



# **2011 Population File for Use with AHRQ Quality Indicators Version 4.3**

**Prepared for:**

Agency for Healthcare Research and Quality  
U.S. Department of Health and Human Services  
540 Gaither Road  
Rockville, MD 20850  
<http://www.qualityindicators.ahrq.gov>

**Contract No. 290-04-0020 (AHRQ SQI-II)**

**Prepared by:**

**Battelle**  
505 King Avenue  
Columbus, OH 43201

**August 2011**

Version 4.3a is a maintenance release of Version 4.3. The differences between the two versions are:

- Version 4.3a includes Version 29 of the Limited License edition of the 3M<sup>TM</sup> APR-DRG Grouper. This grouper corrects an issue found in Version 28 of the Limited License edition of the 3M<sup>TM</sup> APR-DRG Grouper. The grouper is only used with the Inpatient Quality Indicator (IQI) mortality measures. Version 4.3 includes Version 28 of the Limited License edition of the 3M<sup>TM</sup> APR-DRG Grouper, which was incorrectly assigning a Risk of Mortality (ROM) subclass for cases dated on or after 10/1/10.
- Version 4.3a allows users to calculate area-level indicators for years 2010 and 2011, correcting an issue previously identified in Version 4.3.

All other aspects of the software, including measure specifications, remain the same. Thus this document (related to Version 4.3) remains unchanged.

## **Contents**

Overview.....	1
Sources.....	1
Data File Specification.....	2
2010 Census Counts.....	3

## **List of Tables**

Table 1. Data Fields in POP95T11.txt.....	3
-------------------------------------------	---

## Overview

The AHRQ Quality Indicators include area-based measures, which gauge ambulatory care-sensitive conditions for which the population at risk is defined according to the hospital location or the patient's residence. The denominators for these indicators must be constructed from an outside source rather than drawn from a subset of discharges in the user's input file. This document describes the file of population data that is constructed from public-use Census data and provided for use with the Quality Indicators syntax. Population figures through 2011 for use with AHRQ Quality Indicator syntax are provided in the file POP95T11.TXT. These figures were derived in a very similar manner to those previously provided through 2010, however, where the 2010 file contained estimates for the years 2000 through 2008 and modified projections for 2009 and 2010, the 2011 file contains estimates for 2000 through 2009 and modified projections for 2010 and 2011. Additionally, because the most current Census Bureau estimates for the full estimation period were used, all population counts from 2000 through 2009 are subject to change. The 2011 file uses the same intercensal estimates for the years 1995 through 1999 as the 2010 file, so counts for these years did not change.

All counts represent total resident population as of July 1.

## Sources

To quote the Census Bureau, "Estimates are for the past, while projections are based on assumptions about future demographic trends. Estimates generally use existing data collected from various sources, while projections must assume what demographic trends will be in the future." Estimates are generated with the assistance of the Federal State Cooperative Program for Population Estimates (FSCPE) using residence, total births, total deaths, and net migration. With each new issue of July 1 estimates, revisions are made to all years back to the last census. The estimates were obtained at <http://www.census.gov/popest/datasets.html>, under 'County population data sets'. County level detail was taken from 'County estimates by demographic characteristics – age, sex, race and Hispanic Origin'. Files are obtained from "State datasets" under the bullet 'Age, sex, race and Hispanic Origin', however, additional files under the bullet 'Selected Age Groups and Sex' were used to construct 18-24 year age categories from the bureau's 15-19 and 20-24 year age groups. Categories of race and Hispanic Origin were constructed from the Census Bureau variables for 'white-only', 'black-only', etc., and counts from the Bureau's 'two or more' category variables were added to 'other'. Pre-2000 intercensal periods used a specific 'other' race category and did not allow multiple race responses. This constitutes a small discontinuity between pre-2000 and post-2000 years that cannot be avoided due to changes adopted by the Census Bureau.

Projections were generated from data available at <http://www.census.gov/population/www/projections/popproj.html>, under "Interim Projections 2000-2030 based on Census 2000 (released 2005)". Projections for 2010 and 2011 were taken from "Interim State Projections of Population for Five-Year Age Groups and Selected Age Groups by Sex: July 1, 2004 to 2030" (under "Detailed Data Files"),

providing state level detail for 18 to 24 year olds as well as five-year groups. The Bureau states that differences between projected and estimated populations for overlapping years can be significant. Therefore, the projected growth from 2009 to 2010 and from 2009 to 2011 for each state, gender and age group was applied to year 2009 population estimates to avoid discontinuities between projections and estimates at the 2009 versus 2010 division. This means that county-level and race-specific projections for 2010 and 2011 are essentially pro-rata projections from 2009 county-, age-, gender- and race-specific estimates using state-wide age- and gender-specific projections. This same procedure was used for POP95T10.txt, however in that file year 2009 and 2010 projections were applied to year 2008 estimates.

Estimates for 1995 through 1999 intercensal years are from “Intercensal estimates by demographic characteristics (1990-1999)” under ‘County population datasets’ at <http://www.census.gov/popest/datasets.html>. These are unchanged from previous years.

## **Data File Specification**

POP95t11.txt is an ASCII-based text file containing 679,752 records with a fixed logical record length of 150 bytes. The file is structured for use with AHRQ Quality Indicator programs PQSASA2.SAS, PQSASA3.SAS, PSSASA2.SAS, IQSASA2.SAS, IQSASA3.SAS, PDSASA2.SAS and PDSASA3.SAS. As such, any modification to this file will affect the operation of these programs.

A given county is identified by the Federal information processing standards code (FIPS code) for the State in which it is located and by the county’s FIPS code. For each county within the U.S., the file contains 216 records: a record for each unique combination of gender, eighteen age groups, and six race groups. Each physical record represents a gender, age group and race group combination for that county and contains population estimates (rounded to integer values) for that combination for each year from 1995 through 2011.

The file has data for 3,147 counties or “equivalent areas”, defined to constitute primary divisions of their states. “Equivalent areas” include the “independent cities of Baltimore, Maryland, St. Louis, Missouri, Carson City, Nevada, and 39 independent cities in Virginia”. Because they are independent of any contiguous county, they are treated as separate counties with their own population records. Population figures for surrounding counties exclude them. Differences in the record count from previous population files is due to changes in county definitions or such independent cities.

Definitions for State and county FIPS codes can be found at

<http://www.census.gov/datamap/fipslist/AllSt.txt>.

POP95T11.txt is in fixed column format. Table 1 presents the file’s specific fields and the code schema used for each field.

Table 1. Data Fields in POP95T11.txt

Field	Variable	Column Position	Format	Codes
1	State	1-2	Zero Filled Num.	FIPS Code
2	County	3-5	Zero Filled Num.	FIPS Code
3	Sex	7	Numeric	1=Male, 2=Female
4	Age Group	9-10	Numeric	1=0-4 years 2=5-9 years 3=10-14 years 4=15-17 years 5=18-24 years 6=25-29 years 7=30-34 years 8=35-39 years 9=40-44 years 10=45-49 years 11=50-54 years 12=55-59 years 13=60-64 years 14=65-69 years 15=70-74 years 16=75-79 years 17=80-84 years 18=85+ years
5	Race	12	Numeric	1=White, 2=Black, 3=Hispanic, 4=Asian & PI, 5=Amer. Indian, 6=Other
6	1995 Population	13-19	Numeric	Integer Totals
7	1996 Population	20-26	Numeric	
8	1997 Population	27-33	Numeric	
9	1998 Population	34-40	Numeric	
10	1999 Population	41-47	Numeric	
11	2000 Population	48-54	Numeric	
12	2001 Population	55-61	Numeric	
13	2002 Population	62-68	Numeric	
14	2003 Population	69-75	Numeric	
15	2004 Population	76-82	Numeric	
16	2005 Population	83-89	Numeric	
17	2006 Population	90-96	Numeric	
18	2007 Population	97-103	Numeric	
19	2008 Population	104-110	Numeric	
20	2009 Population	111-117	Numeric	
21	2010 Population	118-124	Numeric	
22	2011 Population	125-129	Numeric	

## 2010 Census Counts

As described above, this population file makes most use of detailed county-level intercensal estimates, and ‘fills in’ remaining years from projections. To avoid discontinuities between estimates and projections, the percentage growth by state, gender, and age-group from 2009 to 2010 and from 2009 to 2011 calculated from the projection

data were applied to 2009 estimate data to derive final figures for 2010 and 2011. Like the estimates themselves, these figures were derived independently of the actual 2010 Decennial Census counts. Summing these records for years 2001 through 2009 match estimate totals for these years within a very small rounding error. However, summing all records for 2011 derives a total of 309,642,604, which is 0.3% higher than the published enumerated total of 308,745,538.