donor Histocompatibility (DHS) records are generated as soon as the donor feedback process is completed in DonorNet, or for a living donor, after living donor feedback is completed in Tiedi. These records are completed by the lab that performed the donor testing.

Recipient Histocompatibility (RHS) records are generated when a transplant event is reported through the recipient feedback process in Waitlist. In cases of directed living donor transplants, where the recipient was not on the waiting list, the Recipient Histocompatibility generates after living donor feedback is complete. These records are completed by the lab that performed the recipient testing.

Each record is pre-populated with certain previously collected information from waiting list events, transplant events, and follow-up anniversary events. The data coordinator enters the remaining expected data for each record. The system provides built-in range checks and field data selections from drop-down menus. Tiedi includes a component that validates the data on the record, checking certain fields for data completeness, accuracy, and redundancy. This process ensures that data elements collected on multiple records are consistent and valid across records.

For those members who prefer electronic data transfer to manual data entry, Tiedi provides data import and export utilities. Any record may be imported into Tiedi from the member's own database or spreadsheet. All data is exchanged in an ASCII tab-delimited text format. UNet includes tools and information to assist members in developing import files. Recent UNet statistics demonstrate that over 11% of data records received are imported electronically.

Other UNet Features: Reports, Security, and Support

The features of UNet extend well beyond data collection. The Reports section of UNet provides transplant organizations with up-to-the-minute organization-specific data reports, such as center-specific transplant data reports, OPO reports, and donor transplant confirmation reports. With more reports planned, these capabilities provide member organizations a chance to easily review and verify recent

activity, or to export information for reporting use within one organization.

Because the lives of many patients awaiting transplants depend on the continuous availability of the organ allocation system, UNet is implemented on a highly secure system with numerous built-in storage and service redundancies. The application uses a custom-designed security program to monitor access and ensure encrypted transactions. User support is provided by the UNOS Help Desk and from the UNOS Organ Center, which combine for 24-hour coverage of critical processes.

System data gathered in August 2003 showed that UNet is accessed by over 7,000 registered users from 485 OPTN member organizations. User documentation for this wide variety of users is continually updated and easily accessible. In addition, the UNOS Technology Services department coordinates customized training services for OPTN members — at UNOS headquarters, by phone, by visits to member organizations, and through written and on-screen tutorials. Annual user satisfaction surveys have shown a continued high degree of satisfaction with the system's function and support, contributing to the improved rate of timely return of data records since the implementation of UNet.

DATA QUALITY ASSURANCE PROCESSES

Monitoring the accuracy of the data begins with the edit checks and validation during the data entry process, internal verification processes at UNOS, and a collaborative effort of the OPTN and the SRTR. The UNOS Help Desk takes calls from members who find inaccuracies within fields that can only be modified by UNOS staff (e.g., transplant dates and SSNs) and who need records created or deleted. UNOS also performs electronic searches for inconsistencies in the database, generating discrepancy reports, for example, to find improbable combinations of age, height, and weight fields for each patient that should be queried. The SRTR delivers similar discrepancy reports to the OPTN each month to raise further data quality issues. As problems with records are identified, data quality specialists resolve them through UNet and direct contact with the transplant centers; those affecting large numbers of records may be resolved through programmed updates, while others must be addressed individually. Fields in which UNet allows incorrect data entry are identified on an ongoing basis, and UNet edit checks are regularly revised to reduce opportunities for data entry errors.

Database checks performed to detect problems in the data have included checks among living donor and recipient records for invalid SSNs, as well as checks for inconsistent entry of date of birth, race, gender, and

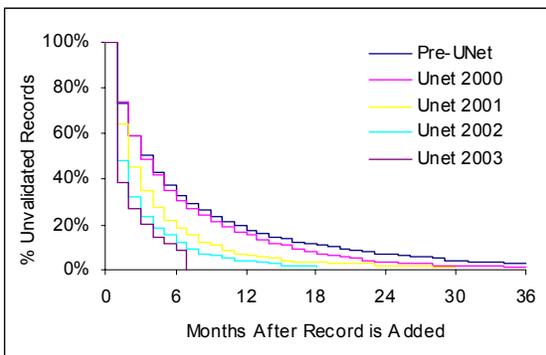
blood type across records for patients wait-listed at multiple transplant programs. Other examples include monitoring persistent waiting list registrations when programs have reported patients as having been transplanted, and ensuring consistency between waiting list registrations removed for transplant and the transplant records themselves.

Compliance with Data Submission

Data submission compliance has improved since the implementation of UNet, which provides availability of records to be completed immediately after the relevant event, as described in the section above regarding Tiedi. Further, there is no lag time caused by paper forms being mailed to and from centers, especially costly in the case of validating incorrect data, which can now be performed nearly instantly through UNet.

In June 2003, the data submission policy was modified to require that 95% of data be complete within three months of the due date, and 100% within six months. Turnaround time for all record types was also shortened. At the same time, the data amnesty plan was implemented to assist OPTN/UNOS members in coming into compliance with data submission standards. Under this plan, certain older transplant follow-up records, due between October 1, 1987, and June 30, 2001, were granted amnesty from monitoring for compliance. All other records due during this time period were required to be submitted by a deadline of June 30, 2003.

Figure II-4. Validation of Follow-Up Forms by Year



Source: SRTR Analysis, August 2003.

Figure II-4 shows the continuing improvement of the timeliness of follow-up record validation. “Survival curves” are presented, starting at the time the record is added to the system for completion, and “surviving” until validation. The highest, right-most curve indicates the time it took for centers to validate records before UNet was implemented: about 11 months until 80 percent of these records were validated. Moving down and to the left, an improvement is shown for each year

that UNet has been in place, ending in just over 4 months to complete the same percentage of records.

This improvement in timing of follow-up record validation is emblematic of the great strides have been made in OPTN data collection, quality and submission of complete validated data in recent years with the employment of technology that provides real-time data collection and quality processes. More timely follow-up data has allowed researchers to choose the most recent cohorts for analyses of follow-up data. Several other factors influence these improvements: the direct entry of data and resolution of inconsistencies by transplant personnel who have access to the source information; a greater emphasis on improving data submission compliance through policies like the data amnesty plan; and improved communications with members, through UNet and direct correspondence, about expected or overdue data. As a result of the continued vigilance of transplant organizations that provide accurate and timely data, meaningful studies can be carried out that will continue contributing to progress in the field of transplantation.

SECONDARY DATA SOURCES AVAILABLE

Additional “secondary” data sources, beyond the primary sources collected by the OPTN, provide means to help determine the accuracy and completeness of data submitted by OPTN members. These data can also expand the scope of available research. For example, additional data sources can help researchers perform the following important tasks:

Evaluate the complete ascertainment of outcomes, improving precision of analyses and answering questions about the quality of transplant data submitted by a transplant center

Examine bias in non-reporting of data, such as in whether transplant recipients who become lost to follow-up have similar outcomes to those who are not

Expand measurement of events not collected by the OPTN, such as death after a candidate is removed from the waiting list

Provide additional ascertainment of other events, such as malignancies from local cancer registries across the country

Offer measures of potentially available donors for evaluating donation practice patterns

Most secondary sources available are integrated using a patient matching system. The SRTR-ESRD person linking table (PLT), maintained by the SRTR as a

central repository for patient identifying data from various sources, provides a common patient identifier that can be used to link patient data across variant sources, including the OPTN data, CMS-ESRD data (most ESRD patients, including kidney transplant recipients and candidates, qualify for Medicare benefits), and the Social Security Death Master File (SSDMF). This PLT, developed collaboratively by URREA and the University of Michigan Kidney Epidemiology and Cost Center (UM-KECC), facilitates probabilistic matching across patient-level sources, finding similarities in patient identifiers such as SSN, HIC, names and nicknames, gender, DOB, etc. — all with allowances for common coding mistakes such as transpositions or entry of the wrong birth year. These patient identifiers also allow researchers to link primary and secondary sources without access to the confidential patient information on which they are based.

Patient Linking Between OPTN Records

One secondary data source facilitated by the PLT is the ability to establish links among different candidates and recipients within OPTN data. These links help a researcher tell, for example, that after a transplant at one center, a patient may have been re-listed or re-transplanted at another center. Analytically, this may tell us that the first transplant has failed, or in some cases, that the patient has died, even after the patient is “lost” by the original transplanting center. These links also allow researchers to analyze the frequency and outcomes from multiple wait-listings, or even to look at the incidence of living donors eventually becoming recipients themselves. Finally, they may help a researcher supplement “missing” data from one registration with that from another registration.

Social Security Death Master File (SSDMF)

The SSDMF, publicly available from the Social Security Administration (SSA), contains over 70 million records created from reports of death to the SSA. Records are reported for both beneficiaries and non-beneficiaries; 90% are reported by family members and funeral homes, the rest are reported by state and federal agencies, banking institutions, postal authorities, etc. This file includes a death date for each decedent, plus identifiers such as SSN, name, and date of birth. Because some non-beneficiaries, children in particular, may be missed, the absence of a particular person in this file does not prove the person is alive, though when combined with other sources, analyses suggest that approximately 98% of posttransplant deaths may be captured (1). These data are timely as well: Of the deaths included in the SSDMF, more than 98% are complete by the end of the third month after a death date. Monthly, these data are

matched to the PLT by the SRTR, confirming SSN matches with plausible names, birth dates, and death dates; implausible matches on these are rejected, acknowledging the possibility of erroneous SSNs in either source. Information from the SSDMF augments both the transplant and candidate analysis files; SSNs recorded for living donors have returned implausible names and dates so frequently that these are not trusted.

CMS-ESRD Database

With data drawn primarily from Medicare records for ESRD recipients, the Centers for Medicare and Medicaid Services End-Stage Renal Disease database provides another auxiliary source of data. Maintained by UM-KECC, this data source combines a range of resources about patients with kidney failure, nearly all of whom are covered by Medicare. The ESRD Medical Evidence Report, combined with detailed Medicare claims files, can indicate both a pre-transplant dialysis history as well as a return to dialysis after transplant, signaling graft failure. These sources, along with the ESRD Death Notification Report and the Standard Information Management System (SIMS) maintained by the ESRD Networks, helps validate outcomes after transplant. These data, updated at least annually, are incorporated into the SRTR candidate and transplant files as auxiliary follow-up information.

National Death Index

The **National Death Index (NDI)** misses only about 5% of all deaths in the US. Compiled by the National Center for Health Statistics (NCHS), the NDI contains data from death certificate information submitted by state vital statistics agencies. Because of the restrictive arrangements with each state agency, the prohibitive cost of using matching large samples, and significant reporting lag, these data are permitted to be used only for validation. However, they are useful in assessing the completeness of other sources regularly available for analyses.

National Cancer Institute SEER

The **Surveillance, Epidemiology, and End Results (SEER)** program of the National Cancer Institute is one of the most complete sources of information on cancer incidence and survival in the United States. For patients in some regions, SEER data from highly accurate cancer registries may validate the posttransplant malignancy data reported on OPTN follow-up records. They may also supplement information for periods before the OPTN began to collect these data.

Hospital and Donation Service Area Data

Other external data sources do not necessarily require direct linking to primary data at a patient level in order to be useful. For example, the OPTN, SRTR, and other researchers have investigated methods to make associations between OPO practice patterns and donor procurement. New data collected by the OPTN include counts of deaths that meet certain eligibility requirements, referred by each hospital to its OPO. The ability to convert potential donors into actual donors may be affected by characteristics of each hospital, such as size (number of beds) and distance from (and size of) a metropolitan area. Hospital-level data such as these may be available from the **American Hospital Association Annual Survey Database**. Relevant cost data are available from CMS in the form of the **Healthcare Cost Report Information System Dataset (HCRIS)**. The **National Center for Health Statistics (NCHS)** provides additional files that can help tabulate numbers of notifiable deaths (those that suggest a suitable donor, given cause, circumstance, and location of death), as well as demographic data about the deceased. The OPTN, SRTR, and other researchers have investigated methods to make associations between OPO practice patterns and donor procurement, considering the suitability for transplantation of deaths in hospitals served by each OPO.

CAVEATS FOR RESEARCHERS

There are a wide range of caveats and potential pitfalls important for consideration by researchers working with these primary and secondary data sources. The varying completeness and accuracy of individual data fields over time, and from source to source, are some of the more straightforward ones. The change in technology of data collection, described above, has brought about improved accuracy and completeness of data items. Changes in scope of data collection are also implemented as new fields are added, and occasionally old fields are removed. Researchers must be aware of time-dependent patterns in specific fields, many of which change at the major “turning points” outlined in Figure II-3, as well as the fast expansion of data collection fields since 1994.

The less obvious caveats involve possible biases in loss-to-follow-up, important events that may happen outside the scope of data collection, and the “annual reporting” nature of posttransplant follow-up. These concerns are discussed below.

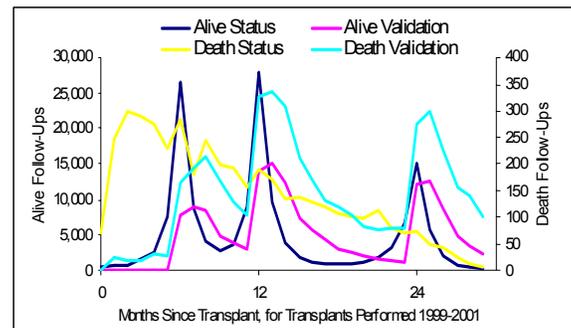
Posttransplant Cohorts: The “Annual Reporting” Nature of Follow-up

For many research questions, one of the first issues to resolve is to define the cohort for analysis. The desire to

have the most recent data on the most recent transplants must be balanced with the need for complete and unbiased data. For an analysis of posttransplant outcomes, it is advisable to allow enough lag time so that all transplants in the cohort have follow-up reported for various transplant “anniversaries” (when follow-up records are due), and to censor at these anniversaries.

To determine the most recent cohort that can be reasonably used for an analysis, one must understand the patterns that describe data submission. Posttransplant follow-up records are due at six months (except thoracic), one year, and each transplant anniversary thereafter. On each follow-up record, the center is asked to report the most recent status (alive, dead, retransplanted, or lost to follow-up) and date of this status, indicating the most recent confirmation of this status. In addition, ad-hoc follow-up records may be submitted at any time to report such events as death, graft failure, retransplant, or loss to follow-up. It is important to note that most analyses consider these events submitted on ad-hoc records as “adverse events.”

Figure II-5. Timing of Patient Status and Validation, Transplant Follow-Up Forms



Source: SRTR Analysis, August 2003.

Figure II-5 depicts the time after transplant until follow-up information is available. The sharpest spikes, occurring at 6, 12, and 24 months, are the dates of last status for patients who are alive at follow-up. Just trailing these in time (x-axis), slightly more rounded peaks indicate the time at which these follow-ups for alive patients were validated in the UNet system. This suggests that when prompted for follow-up on these anniversaries, most centers return timely information about living patients within a few months of the anniversary. However, follow-ups indicating death, also shown in Figure II-5, show a different pattern. As expected, the dates of death show a gradual but consistent decrease in frequency after transplant. However, the validation dates for these reports of death exhibit spikes at anniversaries that are similar to those of living follow-ups but less pronounced. This suggests

Table II-1: Time of Validation of Follow-Up Forms

Timing of Validation	Cumulative Percent Validated by Month:					
	Routine Follow-Up Forms			Interim Follow-Up Forms		
	2000	2001	2002	2000	2001	2002
Validated during month 1	11.3	16.2	25.7	14.0	28.7	44.5
Month 2	26.4	35.0	51.0	25.0	44.4	60.7
Month 3	41.4	54.3	67.4	36.2	59.8	72.7
Month 4	51.2	64.6	76.0	45.4	68.8	79.8
Month 5	58.7	72.1	81.1	53.4	74.9	84.0
Month 6	65.1	77.9	84.7	59.9	80.1	86.9
Month 7	69.5	81.6		64.7	83.7	
Month 8	73.2	84.6		69.2	86.2	
Month 9	76.4	87.2		73.2	88.7	
Month 10	79.0	89.2		77.1	90.2	
Month 11	81.4	91.0		80.4	91.7	
Month 12	83.3	92.3		82.9	93.0	
All not yet validated by 6 months	34.9	22.1	15.3	40.1	19.9	13.1
All not yet validated by 1 year	16.7	7.7	N/A	17.1	7.0	N/A

Source: SRTR analyses, August 2003.

that many centers wait until prompted at the anniversary to report a death that occurred during the interval, though the fact that they are more rounded indicates some ad-hoc reporting of deaths as they occur. The difference in reporting patterns described in this figure has very important implications for choice of cohorts and censor times.

Waiting for transplant anniversaries to get unbiased follow-up. Because adverse events such as death may be reported continuously and non-adverse events may not, the follow-up file over-represents death at any given time during the transplant follow-up, except just after transplant anniversaries. As an example, a researcher calculating a survival rate with data collected as of 18 months after transplant will have follow-up records reflecting the first 12 months for both living and dead patients, plus additional records possible only for patients who died in months 12 through 18. If 18-month survival is needed, it would be useful to base survival for months 12 through 18 on a cohort of people with 24 months of follow-up, since unbiased data may be expected by then. Similarly, one-month survival rates cannot be reliably calculated until at least six months after transplant (1 year for thoracic organs), after the anniversaries have prompted reporting on all patients.

If centers did consistently report all deaths on an ad-hoc basis — continuously as they occurred — a researcher could conclude that at any given point in time (allowing for a reasonable reporting lag), all deaths had been

reported either by the transplant center or a secondary source. Analysts could then forgo censoring at last follow-up and assume all patients to be alive unless records indicated otherwise. However, even though most death data are reported on the interim “death reporting” records, most of these records are filed soon after a transplant anniversary, so it is important to wait for data to be returned after transplant anniversaries. For the same reason, it is important to censor at these anniversaries, since consideration of subsequent events is subject to the same biases.

What is a sufficient time lag after an anniversary? Beyond waiting for a transplant anniversary, it is important to consider how much time elapses between the generation of the follow-up record by the system to prompt reporting and the validation of that record by the transplant center. As detailed earlier, implementation of new data collection mechanisms and more strict rules have shortened the time until validation. Table II-1 shows that the time from the date of record generation until validation is much shorter in 2002 than in previous years. Researchers should consider this cumulative percent at face value but also the diminishing benefit of waiting each additional month. For example, at month 4 after transplant in 2001, only 65% of routine records were validated; however, this represents 70% of the records that would be available by waiting until month 12. A balance needs to be struck between the need for analysis of recent data and the need for complete data. The SRTR typically allows for between 3 and 7 months

of lag time, depending on the need for using the most recent cohort available. Variables indicating a censor date, based on this lag time and reporting pattern, are included in the SRTR research files.

Different analyses, particularly those not related to posttransplant survival, have different time lag and censoring considerations. As described above, a transplant record is generated at the time a patient is removed from the waiting list for transplant. These records may be sufficient for more simple analyses that do not depend on the completion and validation of records, such as counting transplants. Using TRRs for liver, Table II-2 shows that the count of transplants stabilizes very quickly even when the validation takes longer.

Table II-2: Time of Adding Transplant Records

Time	TRR Records for Liver Transplants, 2/1/03 - 2/28/03	
	All Records	Validated Records
Beginning of next month (3/1/03)	422	391
Beginning of month + 1 (4/1/03)	435	401
5/1/03	437	402
6/1/03	437	402
7/1/03	437	402
8/1/03	437	402

Source: SRTR analyses, August 2003.

Which Recipients Are Lost to Follow-Up?

It is inevitable that after transplant, centers may have a difficult time following some recipients. These patients may move away or transfer their care to other medical professionals. Additionally, some centers have a difficult time allocating staff to report on all patients. Patients may become lost to follow-up (LTFU) either when the transplant center reports them as such or does not report follow-up forms at all. About 10% of recipients transplanted with kidneys, livers, hearts, or lungs were LTFU by the end of the third year after transplant. About two-thirds of these had been coded as LTFU by the transplant center, and the other third had no records completed for at least the last 1.5 years before the three-year anniversary.

If the patients who are lost to follow-up for any reason represent a random selection of all patients, then statistical methods of censoring can be applied for correct analysis. Censoring at LTFU may not be necessary for mortality analyses, in which extra data sources allow complete ascertainment of outcomes. However, the analyst needs to beware of the possibility of biases in LTFU and evaluate the possible effects on any analysis, especially those that rely solely on follow-up reporting of events, such as incidence of malignancy or immunosuppressant use.

Kidney recipients are much more likely to be LTFU, probably because of the viability of dialysis as an alternative therapy that removes the patient from the care of the transplant center. Kidney programs are asked to track recipients for two years after graft failure, allowing an even higher probability that the patient may become LTFU while on dialysis; however, even disregarding LTFU after graft failure, kidney recipients are far more likely to not be followed by the transplant center responsible for reporting on them. Other characteristics associated in a multivariate analysis with LTFU by three years include non-white race, younger age at transplant, and being followed by a smaller transplant program.

Differences between patients who are LTFU and patients for whom follow-up reports continue are more important if there are also differences outcomes after becoming LTFU. Using the SSDMF, a secondary source of death information that is expected to be independent of being followed by a transplant center, we ran a time-dependent survival analysis for such patients. At all points in time, patients were coded as either having been LTFU or not. Patients who had been LTFU were found to have about a 12% higher risk of death (RR=1.12, P<.01) than patients who had not been LTFU. These results persisted when the covariates above (race, age, transplant program size) were added to the model.

Extra Ascertainment of Posttransplant Results

Potential bias arising from loss to follow-up stresses the importance of using additional data sources to provide more complete ascertainment of posttransplant outcomes. We have previously shown that, although the patient's transplanting center reported only 77.3% of all deaths following transplants in the 1990s, an additional 6.9% could be found among secondary OPTN sources (linked data) and another 14.3% in the SSDMF (1). This left 0.7% and 0.8% found in CMS-ESRD and NDI data, respectively, allowing us to conclude that these first three sources provided reasonably complete ascertainment of death. Once again, the reliance on secondary sources of data was highest for kidney and

Table II-3: Waiting List and Post-Removal Deaths Reported by Extra Ascertainment: Registrations Removed from the Waiting List, 2000-2002, for Reasons Other Than Transplant, Transfer to Another Center, or Death

Outcome	Reason for Removal From Waiting List					Total
	Condition Improved	Medically Unsuitable	Condition Deteriorated	“Other”	Other Codes	
No death reported	2,326 (90.9%)	642 (70.8%)	2,123 (43.1%)	4,189 (71.7%)	989 (88.3%)	10,269 (66.9%)
Died before removal	32 (1.3%)	47 (5.2%)	390 (7.9%)	404 (6.9%)	10 (0.89)	883 (5.8%)
Died <1 month after removal	29 (1.1%)	23 (2.5%)	905 (18.4%)	153 (2.6%)	11 (1.0%)	1,121 (7.3%)
Died 1-6 months after removal	55 (2.1%)	47 (5.2%)	658 (13.4%)	343 (5.9%)	38 (3.4%)	1,141 (7.4%)
Died >6 months after removal	118 (4.6%)	148 (16.3%)	847 (17.2%)	756 (12.9%)	72 (6.4%)	1,941 (12.6%)
Total (Row percent)	2,560 (16.7%)	907 (5.9%)	4,923 (32.1%)	5,845 (38.1%)	1,120 (7.3%)	15,355 (100.0%)

Excludes deaths after the patient’s next transplant, possibly at a different center. All percentages shown (except those in the last row) are column percentages. Source: SRTR analyses, August 2003.

pancreas because of the availability of alternative therapies.

Even when using additional ascertainment, choice of cohort and censoring rules should reflect the patterns of data submission described above. When incorporating additional sources, researchers must be aware of reporting patterns within those sources — such as the possible loss of Medicare eligibility three years after transplant for CMS data — and be particularly aware of the lag of the most limiting source. The use of extra ascertainment allows mortality analyses to extend beyond LTFU, with the assumption that during the period of study, data from all sources taken together provide complete accounting of death. It is important to choose cohorts and censor dates so that all of these sources are expected to be complete. As described in the Analytic Approaches chapter, the SRTR often uses additional ascertainment for posttransplant mortality analyses, by using time at risk during which we expect the relevant data sources to be reasonably complete and unbiased. Because the lag time for extra sources such as the SSDMF is similar, if not a little shorter, than that for transplant follow-ups, the transplant anniversaries described above may be useful for censoring these analyses as well.

For graft failure data, we do not have a “complete” data source such as the NDI to conduct a similar test. Therefore, it may be problematic to adapt the approach used for patient survival, in which a researcher assumes that the patient is alive unless otherwise indicated. For some organs, retransplant is the only alternative therapy,

and examination of the transplant data file for retransplants for the same patient is sufficient for assuming complete follow-up. For kidney recipients, the alternative therapy of dialysis increases the possibility that graft failure has occurred and that the patient has returned to this other therapy, without the knowledge of the original transplanting center or any new (retransplanting) center. Some additional failure data may be available using CMS-ESRD data. Researchers should evaluate these possibilities for each individual analysis, and should consider whether either patient or graft survival is more appropriate for each organ.

Extra Ascertainment of Waiting List Outcomes

Many analyses used in establishing allocation rules are based on a comparison of outcomes with and without a transplant. Waiting list outcomes may be just as susceptible as posttransplant outcomes to underreporting of death. Transplant centers are responsible for reporting patient outcomes while on the waiting list. However, events that occur after patients are removed from the list — for example, if they are too sick to receive a transplant — are not subject to required reporting. These events may be very relevant for many analyses.

Using several secondary sources of mortality data, including secondary OPTN reporting, SSDMF, and CMS-ESRD, we examined patients who were recently removed (2000-2002) from the waiting list for reasons other than transplant, transfer to another center, or death. There were more than 15,000 registrations removed for reasons such as “condition improved or deteriorated,”

“medically unsuitable for transplant,” “refused transplant,” and “other,” accounting for about 13% of all removals during this time period. This is a substantial enough fraction that additional mortality ascertainment may be important.

Table II-3 shows that more than 7% of these patients died within one month of removal, and twice that many before month 6. Others had died, according to the secondary data sources, before they were removed from the waiting list. Without extra ascertainment, many analyses might not account for the adverse results seen for these patients. While one may expect and account for adverse results for patients removed for “condition deteriorated, too sick to transplant,” high rates for many other removal codes seem unexpected.

A historical examination suggests that over time there has been a small percentage of registrations on the waiting list for patients who have actually died, and most of these patients have an inactive status. Further, this percentage has dropped from close to 2% at the end of 1998 to around 1% in years since. More active waiting list management for liver programs since the implementation of the MELD allocation system, as well as efforts by the OPTN and SRTR to notify transplant centers of actively listed or followed patients who are indicated in the SSDMF as having died, may have contributed to this improvement in accuracy.

CONCLUSION

We have presented the reader with an introduction to the broad scope, in both topic and source, of data available to the transplant researcher. While not intended as a detailed researcher’s guide, the description here of how these data are collected, and how some caveats to these data are addressed, provides important background for both users of existing research and analysts working on new research questions with these data.

REFERENCES

1. Dickinson DM, Ellison MD, Webb, RL. Data Sources and Structure. OPTN/SRTR 2002 Annual Report: 9-26.

CONTRIBUTORS

The following individuals prepared this chapter: David M. Dickinson, MS¹; Paula C. Bryant, MBA²; M. Christian Williams, RN, CPTC²; Gregory N. Levine³; Shiqian Li¹; James C. Welch¹; Berkeley M. Keck, RN, MPH²; Randall L. Webb, BA¹. ¹Scientific Registry of Transplant Recipients / University Renal Research and Education Association; ²Organ Procurement and Transplantation Network / United Network for Organ Sharing; ³Scientific Registry of Transplant Recipients / University of Michigan.