

African American Study of Kidney Disease and Hypertension (AASK) clinical trial

Database Availability

Updated: November 28, 2006

1. Introduction

The African American Study of Kidney Disease and Hypertension Study (AASK) compared the effectiveness of various antihypertensive regimens to slow or prevent progressive renal dysfunction in 1094 African-American with a clinical diagnosis of hypertensive renal disease. AASK was designed to determine what level of blood pressure control prevents a loss of renal function, and whether antihypertensive agents might slow the loss of renal function independent of their effect on blood pressure.

The purpose of this document is to explain details about the available options for users to obtain the data sets included in the AASK database. Data can be obtained as an ORACLE dump file, as Comma Separated Values files, or as SAS files.

2. Database Details

RTI International currently hosts an ORACLE 9i Database with the AASK clinical trial data. Details about this database follow.

Database Size

RTI's AASK ORACLE database size is approximately: 1,400 MB.

Tablespace

This database requires a table space with the following properties:

Table space name: USERS
pctfree: 10
initrans 1
maxtrans 255
storage:
 initial 2080K
 minextents 1
 maxextents unlimited

Tables

The AASK database contains 60 tables. Table names and descriptions are shown on Exhibit 1. Appendix A shows details for all field on each one of the 60 tables on this database.

Exhibit 1: AASK Tables

	Table Name	Description
1	ADHERE_REVIEW2_ARCHIVE	
2	ADHERE_REVIEW_ARCHIVE	
3	ANNUAL_PHYS_ARCHIVE	
4	BP_ARCHIVE	
5	BP_NON_RZ_ARCHIVE	
6	CARDIO_REVW_ARCHIVE	
7	CARDIO_REVW_ARCHIVE_200249	
8	CBL_MAILING_SERUM_ARCHIVE	
9	CBL_MAILING_URINE_ARCHIVE	
10	CBL_RESULT_SERUM_ARCHIVE	
11	CBL_RESULT_URINE_ARCHIVE	
12	CC_DEATH_ARCHIVE	
13	CC_DEATH_ARCHIVE_200249	
14	CC_STOP_POINT_ARCHIVE	
15	CERTIFICATION_ARCHIVE	
16	CLOSE_OUT_ARCHIVE	
17	CMS_DEATH_ARCHIVE	
18	CMS_STOP_POINT_ARCHIVE	
19	CMS_STOP_POINT_ARCHIVE_200249	
20	DEMOGRAPHIC_ARCHIVE	
21	DIAGNOSIS_DETAIL_ARCHIVE	
22	EKG_ARCHIVE	
23	ELIGIBILITY_ARCHIVE	
24	GENETIC_INFO_ARCHIVE	
25	GFR_MAILING_ARCHIVE	
26	GFR_RESULT_ARCHIVE	
27	HOSPITAL_ADMISSION_ARCHIVE	
28	HOSPITAL_ARCHIVE	
29	LOCAL_LAB_ARCHIVE	
30	MAP_CC_REVW_ARCHIVE	
31	MAP_CMT_REVW_ARCHIVE	
32	MED_ARCHIVE	
33	MED_COMPLIANCE_ARCHIVE	
34	MED_COMPLIANCE_ARCHIVE_200249	
35	MISSED_VISIT_RSN_ARCHIVE	
36	MISSED_VISIT_RSN_ARCH_200249	
37	OPEN_LABEL_MEDS_ARCHIVE	
38	PATIENT_ARCHIVE	
39	PHYSICAL_EXAM_ARCHIVE	
40	POST_AASK_ARCHIVE	
41	QC_GFR_MAILING_ARCHIVE	
42	QC_GFR_RESULT_ARCHIVE	

43	QC_ID_ARCHIVE
44	QC_ID_CBL_SERUM_ARCHIVE
45	QC_ID_CBL_URINE_ARCHIVE
46	QC_ID_GFR_ARCHIVE
47	QC_MAILING_SERUM_ARCHIVE
48	QC_MAILING_URINE_ARCHIVE
49	QC_RESULT_SERUM_ARCHIVE
50	QC_RESULT_URINE_ARCHIVE
51	RANDOMIZATION_ARCHIVE
52	RENAL_FAILURE_FU_ARCHIVE
53	RENAL_FAILURE_NOTIFY_ARCHIVE
54	RE_INTRO_RAND_ARCHIVE
55	SCRN_VISIT_ARCHIVE
56	SF_36_ARCHIVE
57	SYMPTOM_ARCHIVE
58	TRANSFER_ARCHIVE
59	VISIT_ARCHIVE
60	VISIT_ARCHIVE_200249

3. Data export options

3.1 ORACLE “dump” file

Oracle provides the Export/Import utility combination for data transfer. An Export “dump” file is in binary format; it is not editable, but can be used with a text editor to view portions of this file. The dump file contains data, as well as all the DDL commands to re-create the data dictionary. Editing the export dump file will lose some binary information, making the file unusable.

For the most part, export dump files are not specific to a version of Oracle. Dump files also are not Operating System specific; as a result, the Export/Import utilities can be used to transfer data between databases on different platforms.

The Export/Import utilities can be used to move information as follows:

- Between databases on the same version of Oracle
- Between databases on different versions of Oracle
- Between databases on different operating system platforms
- From one schema to another schema
- From one tablespace to another tablespace
- In the entire database

The “dump” file for the AASK database is 304,992,256 bytes in size.

Users can obtain this dump file from RTI International and they can import data into their ORACLE database, by following these steps:

- Create ORACLE database
- Create tablespace “USERS”
- Create schemas users
- Grant any of the users these privileges:
 - Select any Dictionary
 - Unlimited tablespace
- Use import/export utilities to import “dump” file

3.2 Comma Separated Values Files

The comma-separated values (or CSV) file format is a delimited data format that has fields separated by the comma character and records separated by newlines. Fields that contain a comma, newline, or double quote character, or which start or end with whitespace that is to be preserved, must be enclosed in double quotes. However, if a line contains a single entry which is the empty string, it may be enclosed in double quotes. If a field's value contains a double quote character it is escaped by placing another double quote character next to it. The CSV file format does not require a specific character encoding, byte order, or line terminator format.

CSV is one implementation of the Delimiter-separated values file format, which can use any character as the delimiter. Many applications in fact allow .csv-named files to use any delimiter character. However CSV differs from other delimiter separated file formats in using a " (double quote) character around fields that contain reserved characters (such as commas or newlines). Most other delimiter formats either use an escape character such as a backslash, or have no support for reserved characters.

RTI has created CSV files for all ORACLE tables in the AASK database. There is one CSV file for each ORACLE table. Some CSV file sizes are as small as 700 bytes and others as large as 20,000,000 bytes, depending on the records contained on each table.

3.3 SAS Dataset files

The SAS7BDAT file format is a proprietary SAS file format. It is the file extension for the SAS system data set file.

SAS 9 File Format

In order to provide longer file names and variable names, the file format used in SAS 7 and 8 is different from the file format used in SAS 6. For SAS 9, the file format is basically the same as in SAS 7 and 8. The Base SAS engine is the same engine, except that for SAS 9, longer format names can be defined and used. Longer format names in SAS 7 or 8 can not be used.

SAS files created with SAS 7 and 8 are compatible with SAS 9. However, a SAS file that is created with a 32-bit version of SAS has a different data representation than a SAS file from a 64-bit version of SAS. The data representation is the format in which data is represented on a computer architecture or in an operating environment. Therefore, if you have SAS files that were created on a 32-bit version of SAS and you have upgraded to a 64-bit version of SAS, you will have processing limitations due to the different data representation.

SAS 9 Filename Extensions

Because SAS needs to distinguish among the different file types and versions, SAS automatically assigns a specific extension to each file when the file is created. For example, in order to distinguish SAS 7 and 8 files from SAS 6 files, the extensions are different. For SAS 9, the file extensions are the same as the file extensions in SAS 7 and 8. The following table lists the file extensions for a SAS data file (SAS data set with member type DATA) in SAS 6, 7, 8, and 9 for different operating environments:

File Extensions for a SAS Data File in Different Operating Environments				
Engine Name	UNIX	OpenVMS Alpha	Windows	z/OS (1)
V6	.ssd01	.SASEB\$DATA	.sd2	not available
V7	.sas7bdat	.sas7bdat	.sas7bdat	.sas7bdat
V8	.sas7bdat	.sas7bdat	.sas7bdat	.sas7bdat
V9	.sas7bdat	.sas7bdat	.sas7bdat	.sas7bdat

(1): applies to SAS data sets that reside in the hierarchical file system of UNIX System Services. Source:

<http://post.queensu.ca:8080/SASDoc/getDoc/en/lrcon.hlp/a001331051.htm>

RTI has created SAS files for all ORACLE tables in the AASK database. There is one SAS file for each ORACLE table.