SINGLE VENTRICLE RECONSTRUCTION TRIAL
PUBLIC USE DATASET

ABOUT THE STUDY

The NHLBI Single Ventricle Reconstruction (SVR) Trial was conducted by the Pediatric Heart Network (PHN) at 15 centers from 2005 to 2009. The PHN screened a total 920 newborns; 664 were medically eligible and 555 of these patients were randomized. The primary aim of the trial was to compare one-year transplant-free survival of newborns randomized to receive either a modified Blalock-Taussig shunt (MBTS) or a right ventricle to pulmonary artery shunt (RVPAS) as part of the Norwood procedure. Study measurements were made at baseline (pre-Norwood), at the Norwood procedure, prior to the Stage II procedure (mean age 5 months), and at age 14 months. Core laboratories were used for interpretation of two-dimensional (2D) and three-dimensional (3D) echocardiograms and for genetic analysis (ApoE gene).

The trial design has been summarized in Ohye et al. (*AHJ* 2008) and in the study protocol (available to users with approved logins). Table 1 provides key subject characteristics. Additional information on available sample sizes for measurements may be found in Ohye et al. (*NEJM* 2010; http://www.pediatricheartnetwork.org/publications/Fontan_Mainresultspaper.pdf), as well the published articles on specialized topics (see posted Bibliography at http://www.pediatricheartnetwork.org/pubSVR.asp).

Table 1. Key SVR Trial Analytic Cohort (N=549) Characteristics

Characteristic	Mean±SD or n (%)
Assigned to MBTS (*trt*)	275 (50%)
MBTS in place at end of Norwood operation (*ctrt*)	268 (49%)
Male (*GENDER*)	340 (62%)
White, non-Hispanic	346 (64%)
Birth weight, kg (*BWT*)	3.10±0.54
Gestational age<37 wk (*preterm*)	64 (12%)
Hypoplastic left heart syndrome (*HLHS*)	474 (86%)
Aortic atresia (*ATRESIA*)	343 (62%)
Norwood perfusion type (*newrcpdhca*)	
Deep hypothermic circulatory arrest (DHCA) only	296 (54%)
Regional cerebral perfusion (RCP) only*	130 (24%)
DHCA and RCP	118 (22%)
Unknown	5 (--)
Death or Cardiac Transplant by 1 yr post-rand. (*dtx1yr*)	172 (31%)
Pre-transplant death by 1 yr post-rand.	159
Transplant by 1 yr post-rand.**	13

*Patients with ≤10 minutes of DHCA were classified as RCP only

**5 of 13 died post-transplant

DATA AND DOCUMENTATION

The following datasets and descriptor files are available for download. A login and password (request access via http://www.pediatricheartnetwork.org) are required for download capability. The lock date used for creation of the public dataset is October 22, 2012. Privacy protection of these data is described in Appendix A.

1. Annotated study data collection forms (PDF) – These contain the SAS variable names next to each data field on the form. These form documents also include some created variables and their definitions.

2. SAS version 9.3 datasets

3. The file *svrformats.sas7bcat* – Include this file in your program using:
```
options fmtsearch = (fmtlib.svrformats64);
```
where fmtlib is specified using a libname statement as the path name.

4. SAS Proc Contents for each dataset (PDF)

5. Excel datasets (with variable formats applied) – These data have a .csv extension, which means that the file may also be opened either in Excel, OR in a text editor, appearing as a comma-delimited file.

STUDY RESOURCES

Resources posted on the pediatricheartnetwork.org website include:

- SVR Trial Design paper and Main Results paper (see http://www.pediatricheartnetwork.org/publications/SVR20Design20Paper.pdf and http://www.pediatricheartnetwork.org/publications/SVR_comparison1980-1992.pdf)

- SVR Trial bibliography (see http://www.pediatricheartnetwork.org/pubSVR.asp)

- SVR Trial protocol (with login access)

DATA USE POLICY

- REQUIRED ACKNOWLEDGEMENTS: All presentations and publications using these data must include the following statement: *"The NIH/NHLBI Pediatric Heart Network Single Ventricle Reconstruction Trial dataset was used in preparation of this work. Data were downloaded from https://www.pediatricheartnetwork.com/pud_login.asp?study_id=SVR on mm/dd/yyyy"."*

- PAPER, ABSTRACT, and PRESENTATION TITLES: Titles may, at the authors' discretion, mention the PHN database but should not imply that the work is from the PHN. An example of an acceptable phrase would be, "an analysis of the Pediatric Heart Network public database." Whether or not the title makes mention of the PHN, acknowledgement should be made as described above.

- All users are requested to send a copy of published abstracts and articles to the PHN Data Coordinating Center at New England Research Institutes (PHNpubs@neriscience.com) within one month of publication. This will allow the PHN and the NHLBI to document the continued impact of this study on the field.

- The login and password provided to each user are valid for 6 months. If a user decides to complete analyses leading to more than one presentation or publication in that time period, it is requested that they notify the PHN Data Coordinating Center at New England Research Institutes of their additional analysis topics, solely for the purposes of tracking.

- The login and password to access the public dataset is provided to a single user. If a colleague would like to access the public dataset for a different analysis topic, a separate request for login and password should be submitted via the www.pediatricheartnetwork.org website.

- As an approved user, you agree that you will not attempt to establish the identities of research participants through use of this dataset.

- As an approved user, you agree to not place these data in other public locations.

TIPS ON USING THESE DATA

- Identification numbers for study subjects and study sites have been re-assigned for privacy protection.

  subj_id: Subject ID ranging from 1 to 920

- The study data are contained in a large number of individual forms. These may be used jointly by merging on *blind_id*. No dataset has more than one record per subject. A single dataset called KEYINFO, containing 24 of the most commonly used raw and created variables, including survival time variables, KEYINFO, is also provided. It contains 549 records.

- One dataset, R100, contains data on both screened (N= 920) and randomized subjects (555 of the 920). The 555 randomized patients can be identified using *CONSENT*=2 (Yes). The SVR publications, however, utilize a total of 549 subjects (*main549cohort*=1ma (Yes)). There were 5 subjects who never underwent the Norwood procedure, and 1 subject who withdrew from the trial in week 1. These 6 subjects (*blind_id* 401,455,515,683,691,731) are excluded from PHN analyses. There are also 5 subjects who died in the operating room during the Norwood procedure (*diedinOR*=1 in KEYINFO dataset).

- The treatment assignment variable is located in the KEYINFO dataset and is called *trt*. The non-intent-to-treat variable is also located in the KEYINFO dataset and is called *ctrt*. The variable *ctrt* classifies patients according to the shunt in place when the patient left the operating room at the end of the Norwood procedure.

- The raw data collected are contained in the original variables (denoted by upper case variable names). Prior to analysis, these variables must have special values (typically negative numbers, see Appendix B) set to missing. Created variables (denoted by lower case variable names) already contain a SAS missing value if the measurement is unavailable.

- To select for 2D or 3D echocardiograms (Forms R301 and R302, respectively) that have data acceptable for analysis, use ACCEPTABLE=1. Unacceptable echocardiograms have no qualitative or quantitative measurements recorded. Methodology for 2D and 3D image acquisition, central measurement, and indexed variables (i.e., z-scores where appropriate or indexing to certain transformations of BSA), respectively is found in Frommelt et. al, *Circulation* 2012; 125:2630-2638 and Marx et al., In press (Nov 2013), *Circulation CV Imaging*.

- Cause of death (COD) was assigned by a five-member panel that arrived at consensus decision on each event after independent review by each member (see Ohye et al., *JTCVS* 2012**;** 144:907-14). This adjudicated COD (character variable, named CODE) is located in the R801 dataset and correspondingly, site-level determinations of cause of death (on the R108 and R188 forms) are omitted. For privacy reasons, the COD narratives from the sites are also omitted. A code list for the causes of death is provided in the annotated data form and as a separate code list for download. A total of 184 deaths, which occurred through the end of the main SVR trial (2009), were adjudicated. This includes 159 pre-transplant deaths by 1 yr; 5 post-transplant deaths by 1 yr; 19 pre-transplant deaths after 1 year; and 1 post-transplant death after 1 year. The post-transplant deaths occurred in *blind_id* = 25, 110, 606, 657, 817, and (after one year) 555.

- Survival times and related measures are calculated from dates on the R108 and R188 forms. The calculated variables are located in the KEYINFO dataset. Most of these are calculated with respect to the trial primary endpoint (death or cardiac transplant at 1 year post-randomization; *dtx1yr*). In addition, the status of each subject at the end of the trial (one year after the last randomized subject), was recorded. This longer-term status is summarized by eventyn (1=yes; 2=no) and *time_event* (in days).

- It is often of interest to construct totals of SAEs and complications (the term used for all adverse events not included in the SVR SAE definition). Such totals must be done utilizing a large number of data forms, sometimes including R800, R105, and R206, in addition to the more obvious staged surgery forms (R103, R104, etc.). For convenience, a few pre-calculated event totals are included in the KEYINFO dataset.

- Anatomic diagnosis is best described using the detailed code contained in the variable *adx*, and the corresponding Code List O. The *HLHS* variable is created based on *adx* in dataset R100.

- Twenty-two subjects were not discharged from the hospital after the Norwood procedure and prior to the Stage II procedure.

- Socioeconomic status (SES) is summarized for subjects living in the U.S. using a calculated variable *sescore* in the KEYINFO dataset. SES was assigned using a U.S.-census-based score derived from 6 measures related to income, housing, and occupational-related features of the subject's census block tract at the time of randomization (see Diez Roux AV, Merkin SS, Arnett D, Chambless L, Massing M, Nieto FJ, et al. Neighborhood of residence and incidence of coronary heart disease. N Engl J Med 2001 Jul 12;345(2):99-106) and the PHN paper Tweddell et al. *JTCVS* 2012. In addition, a separate U.S. census-based measure of the percentage of persons living below the federal poverty level in the subject's census block tract, called *poverty*, is in KEYINFO.

- All anthropometric z-score calculated variables (weight-for-age, length-for-age, weight-for-length, etc.) in the SVR trial datasets are based on the WHO standard (see de Onis M, Garza C, Onyango AW, Borghi E. Comparison of the WHO child growth standards and the CDC 2000 growth charts. J Nutr 2007; 137(1):144-148).

- The second edition of the Bayley Scales of Infant Development® were used in the SVR Trial (BSID–II). See Bayley N. Bayley Scales of Infant Development, Second Edition. Second Edition ed. San Antonio, TX: The Psychological Corporation; 1993. The summary scores *PDISCORE* and *MDISCORE* are located in the R114 dataset.

- The SVR Trial code lists are posted along with the datasets. Some variables are formatted (such as Complications), but with short versions of the names, and for some, the code numbers are retained and the code list must be referenced (e.g., medications, cardiac cath interventional procedures).

ADDITIONAL ASSISTANCE

If you have questions about the study dataset that this documentation and the above resources (protocol, articles) have not answered, please contact Victor Zak (vzak@neriscience.com) at the PHN Data Coordinating Center (617-923-7747).

APPENDIX A

Implementation of Privacy Protection Rules for Public Use of the PHN SVR Trial Dataset

Variables that could lead to subject identification were eliminated in the public dataset. Steps included:

1. Removal of original study ID number (replaced with *blind_id*, a random consecutive numbering ranging from 1 to 920), and removal of acrostic. Of note, no names, addresses, zip code, or medical record numbers were ever contained in the original study dataset.

2. All dates in the original datasets were removed, and replaced with "Age at event/intervention/procedure" in years (to 2 decimal places).

3. Free (write-in) text variables remain in the public dataset, with the exception of narratives for cause of death and for adverse events. These often provide highly relevant information for interpretation of the data. However, any write-in string that referred to a specific date, a particular medical center or a particular MD was blinded or omitted.

4. Outliers for continuous variables and small group sizes for categorical variables were retained in the dataset for public use due to their importance in interpretation of the data and low likelihood of unblinding any user to a subject identity unless the user already had access to the particular medical center's data for valid reasons.

APPENDIX B

Special Value Codes

-9 = missing

-8 = don’t know/indeterminate

-7 = refused to answer

-6 = not recorded

-5 = measurement could not be reliably recorded or is not interpretable (study technically inadequate)

-4 = illegible

-2 = programmed skipped field based on results of or response to a previous question

-1 = not applicable/structure not present