



ANALYTIC METHODOLOGY FOR ASSESSMENT OF COLORECTAL CANCER SCREENINGS AND DISEASE PREVALENCE IN ARKANSAS

ACHI

September 30, 2021

Table of Contents

Introduction	3
Methods	3
Data Sources	3
Study Population	4
Administrative Claims-based Screening Analyses	5
Estimating Additional Screenings	6
Cancer Registry Data Analyses	7
Limitations	8
Appendices	10
Appendix A: SAS® Code for Claims-based Analyses	10
Appendix B: List of Procedure Codes for Screenings Analyses	14
Appendix C: SAS® Code for Cancer Registry Analyses	15
Appendix D: List of Recodes used for Cancer Registry Analyses	18

For general inquiries, please contact:

Mike Motley, MPH, Director of Analytics
Arkansas Center for Health Improvement
1401 West Capitol Avenue, Suite 300 (Victory Building)
Little Rock, AR 72201
mwmotley@achi.net
501-526-2244

Suggested Citation

ACHI. (2021). *Analytic Methodology for Assessment of Colorectal Cancer Screenings and Disease Prevalence in Arkansas*. Arkansas Center for Health Improvement. Little Rock, AR

Introduction

At the request of Fight Colorectal Cancer, ACHI conducted an assessment to identify gaps in access to colorectal screenings, including those related to follow-up colonoscopies and cost-sharing. Fight Colorectal Cancer is a patient-empowerment and advocacy organization with a mission to cure colorectal cancer (CRC) through patient support, policy change, and research. The assessment examined current prevalence of and access to CRC screenings and CRC disease incidence. ACHI also projected the number of Arkansans who could be screened based on the final 2021 United States Preventive Services Task Force (USPSTF) recommendation to lower the age to start CRC screenings to 45. Where available, ACHI generated demographic and geographic profiles to identify differences in marginalized or underserved populations.

This document provides guidance on the methodology used by ACHI to complete the colorectal cancer assessment. Appendices are included that contain specific examples of SAS code used to complete this analysis. More specifically, this document provides guidance on the following:

- Data sources used for the analyses
- Study population inclusion and exclusion criteria
- Methods for Arkansas Cancer Registry data analyses of colorectal cancer
- Methods for administrative claims-based analyses of screenings
- Procedure codes used and other designations of utilization
- Approach to estimates of potential future screenings
- Limitations of the data sources and the of the analysis overall
- Examples of SAS® code used to create the analytic programs used in this assessment

DATA SOURCES

Data for the assessment are from the Arkansas Healthcare Transparency Initiative's All-Payer Claims Database (APCD),ⁱ housed at ACHI. The APCD contains data for the majority of healthcare covered lives in Arkansas including medical, pharmacy, and dental claims, enrollment and provider files, as well as death and birth certificates and cancer registry data. The APCD version used for this analysis includes data from 2013 through December 2018 for Medicare and 2013 through June 2020 for Arkansas Medicaid, fully insured private payers, and self-insured payers receiving state funds. Cancer registry data is from January 2013 through December 2017. Data are submitted to the APCD on a quarterly basis pursuant to a mandate authorized by the Arkansas Healthcare Transparency Initiative Act of 2015. The data include some geographic and demographic information and a unique identifier permitting tracking of an individual over time and across datasets, but they do not include direct personal identifiers such as name or street address.

ⁱ Arkansas APCD Website <https://www.arkansasapcd.net/Home/>

STUDY POPULATION

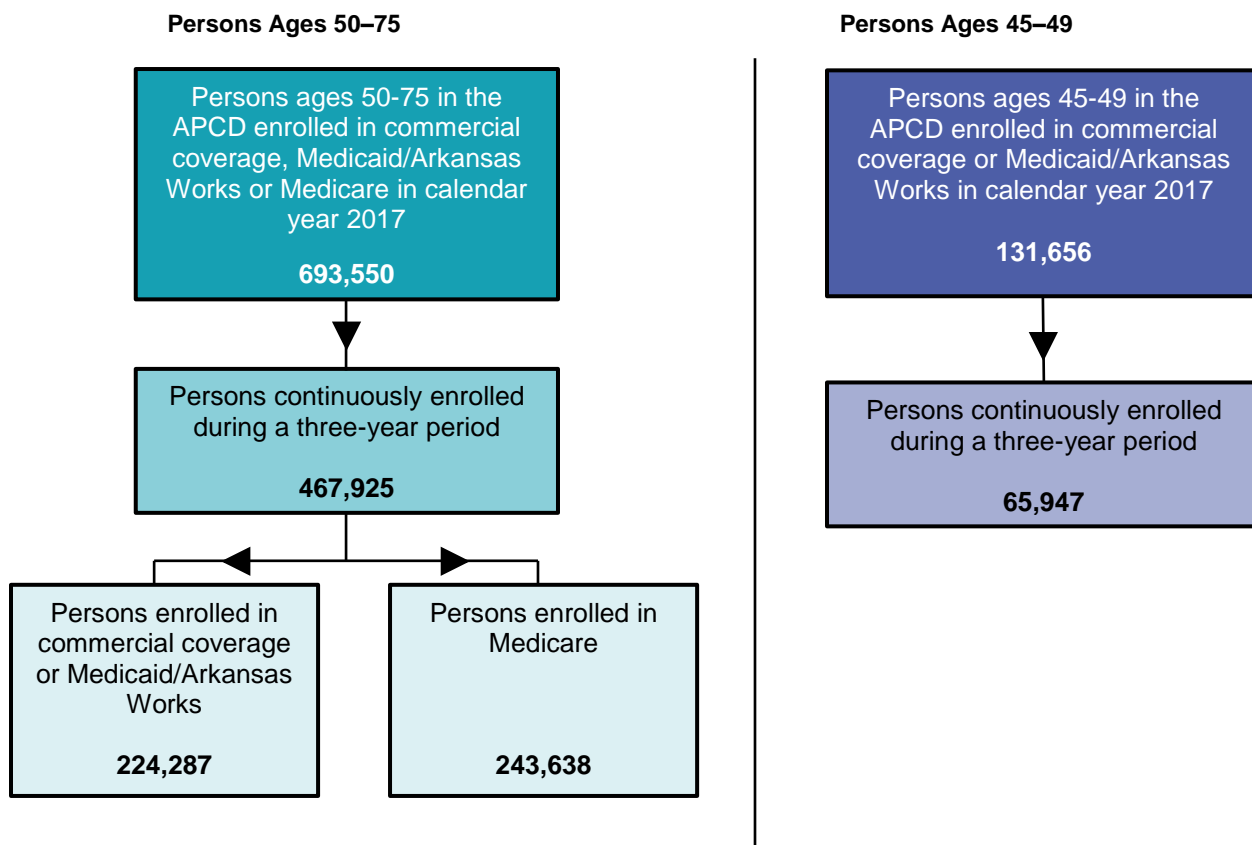
The overall study population for the CRC screening portion of this analysis included persons enrolled in commercial coverage, Medicaid (including Medicaid expansion “Arkansas Works”), and Medicare. Figure 1 shows the flow of the individuals from the APCD that were included in the study population for year 2017. The study population flow is only displayed for the 2017 cohort because this is the first year of overlap for the selected overall study periods for Commercial and Medicaid/Arkansas Works enrollees and the Medicare beneficiary population.

To be included in our overall study population, individuals had to be continuously enrolled for the entire study period; therefore, the same individuals present in the 2017 cohort are included in subsequent study years. Because this analysis includes separate age groups and multiple study years, some individuals in the 45- to 49-year-old age group moved into the 50- to 75-year-old age group into subsequent study years.

The inclusion criteria among the commercial and Medicaid/Arkansas Works enrollees included individuals age 45 as of January 1, 2017, to age 75 as of December 31, 2019, who were continuously enrolled from 2017 to 2019. Continuously enrolled is defined as having no gap in coverage or having one coverage gap of 45 days or less.

For Medicare enrollees, the inclusion criteria included individuals who reached age 65 as of January 1, 2016, to age 75 as of December 31, 2018, and who were continuously enrolled from 2016 to 2018.

FIGURE 1: FLOW CHART OF THE STUDY POPULATION, 2017



ADMINISTRATIVE CLAIMS-BASED ANALYSES OF CRC SCREENINGS

This assessment includes analyses of frequency of CRC screenings among individuals ages 50-75 enrolled in commercial coverage, Medicaid/Arkansas Works, or Medicare, including utilization of and out-of-pocket costs for colorectal cancer stool-based screenings and follow-up colonoscopies by age and payer type. These analyses also includes a projection of future utilization for persons ages 45–49 based on new USPSTF recommendations

For this assessment we analyzed commercial, traditional Medicaid (including Medicaid expansion “Arkansas Works”), and Medicare medical claims of Arkansas residents ages 50–75 and 45–49 who met the guidelines for a colorectal cancer screening based on the 2020 USPSTF recommendation.¹ Arkansans included in the study population for the study duration of January 2017 to December 2018 were assessed for evidence of any history of stool-based tests and follow-up colonoscopies. For the purposes of the assessment, a follow-up colonoscopy was defined as a colonoscopy completed within six months of a stool-based test. Stool-based tests included guaiac fecal occult blood tests (gFOBT), fecal immunochemical tests (FIT), and the stool DNA test Cologuard.

For individuals ages 50–75 who were screened and had follow-up colonoscopies, screening rates and counts for stool-based tests and follow-up colonoscopies were computed by coverage type, time to follow-up colonoscopy (one, three, or six months), race (Medicare only), and urban or rural geographic area. As current Medicare data are available through December 2018, follow-up colonoscopies within six months of an abnormal stool-based tests that occurred after June 30, 2018, may not be reflected in the data reported to the APCD. Consequently, analyses regarding follow-up colonoscopies within six months in 2018 that include Medicare data are preliminary and only include stool-based test data for January 1–June 30, 2018.ⁱⁱ Data for stool-based tests and follow-up colonoscopies within six months for full year 2018 are available for Medicaid/Arkansas Works and commercial coverage analyses.

Claims payment and evidence of any cost-sharing for a stool-based test were assessed for variation by payer type and test type. Screening rates and counts for individuals ages 45–49 were also computed to project the stool-based test and follow-up colonoscopy rates and counts for persons in that age group.

These analyses use Medicare fee-for-service, commercial, and Medicaid enrollment and claims information from 2013 through 2018, and references the appropriate CRC colonoscopy and stool-based test procedure codes (for FIT, gFOBT, and Cologuard). Claims and enrollment records were extracted for individuals who met the age-specific and continuous enrollment inclusion criteria for each payer.

ⁱⁱ For the time to follow-up colonoscopy analysis, dates were extended for follow-up colonoscopies that occurred within one month or three months from the stool-based test in partial year 2018. These dates included: January 1, 2020, through November 30, 2020, for the one-month analysis and January 1, 2020, through September 30, 2020 for the three-month analysis.

To establish the eligible study population, records were extracted by payer group according to the continuous enrollment criteria included above. Eligible Medicare enrollees who were also enrolled in Medicaid or commercial populations at some previous time during the study period are included in the Medicare group, and excluded appropriately from the Medicaid and commercial groups. For these individuals, any relevant claims history (e.g., stool-based CRC screening history) that occurred during the study period was included, even if it occurred on commercial or Medicaid claims for that individual. Age categories were created for the following groups: 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, and 80 and older. These age groups were selected in part due to alignment with previous and new USPSTF guidelines for CRC screening age recommendations.

Analysis of stool-based tests and follow-up colonoscopies was done using appropriate procedure codes for each test type. These codes are included in the appendices of this report.

Counts of stool-based screenings were stratified by age group and divided by the overall study population for each age group to calculate utilization rates. Variables designating amounts paid by the payer and for patient out of pocket costs (the sum of any co-pay, coinsurance, and deductibles) were flagged to calculate the proportion of members whose stool-based screenings were paid for by the payer, and/or for which any out of pocket costs were paid. Percentages were calculated for each of these occurrences based on the overall number of individuals receiving a stool-based screening by age group. For iterative review and validation we calculated the mean, median, minimum, maximum, and inner-quartile range for all payer-paid amounts and out-of-pocket costs.

ESTIMATING THE NUMBER OF ADDITIONAL PATIENTS WHO WOULD UTILIZE A STOOL-BASED SCREENING BASED ON THE UPDATED USPSTF CRC RECOMMENDATION

To produce a simple estimate of the number of additional Arkansans who would be screened based on the 2021 final USPSTF recommendation of CRC screenings for 45–49 year olds, we used the 2017 rate of stool-based screening among 50–54 year olds as a proxy utilization rate. The 50–54 age group was chosen for the proxy utilization rate (as opposed to the entire eligible age range) because it is assumed that their behavior in seeking evidence-based screenings would be more comparable to that of 45–49 year olds than that of the entire age-eligible population (50–75). The utilization rate among 50–54 year olds was multiplied by the state population estimate for 45–49 year-olds to derive an estimated number of individuals of that age who would be screened in alignment with the updated USPSTF recommendation. The rate of follow-up colonoscopy utilization among 50–54 year olds who had a stool-based screening also was used to estimate the number of individuals aged 45–49 who would have a follow-up colonoscopy. For this estimate, we multiplied the proxy follow-up colonoscopy rate by the estimated number of individuals ages 45–49 who would have a stool-based CRC screening.

ANALYSIS OF CANCER REGISTRY DATA FOR CRC INCIDENCE AND STAGE AT DIAGNOSIS

To assess CRC incidence and the number of late-stage diagnoses by age group, ACHI analyzed cancer data from the Arkansas Cancer Registry. Cancer registry data availability may vary by state. The Arkansas APCD includes Arkansas Cancer Registry data as part of the state's broader Healthcare Transparency Initiative. Diagnoses of CRC were stratified by age group for each study year.

The number of CRC late-stage diagnoses was part of this analysis, and results were stratified by age groups for each study year.

Colorectal cancer incidence was identified by using cancer registry "recodes," which are a variable derived from the cancer primary site variable (cncr_primarysit) and the cancer histology variable (cncr_histology3). The National Cancer Institute designates recodes by group, and the information is documented on their site.ⁱⁱⁱ Colon and rectum recodes from SEER Site Recode ICD-O-3/WHO 2008 were used to identify the number of individuals who had a CRC diagnosis in 2016 and 2017. Cancer registry data was queried to include all colorectal cancers in the study period by using recodes 21041, 21042, 21043, 21044, 21045, 21046, 21047, 21048, 21049, 21051, 21052, and 21060, which are designated for colorectal cancer. Inclusion criteria for record extraction for this analysis were any records with a CRC diagnosis date occurring in 2016 or 2017.

Cancer stage was obtained from the "devss2000" field, defined as "derived SEER summary stage 2000" and defined as: 0, 1 = Localized; 2, 3, 4, 7 = Advanced; 9 = Unstaged. There were none in our data with a code of 8.

To enable stratification by age group, a new field containing each patient's age at the time of diagnosis was added, then another new field separating patients into five-year age groups. To calculate rates of incidence, an age group table was created and joined to a table containing the same year/age group-specific state populations on age groups to perform the following calculations: count by unique CRC patient ID (cncr_pid), rate of incidence as $\text{count}(\text{distinct ID}/\text{state age group population estimate}) * 10,000$. This output was then organized by age groups. Age-specific population estimates were derived from the U.S. Census American Communities Survey estimates for 2016 and 2017, for the following age groups: 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, and 80 and older. These age groups were selected in part due to alignment with previous and new USPSTF guidelines for CRC screening age recommendations.

Rates of incidence were age-adjusted using the 2000 U.S. Standard Population per 100,000. To complete this step the output table containing the unadjusted incidence rates was joined to the table containing 2000 U.S. Standard Population. The new table included all fields from the previous unadjusted table as well as the following calculations of the age-weighted incidence as a proportion of total count of incidence for each age group.

ⁱⁱⁱ NIH Recodes classification list: https://seer.cancer.gov/siterecode/icdo3_dwhoheme/index.html

LIMITATIONS

As with any claims-based data analysis, this analysis is subject to inherent provider-level billing and coding variation. While this is a known limitation of healthcare claims-based data analysis, the team at ACHI uses evidence-based research methods and conducts multi-layer data and analytic validation processes. Additionally, as previously discussed, analyses with follow-up colonoscopies within six months in 2018 that included Medicare data are preliminary and only include stool-based test data for January 1–June 30, 2018, due to availability of Medicare data. The data also did not allow us to assess whether stool-based tests that were administered resulted in a positive screen, but follow-up colonoscopies observed in the the data were in the range of reported positive screens for gFOBTs, FITs, and Cologuard.^{2,3,4}

Claims-based costs of follow-up colonoscopies are not included in the analysis due to resource limitations and the complexity of defining all costs associated with a follow-up colonoscopy (across multiple insurance claim lines), as opposed to identifying follow-up colonoscopy occurrences based on singular procedure codes. Assessment of costs associated with follow-up colonoscopies warrants further analysis.

As with claims-based data, there is also an inherent lag in availability of cancer registry data. At the time of the analysis, the most recent year of available cancer registry data was 2017.

Race and ethnicity information is not readily available for the entire study population, however that information is available for the study population subset of individuals ages 65–75 who had Medicare coverage in 2017.

Appendices

APPENDIX A:

EXAMPLE OF SAS CODE FOR ADMINISTRATIVE CLAIMS-BASED ANALYSIS OF CRC SCREENINGS

Note: This code has been generalized for use on other data sources. The original code was developed to be utilized on the Arkansas APCD.

Step 1.

Determine which denominator will best suit your policy question and the data that you have available. For example, state population estimates may be most suitable for your analyses.

Step 2.

*/*Extract medical claims that contain one of the screening procedure codes, restricting by time period of interest.*/*

```
%let screening_procs = "G0328" "G0464" "82270" "82274" "81528";
```

```
proc sql;
```

```
create table stool_tests as
```

```
select distinct *
```

```
  ,case
```

```
    when ProcedureCode in ("82270") then "gFOBT"
```

```
    when ProcedureCode in ("G0328", "82274") then "FIT"
```

```
    when ProcedureCode in ("G0464") then "FIT-DNA"
```

```
    when ProcedureCode in ("81528") then "Cologuard"
```

```
end as Test_Type,
```

```
  ,year(DateOfServiceFrom) as stool_test_year
```

```
  ,DateOfServiceFrom as stool_test_date " format=date9.,
```

```
  ,year(DateOfServiceFrom) - year(MemberDateOfBirth) as test_age
```

```
  ,case
```

```
    when Calculated test_age le 44 then 'less 44'
```

```
    when 45 le Calculated test_age le 49 then '45 to 49'
```

```
    when 50 le Calculated test_age le 54 then '50 to 54'
```

```
    when 55 le Calculated test_age le 59 then '55 to 59'
```

```
    when 60 le Calculated test_age le 64 then '60 to 64'
```

```
    when 65 le Calculated test_age le 69 then '65 to 69'
```

```
    when Calculated test_age ge 70 then '70 plus'
```

```
end as test_age_cat
```

```
  ,case
```

```
    when 50 le Calculated test_age le 55 then '50 to 55'
```

```
end as Fifty_to_55
```

```
  ,sum(CopayAmount,CoinsuranceAmount,DeductibleAmount) as Amt_OOP
```

```
  ,sum(CopayAmount,CoinsuranceAmount,DeductibleAmount, PaidAmount) as Total_Paid
```

```
from medical_claims
```

```

        where year(DateOfServiceFrom) in (2016,2017,2018,2019) and ProcedureCode in
        (&screening_procs.);
quit;

/*Extract medical claims that contain one of the follow-up procedure codes, restricting by time
period of interest.*/
%let follow_up_procs = "45385" "45380" "45378" "G0121" "G0105" "45384" "45388" "45385"
"45380" "45378";
proc sql;
create table Follow_up_extract as
select distinct *
, sum(CopayAmount, CoinsuranceAmount, DeductibleAmount) as Amt_OOP
, sum(CopayAmount, CoinsuranceAmount, DeductibleAmount, PaidAmount) as Total_Paid
from medical_claims
where year(DateOfServiceFrom) in (2016,2017,2018,2019) and ProcedureCode in
(&follow_up_procs.);
quit;

/*Combine information from screening procedure claims with follow-up procedure claims
joining on patient ID. Create indicators identifying follow-up procedures occurring within 1, 3,
and 6 months from the stool test.*/
proc sql;
create table Follow_ups as /* 52,978 */
select distinct *
/*Flag indicating that follow-up occurred within 6 months of stool test.*/, case
when b.stool_test_date le DateOfServiceFrom le (b.stool_test_date + 180) then 1
else 0
end as Six_Month_Followup
/*Flag indicating that follow-up occurred within 3 months of stool test.*/, case
when b.stool_test_date le DateOfServiceFrom le (b.stool_test_date + 90) then 1
else 0
end as Three_Month_Followup
/*Flag indicating that follow-up occurred within 1 month of stool test.*/, case
when b.stool_test_date le DateOfServiceFrom le (b.stool_test_date + 30) then 1
else 0
end as One_Month_Followup
, DateOfServiceFrom - b.stool_test_date as Days_after_Stool_Test
from Follow_up_extract a left join stool_tests b on a.patient_id = b.patient_id
where not missing(stool_test_date);
quit;

/*Calculate measures of interest. */
proc sql;
/*number of individuals with a stool-based test*/

```

```

create table stool_test as
  select distinct stool_test_year, count(distinct patient_id) as Stool_tests
  from stool_tests
  where stool_test_year = &yr.;

/*number of individuals with a stool-based test that was at least partially paid for by a payer*/
create table Stool_Amt_Payer_Paid as
  select distinct stool_test_year, count(distinct patient_id) as Stool_Amt_Payer_Paid
  from stool_tests
  where stool_test_year = &yr. and PaidAmount > 0;

/*number of individuals with a stool-based test that had an out of pocket cost, regardless of
payer paid amount*/
create table Stool_Amt_Out_ofPocket as
  select distinct stool_test_year, count(distinct patient_id) as Stool_Amt_Out_ofPocket
  from stool_tests
  where stool_test_year = &yr. and Amt_OOP > 0;

/*number of individuals with a follow-up procedure*/
create table Follow_up as
  select distinct stool_test_year, count(distinct patient_id) as Follow_up
  from Follow_up
  where stool_test_year = &yr.
  group by stool_test_year;

/*number of individuals with a follow-up procedure that was at least partially paid for by a
payer*/
create table Follow_up_Amt_Payer_Paid as
  select distinct stool_test_year, count(distinct patient_id) as Follow_up_Amt_Payer_Paid
  from rdrive.Follow_up
  where and stool_test_year = &yr. and PaidAmount > 0
  group by stool_test_year;

/*number of individuals with a follow-up procedure that had an out of pocket cost, regardless
of payer paid amount */
create table Follow_up_Amt_Out_ofPocket as
  select distinct stool_test_year, count(distinct patient_id) as Follow_up_Amt_Out_ofPocket
  from rdrive.Follow_up
  where and stool_test_year = &yr. and Amt_OOP > 0
  group by stool_test_year;
quit;

/*Build an output table grouped by stool test year.*/
proc sql;

```

```

create table Year_&yr. as
  select distinct a.stool_test_year
/*chosen denominator based on available options*/
  ,denominator
/*distinct number of individuals with a stool test*/
  ,b.Stool_tests
/*percent of individuals with a stool test*/
  ,b.Stool_tests/denominator as Rate_of_Stool_test
/*distinct number of individuals with any stool test cost covered by payer (partially or totally)*/
  ,c.Stool_Amt_Payer_Paid as Stool_Amt_Payer_Paid
/*percent of individuals with payer-paid stool test */
  ,c.Stool_Amt_Payer_Paid/b.Stool_tests as Rate_paid_Stool
/*distinct number of individuals with any out of pocket costs for stool tests*/
  ,d.Stool_Amt_Out_ofPocket
/*percent of members who had stool test with out-of-pocket costs */
  ,d.Stool_Amt_Out_ofPocket/b.Stool_tests as Pct_Stool_Amt_Out_ofPocket
/*distinct number of individuals with a follow up colonoscopy*/
  ,e.Follow_up
/*distinct number of individuals whose follow-up colonoscopy was paid by payer (partially or
totally)*/
  ,f.Follow_up_Amt_Payer_Paid as Follow_up_paid
/*percent of individuals who had a follow up colonoscopy after a stool test*/
  ,e.Follow_up/b.Stool_tests as Pct_Follow_up
/*percent of individuals whose follow up colonoscopy was paid by payer (partially or totally)*/
  ,f.Follow_up_Amt_Payer_Paid/e.Follow_up as Pct_Follow_up_Amt_Payer_Paid
/*distinct number of individuals with any out of pocket costs for follow ups*/
  ,g.Follow_up_Amt_Out_ofPocket
/*percent of individuals who had follow-up colonoscopy with out-of-pocket costs*/
  ,g.Follow_up_Amt_Out_ofPocket/f.Follow_up as Pct_Follow_up_Amt_Out_ofPocket
/*distinct number of individuals who had follow-up colonoscopy without any out-of-pocket
cost*/
  ,e.Follow_up - g.Follow_up_Amt_Out_ofPocket as Follow_No_OOP
/*percent of individuals who had follow-up colonoscopy without any out-of-pocket costs */
  ,Calculated Follow_No_OOP/e.Follow_up as Pct_Follow_No_OOP
  from stool_tests a
  left join stool_test b on a.stool_test_year = b.stool_test_year
  left join Stool_Amt_Payer_Paid c on a.stool_test_year = c.stool_test_year
  left join Stool_Amt_Out_ofPocket d on a.stool_test_year = d.stool_test_year
  left join Follow_up e on a.stool_test_year = e.stool_test_year
  left join Follow_up_Amt_Payer_Paid f on a.stool_test_year = f.stool_test_year
  left join Follow_up_Amt_Out_ofPocket g on a.stool_test_year = g.stool_test_year;
quit;

```

APPENDIX B:

LIST OF PROCEDURE CODES USED FOR ADMINISTRATIVE CLAIMS-BASED ANALYSIS OF CRC SCREENINGS

Screening Type	Code	Code System	Definition
Stool-based: FOBT	G0328	HCPCS	Colorectal cancer screening; fecal occult blood test, immunoassay, 1-3 simultaneous (G0328)
Stool-based: FIT-DNA	G0464	HCPCS	Colorectal cancer screening; stool-based DNA and fecal occult hemoglobin (e.g., kras, ndrg4 and bmp3) (G0464)
Stool-based: FOBT	82270	CPT	Blood, occult, by peroxidase activity (e.g., guaiac), qualitative; feces, consecutive collected specimens with single determination, for colorectal neoplasm screening (i.e., patient was provided 3 cards or single triple card for consecutive collection)
Stool-based: FOBT	82274	CPT	Blood, occult, by fecal hemoglobin determination by immunoassay, qualitative, feces, 1-3 simultaneous determinations
Stool-based: Cologuard	81528	CPT	Oncology (colorectal) screening, quantitative real-time target and signal amplification of 10 DNA markers (KRAS mutations, promoter methylation of NDRG4 and BMP3) and fecal hemoglobin, utilizing stool, algorithm reported as a positive or negative result.
Colonoscopy	45385	CPT	Colonoscopy, flexible, proximal to splenic flexure; with removal of tumor(s), polyp(s) or other lesion(s) by snare technique
Colonoscopy	45380	CPT	Colonoscopy, flexible, proximal to splenic flexure, with biopsy, single or multiple
Colonoscopy	45378	CPT	Colonoscopy, flexible, proximal to splenic flexure; diagnostic, with or without collection of specimen(s) by brushing or washing, with or without colon decompression (separate procedure)
Colonoscopy	G0121	HCPCS	Colorectal cancer screening; colonoscopy on individual not meeting criteria for high risk
High-Risk Colonoscopy	G0105	HCPCS	Colorectal cancer screening; colonoscopy on individual at high risk
Colonoscopy	45384	CPT	Colonoscopy, flexible, proximal to splenic flexure; with removal of tumor(s), polyp(s) or other lesion(s) by hot biopsy forceps or bipolar cautery
Colonoscopy	45388	CPT	Colonoscopy, flexible; with ablation of tumor(s), polyp(s), or other lesion(s) (includes pre and post-dilation and guide wire passage, when performed)
Barium Enema	45385	CPT	Radiologic examination, colon; contrast (e.g., barium) enema, with or without KUB
Barium Enema	45380	CPT	Radiologic examination, colon; air contrast with specific high density barium, with or without glucagon
Barium Enema	45378	CPT	Colorectal cancer screening; alternative to G0105, screening colonoscopy, barium enema

APPENDIX C:

EXAMPLE OF SAS CODE FOR ANALYSIS OF CANCER REGISTRY DATA FOR CRC INCIDENCE AND STAGE AT DIAGNOSIS

Note: This code is based on the Arkansas Cancer Registry Database

Step 1.

Prepare the 2000 US Standard Million Population table downloaded from <https://seer.cancer.gov/stdpopulations/stdpop.19ages.html>. This table was copied into an Excel file and prepared for use by combining the age groups of 00 years and 01-04 years into a 0 to 5 years age group and combining the age groups of 80-85 years and 85+ years into an 80+ years age group.

Prepare a state population table using the same age groups as the standard table.

Step 2.

```
/*Select all records that have a recode that identifies Colorectal Cancer, restricting by time period of interest.*/
```

```
%let year_of_interest = 2016;
```

```
%let yr = 16;
```

```
proc sql;
```

```
create table coloreccancer&yr as
```

```
select *
```

```
from cancer_registry_table
```

```
where recode in (21041, 21042, 21043, 21044, 21045, 21046, 21047, 21048, 21049,  
21051, 21052, 21060) and cnr_dxdate = &year_of_interest;
```

```
quit;
```

```
proc sql;
```

```
create table colorecagegroups&yr as
```

```
select distinct *
```

```
/*Create an age group field that aligns with the age field in the 2000 US Standard Million Population table to link the two tables together.*/
```

```
, case
```

```
when 0 le cnr_dxage < 5 then "0-4"
```

```
when 5 le cnr_dxage < 10 then "5-9"
```

```
when 10 le cnr_dxage < 15 then "10-14"
```

```
when 15 le cnr_dxage < 20 then "15-19"
```

```
when 20 le cnr_dxage < 25 then "20-24"
```

```
when 25 le cnr_dxage < 30 then "25-29"
```

```
when 30 le cnr_dxage < 35 then "30-34"
```

```
when 35 le cnr_dxage < 40 then "35-39"
```

```
when 40 le cnr_dxage < 45 then "40-44"
```

```
when 45 le cnr_dxage < 50 then "45-49"
```

```
when 50 le cnr_dxage < 55 then "50-54"
```

```
when 55 le cnr_dxage < 60 then "55-59"
```

```

    when 60 le cnr_dxage < 65 then "60-64"
    when 65 le cnr_dxage < 70 then "65-69"
    when 70 le cnr_dxage < 75 then "70-74"
    when 75 le cnr_dxage < 80 then "75-79"
    when cnr_dxage ge 80 then "80+"
    else "?"
end as AgeGroup
/*Create an cancer stage field based on the summary stage field in the cancer table*/
, case
    when cnr_summarystg00 in ('0', '1') then "Localized"
    when cnr_summarystg00 in ('2', '3', '4', '7') then "Advanced"
    else "Unstaged"
end as Stage
    from coloreccancer&yr;
quit;

```

/*The code below will produce a table with age group, number of individuals in age group with CRC, the state population per age group, the crude rate, the 2000 US Standard Population, the proportion of the US Standard Population, and the age-weighted incidence per 100,000 of the 2000 US Standard Population.*/

```

proc sql;
select sum(US_STD_POP_2000) into :uspop from standardpop;

create table agegroup&yr as
select distinct
a.AgeGroup
/*number of individuals in age group with CRC*/
,count(distinct a.cnr_pid) as incidence_of_cancer
/*the state population per age group*/
,b.StatePop format = comma16.
/*crude rate*/
,(count(distinct a.cnr_pid)/b.StatePop)*100000 as crude_incidence_rate
/*2000 US Standard Population*/
,b.US_STD_POP_2000
/*proportion of the US Standard Population*/
,(c.US_STD_POP_2000/&uspop.) as USstandardpopulation format = 8.3
    label = 'Proportion of US Standard Population'
/*the age-weighted incidence per 100,000 of the 2000 US Standard Population*/
,(c.US_STD_POP_2000/&uspop.)*(count(distinct a.cnr_pid)/b.StatePop)*100000) as weighted
    label = 'Age-weighted incidence'
from colorecagegroups&yr a left join statepop b on a.AgeGroup = b.AgeGroup
    left join standardpop c on a.AgeGroup = c.AGE_GROUP
group by a.AgeGroup;

```

```
quit;
```

```
/*The code below will produce a table with stage, number of individuals in stage group with  
CRC, the crude rate, and the age-weighted incidence per 100,000 of the 2000 US Standard  
Population.*/
```

```
proc sql;
```

```
select sum(US_STD_POP_2000) into :uspop from standardpop;
```

```
/*Calculating age weighted adjustments by stage group.*/
```

```
create table stageagegroup&yr as
```

```
select distinct
```

```
  a.AgeGroup
```

```
  a.Stage
```

```
  ,count(distinct a.cnr_pid) as incidence_of_cancer
```

```
  ,b.StatePop format = comma16.
```

```
  ,(count(distinct a.cnr_pid)/b.StatePop)*100000 as crude_incidence_rate
```

```
  ,b.US_STD_POP_2000
```

```
  ,(c.US_STD_POP_2000/&uspop.) as USstandardpopulation format = 8.3
```

```
    label = 'Proportion of US Standard Population'
```

```
  ,(c.US_STD_POP_2000/&uspop.)*(count(distinct a.cnr_pid)/b.StatePop)*100000) as weighted
```

```
    label = 'Age-weighted incidence'
```

```
  from colorecagegroups&yr a left join statepop b on a.AgeGroup = b.AgeGroup
```

```
    left join standardpop c on a.AgeGroup = c.AGE_GROUP
```

```
  group by a.AgeGroup, a.Stage;
```

```
quit;
```

```
/*Determining final age weighted adjustments by stage group.*/
```

```
proc sql;
```

```
create table stagepop_rolled_&yr as
```

```
select distinct
```

```
  Stage
```

```
  , sum(incidence_of_cancer) as incidence_of_cancer_bystage
```

```
  , sum(crude_incidence_rate) as CrudeRate format=6.2
```

```
  , sum(weighted) as AgeWeightedRate format=6.2
```

```
  from stageagegroup&yr
```

```
  group by Stage;
```

```
quit;
```


APPENDIX D:

LIST OF RECODES USED FOR ANALYSIS OF CANCER REGISTRY DATA FOR CRC INCIDENCE AND STAGE AT DIAGNOSIS

Site Group	ICD-O-3 Site	ICD-O-3 Histology (Type)	Recode
Colon excluding Rectum			
Cecum	C180	excluding 9050-9055, 9140, 9590-9992	21041
Appendix	C181		21042
Ascending Colon	C182		21043
Hepatic Flexure	C183		21044
Transverse Colon	C184		21045
Splenic Flexure	C185		21046
Descending Colon	C186		21047
Sigmoid Colon	C187		21048
Large Intestine, NOS	C188-C189, C260		21049
Rectum and Rectosigmoid Junction			
Rectosigmoid Junction	C199	excluding 9050-9055, 9140, 9590-9992	21051
Rectum	C209		21052
Anus, Anal Canal and Anorectum	C210-C212, C218		21060

¹ U.S. Preventive Services Task Force. Final Recommendation Statement, Colorectal Cancer: Screening. May 18, 2021. <https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/colorectal-cancer-screening>

² Lin JS, Perdue LA, Henrikson NB, et al. Screening for Colorectal Cancer: An Evidence Update for the U.S. Preventive Services Task Force [Internet]. Rockville (MD): Agency for Healthcare Research and

Quality (U.S.); 2021 May. (Evidence Synthesis, No. 202.) Appendix D, Tables. Available from:
<https://www.ncbi.nlm.nih.gov/books/NBK570907/>

³ Finney Rutten et al. (September 2020). Colorectal cancer screening completion: An examination of differences by screening modality. *Prev Med Rep.* 2020;20:101202. doi:10.1016/j.pmedr.2020.101202

⁴ Imperiale, T.F. et al. 2014. Multitarget Stool DNA Testing for Colorectal-Cancer Screening. *NEJM.*
<https://www.nejm.org/doi/full/10.1056/nejmoa1311194#>