
National Health Statistics Reports

Number 69 ■ November 1, 2013

Self-report of Diabetes and Claims-based Identification of Diabetes Among Medicare Beneficiaries

by Hannah R. Day, Ph.D., and Jennifer D. Parker, Ph.D., Office of Analysis and Epidemiology

Abstract

Objective—This report compares self-reported diabetes in the National Health Interview Survey (NHIS) with diabetes identified using the Medicare Chronic Condition (CC) Summary file.

Background—NHIS records have been linked with Medicare data from the Centers for Medicare & Medicaid Services. The CC Summary file, one of several linked files derived from Medicare claims data, contains indicators for chronic conditions based on an established algorithm.

Methods—This analysis was limited to 2005 NHIS participants aged 65 and over whose records were linked to 2005 Medicare data. Linked NHIS participants had at least 1 month of fee-for-service Medicare coverage in 2005. Concordance between self-reported diabetes and the CC Summary indicator for diabetes is compared and described by demographics, socioeconomic status, health status indicators, and geographic characteristics.

Results—Of the Medicare beneficiaries in the 2005 NHIS, 20.0% self-reported diabetes and 27.8% had an indicator for diabetes in the CC Summary file. Of those who self-reported diabetes in NHIS, the percentage with a CC Summary indicator for diabetes was high (93.1%). Of those with a CC Summary indicator for diabetes, the percentage self-reporting diabetes was comparatively lower (67.0%). Statistically significant differences by subgroup existed in the percentage concordance between the two sources. Of those with self-reported diabetes, the percentage with a CC Summary indicator differed by sex and age. Of those with a CC Summary indicator for diabetes, the percentage with self-reported diabetes differed by age, self-rated health, number of self-reported conditions, and geographic location.

Conclusions—Among Medicare beneficiaries who self-reported diabetes in NHIS, a high concordance was observed with identification of diabetes in the CC Summary file. However, among Medicare beneficiaries with an indicator for diabetes in the CC Summary file, concordance with self-reported diabetes in NHIS is comparatively lower. Differences exist by subgroup.

Keywords: record linkage • National Health Interview Survey • Chronic Condition Summary file

Introduction

Diabetes is an endocrine disorder characterized by impaired insulin production or insulin resistance, which leads to increased blood glucose levels (1). In adults aged 65 and over, diabetes can be classified as Type 1 (previously called juvenile diabetes), or Type 2 (previously called adult-onset diabetes) (1). Type 1 diabetes is an autoimmune disorder with a strong genetic component, and Type 2 diabetes is an acquired disease with evidence of genetic components. Type 2 diabetes is responsible for 90%–95% of diabetes cases in the United States (1). Risk factors for Type 2 diabetes include age, obesity, ethnicity, and insufficient physical activity (1). Complications of the disease include heart disease, myocardial infarction, stroke, lower extremity infection, amputation, and blindness (1–3). In older adults, diabetes is associated with higher morbidity and all-cause mortality (2). Diabetes is one of the 10 most expensive disease categories among Medicare beneficiaries (4). Higher costs result from greater service utilization, especially more hospital stays and longer lengths of stay, more physician visits, and greater



U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics



medication use compared with adults not having diabetes (5).

Most studies of diabetes in the Medicare population use administrative Medicare records collected for billing purposes. The use of Medicare claims to identify people with diabetes has been examined using the Medicare Current Beneficiary Survey (MCBS), a nationally representative sample of the Medicare population with survey records linked to Medicare claims. One recent MCBS study using survey data from 2004 estimated that 20.3% of noninstitutionalized Medicare beneficiaries reported being told by a doctor that they had diabetes (6). A previous comparison of MCBS self-report data for 1992–1993 with Medicare claims data from 1992–1993 found the percentage of patients identified in claims data varied widely depending on the algorithm and reference period used (7). Using an algorithm looking for “one or more diagnoses of diabetes in any Medicare claim file” over a 2-year period, 79% of those who self-reported diabetes were identified in the Medicare data, but only 71% of those with a claim for diabetes in the Medicare data had self-reported diabetes (7). These studies highlight that concordance between self-report and claims data is not always 100%. People who self-report diabetes may not always be identified by Medicare claims, and not all people who have Medicare claims for diabetes will self-report the disease (7). A previous National Center for Health Statistics (NCHS) study using data from the Health Interview Evaluation Survey, conducted in 1990, compared self-report with medical records and found more instances where diabetes was present in the medical record but not reported in the interview, than instances where diabetes was reported in the interview but not present in the medical record (8).

Differences between survey and administrative data can lead to varying prevalence estimates, with estimates based on self-report often being lower than claims-based estimates. Examples of diabetes prevalence estimates for noninstitutionalized adults aged 65 and over based on self-report include the

20.3% prevalence estimate in the noninstitutionalized 2004 MCBS population, an estimate of 20.5% in adults aged 65 and over from the 2011 National Health Interview Survey (NHIS), and an estimate of 17.0% in the interview component of the 2005–2006 National Health and Nutrition Examination Survey (NHANES) (6,7,9,10). In addition to identifying diabetes using self-report or claims, the NHANES study also conducts examinations with fasting plasma glucose and oral glucose tolerance tests. In NHANES, as much as 46.2% of diabetes was neither diagnosed nor self-reported (9). The prevalence for adults aged 65 and over based on examination data in 2005–2006 NHANES was nearly twice that based on interview, at 31.6% (9).

In contrast, based on Medicare claims data, Margolis et al. estimated a diabetes prevalence of 26.5%–28.0% in Medicare fee-for-service (FFS) beneficiaries aged 65 and over in 2006–2008 (11), and another study estimated a prevalence of 24.3% in adult Medicare beneficiaries of all ages using the 2005 Medicare Chronic Conditions (CC) Summary file (12).

The differences between self-report, administrative, and examination data have important implications for researchers who wish to use either survey or administrative data to examine chronic diseases in subpopulations, or who wish to adjust for diabetes in these subpopulations. Knowing which subpopulations differ significantly with respect to what percentage of participants who self-report have administrative indicators of diabetes, or the percentage of participants with administrative indicators who self-report diabetes, will help researchers know the limitations of using either source of data in their studies and help them choose the best data source, or combination of data sources, for their study population.

There are strengths and weaknesses in using either self-report or administrative databases for studies of diabetes. Self-report data include detailed information gathered directly from participants and are subject to possible response bias. However,

population health surveys provide additional information on characteristics not included in administrative data, including income, education, and marital status. In contrast, administrative databases often include detailed cost and service use information and allow for longitudinal tracking of patients, but administrative data are created for billing purposes and often lack demographic, health, and risk factor information (13). Further, although Medicare covers the majority of adults aged 65 and over, the Centers for Medicare & Medicaid Services (CMS) receives only the detail needed to identify diabetes from claims for FFS beneficiaries. Medicare beneficiaries have the option to enroll in traditional FFS or in Medicare Advantage (MA)—also referred to as managed care organizations or Medicare Part C. In 2005, 85% of Medicare beneficiaries were in the FFS program (14). Some studies have found FFS beneficiaries to be older and in poorer health (15–17); however, other studies have shown that the difference between MA and FFS may be decreasing (18,19). The accuracy of survey response in FFS compared with MA enrollment is unknown.

The CC Summary file was created from CMS claims to allow researchers to easily identify Medicare beneficiaries with claims for 21 selected chronic conditions (including diabetes, selected cancers, heart disease, hip fracture, and others), based on published algorithms. To identify individual diseases or conditions from claims data without the CC Summary file, multiple Medicare files must be combined, using various algorithms to identify claims specific to each disease (7). The CC Summary file has been used to study prevalence of comorbidity (20), multiple chronic conditions (12), functional capacity of home health care patients (21), arthritis screening and prevention (22), and other topics important to the Medicare population (23).

The goal of this report is to compare self-reported diabetes in the 2005 NHIS—currently the most recent year of NHIS linked to Medicare data—with an indicator for diabetes in the 2005 CC Summary file. The NHIS–Medicare linked data provide a

unique opportunity to examine and compare the prevalence of diabetes derived from each data source. The large size of NHIS and its detailed questionnaire allow identification of demographic, geographic, and health-related factors associated with discrepancies between the data sources.

Data Sources

NHIS

NHIS uses a cross-sectional, multistage area probability design to collect data on demographic and health-related information from the civilian noninstitutionalized U.S. population. NHIS has been conducted continuously since 1957 by the Centers for Disease Control and Prevention's NCHS, although both the survey design and instrument have changed over time. Currently, the design oversamples black, Hispanic, and Asian persons. The NHIS sample is currently drawn from each state and the District of Columbia. The survey consists of household, family, Sample Adult, Sample Child, and supplemental components. Limited demographic information is collected on all individuals living in a house for the household file. The family component of the questionnaire collects demographic information on each member in a family from a knowledgeable adult (family respondent), along with information on health status, limitations, health care access utilization, and insurance. From each family, one sample adult is randomly selected to answer more in-depth questions about their own health. Information on one sample child per family is collected if children are present in the family.

Self-reported diabetes in the 2005 NHIS was measured by asking sample adults, "Other than during pregnancy, have you EVER been told by a doctor or health professional that you have diabetes or sugar diabetes?" Participants could respond "yes," "no," "borderline diabetes," "refused," or "don't know." For this report, participants were considered to have diabetes if they responded "yes"; participants who

responded "no" were considered not to have diabetes by self-report.

Descriptive characteristics examined in this report include demographic characteristics (age, race and ethnicity, marital status, and sex), socioeconomic status (education and poverty level), geographic characteristics (metropolitan residence status and region of country), and health risk factor variables (self-rated health, number of reported chronic health conditions, and obesity). Age at the time of interview was categorized into two groups: 65–74, and 75 and over.

Race and ethnicity are collected separately in NHIS, first by asking if the participant is Hispanic or Latino, then by asking the participant's race or races. For this report, race and ethnicity were combined into four categories: non-Hispanic white, non-Hispanic black, Hispanic, and all other races and ethnicities. Estimates for beneficiaries reporting all other races and ethnicities are not presented due to the small number of people in this group; however, these beneficiaries are included in statistics based on other variables.

Socioeconomic status was represented by examining education and poverty level. Two education categories were created: high school education or lower (no high school diploma, high school diploma, or General Educational Development high school equivalency certificate), and some college or more. Because income data are often missing, multiply imputed income files were used to categorize annual family income into groups of below 100% of the federal poverty level (poor), 100%–199% of poverty level (near poor), or 200% of poverty level or more (non-poor). Five sets of multiply imputed income were combined and analyzed using established methods (24).

Geographic variables of metropolitan residence status [within a metropolitan statistical area (MSA) and outside an MSA] and region of country (Northeast, Midwest, South, and West) were used. Definitions of MSA compared with outside an MSA used the 2006 NCHS urban–rural classification scheme based on the 2000 census and

Office of Management and Budget standards for defining metropolitan and micropolitan statistical areas (25). The NCHS Urban–Rural Classification Scheme for Counties includes six categories: large central metropolitan, large fringe metropolitan, medium metropolitan, small metropolitan, micropolitan, and noncore. Due to small cell sizes, subgroups smaller than the metropolitan categories (large central metropolitan, large fringe metropolitan, medium metropolitan, and small metropolitan) were not examined compared with nonmetropolitan categories (micropolitan and noncore) (25).

Health risk factor variables from NHIS included participant self-perceived health status (fair or poor compared with excellent, very good, or good), obesity [body mass index (BMI) greater than 30 compared with 29 or less], and number of reported health conditions. The total number of health conditions variable was created from a set of self-report NHIS variables and was dichotomized into 0–1 or 2 or more conditions. Conditions included in the total number of health conditions variable included hypertension, coronary heart disease, angina, myocardial infarction, stroke, emphysema, asthma, kidney disease, liver disease, and cancers other than nonmelanoma skin cancer. All NHIS self-reported chronic conditions that were similar to those in the CC Summary file were included in this count. Because the self-reported conditions and the CC Summary file differed slightly with regard to the conditions available, this count was intended to be an indicator of self-reported number of conditions only.

Linked NCHS–CMS data

The NCHS record linkage program has linked various NCHS surveys to administrative records from CMS and the Social Security Administration (SSA) under an interagency agreement among NCHS, CMS, SSA, and the Office of the Assistant Secretary for Planning and Evaluation (26). The 2005 NHIS has been linked to 1999–2007 Medicare claims (26). NHIS participants

who had sufficient personal identifier information (PII), who did not refuse to provide their Social Security number (SSN) or Health Insurance Claim number, and who had an SSN verified by the SSA Enumeration Verification System were eligible for linkage to CMS and SSA records (26). Among the linkage-eligible, if a participant's PII matched the CMS denominator file on SSN, date of birth (month, day, year), and sex, the participant was considered to be a match to CMS records (26). In 2005, the family respondent and sample adult were asked to provide their SSN. Initial data were available for 31,428 sample adults in NHIS 2005. Of these, 6,078 were aged 65 and over at the time of survey and, of these, 2,554 sample adults aged 65 and over were linkage-eligible. Of these, 98% of the sample adults ($n = 2,499$) were linked with the 2005 Medicare CC Summary file.

Available Medicare files for 1999–2007 include the denominator file (with information on demographics and plan coverage), the carrier file (noninstitutional claims and provider claims), the outpatient file (institutional outpatient claims), home health agency claims (claims for home health services), the hospice file, and a durable medical equipment file. Utilization claims data can be obtained from the Medpar file, which contains information on inpatient hospitalizations and skilled nursing facilities. The CC Summary file is available for 2005 and later, and Part D files are available for 2006 and later. In addition, NCHS creates a Summary Medicare Expenditure and Claims file using information from the denominator and claims files for use with the NCHS–Medicare linked data files.

This analysis uses the 2005 NHIS data linked to the 2005 Medicare CC Summary file and the 2005 denominator file. The 2005 Medicare denominator file includes enrollment information and program participation. Only beneficiaries enrolled in traditional FFS are included in this analysis, because Medicare does not receive claims for beneficiaries enrolled in MA programs. The CC Summary file summarizes the presence of 21 conditions by applying algorithms

to the various Medicare claims data. The 2005 CC summary file includes flags representing claims and eligibility and a variable indicating the date on which the clinical component of a condition was first recorded during 1999–2005. The algorithm used to define diabetes in the CC Summary file identified beneficiaries as having diabetes if they had, within a 2-year reference period, at least one inpatient skilled-nursing facility or home health agency claim, or two hospital outpatient or carrier claims (claims data submitted by noninstitutional providers, including physicians and supplementary providers), with codes for diabetes (27).

Methods

Analysis

Study population

The analytic sample includes linkage-eligible 2005 NHIS sample adults who were aged 65 and over at the time of interview, were linked to the CMS denominator file for 2005, and had at least 1 month of FFS claims in 2005.

Of the 2,499 sample adults aged 65 and over whose records linked with the 2005 Medicare CC Summary file, 13.3% ($n = 331$) were excluded due to not having at least 1 month of FFS coverage in 2005. The single survey participant who responded “don't know” to the NHIS question about diabetes, and the 58 participants who responded “borderline” to the question about diabetes, were not included in the primary analysis. Participants who responded “borderline” were retained for sensitivity analysis, described in the “Results” section.

In the 2005 NHIS, 20.0% of Medicare beneficiaries aged 65 and over reported they had ever had diabetes (Figure 1, Table 1).

A total of 27.9% of Medicare beneficiaries aged 65 and over in the 2005 NHIS had a 2005 Medicare CC Summary indicator for diabetes, indicating one or more claims for diabetes from 1999 through 2005 (Figure 1, Table 1).

Statistical analysis

To account for the complex survey design, SAS-callable SUDAAN version

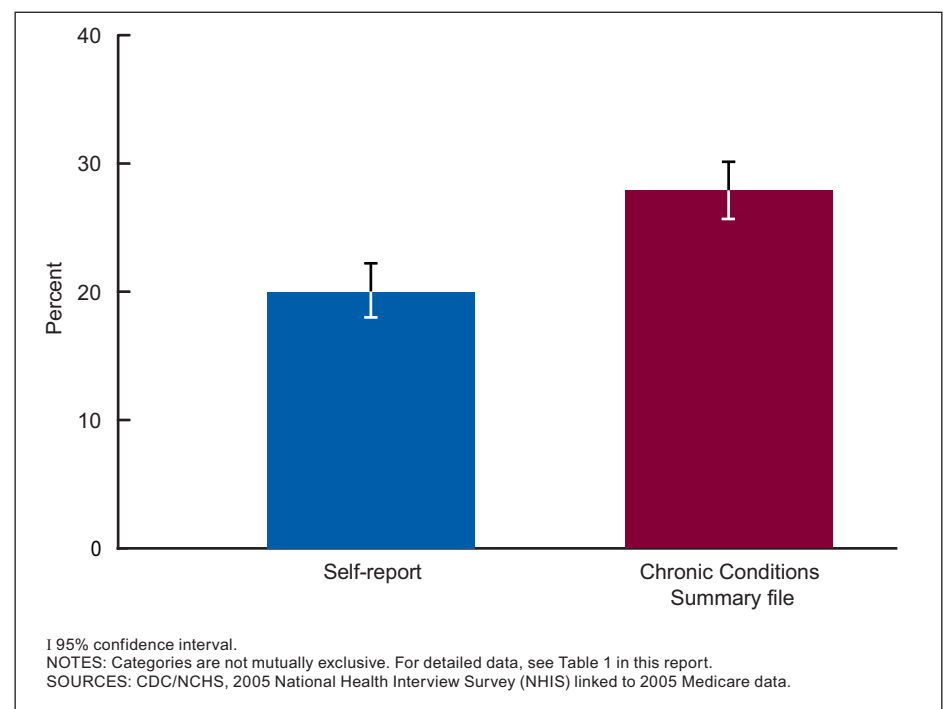


Figure 1. Diabetes by self-report or Chronic Conditions Summary indicator, among Medicare beneficiaries aged 65 and over at NHIS interview

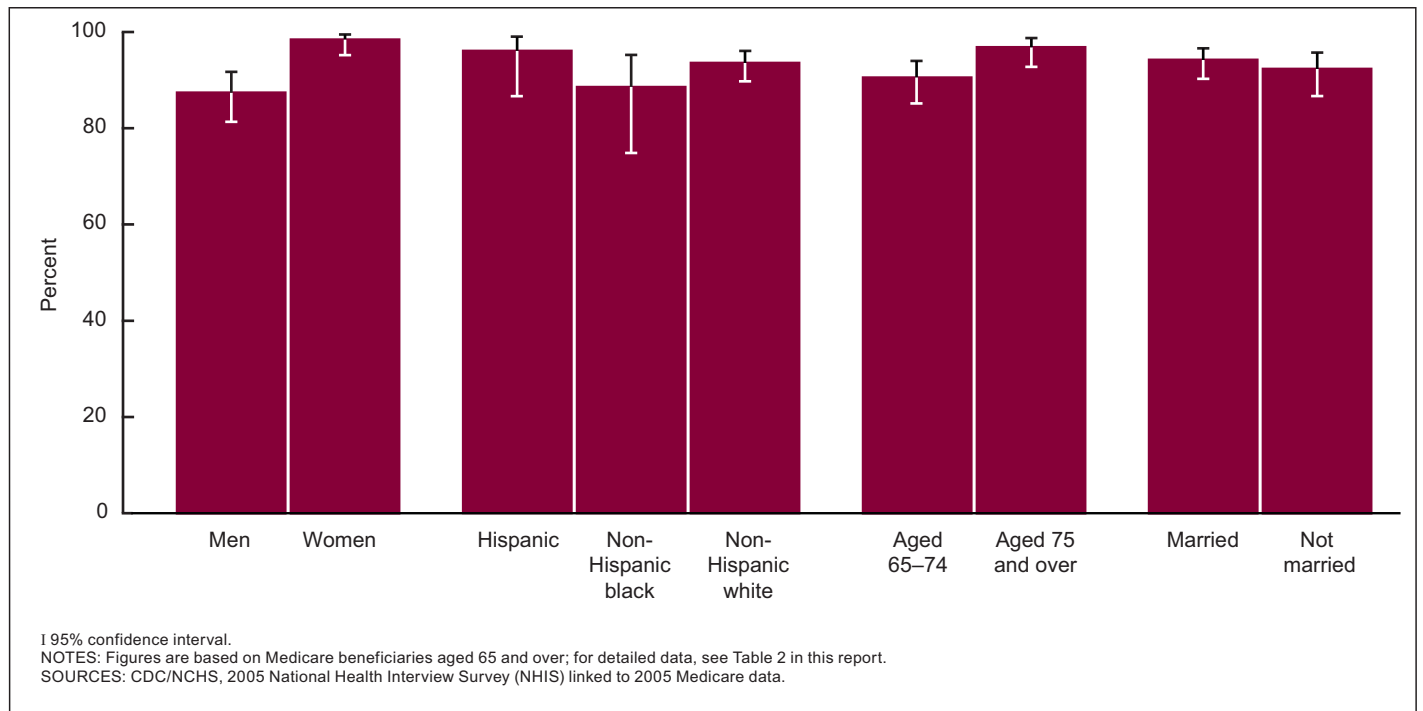


Figure 2. NHIS Medicare beneficiaries with diabetes in Chronic Conditions Summary file among those who self-reported diabetes, by demographic characteristics

9.3 statistical software was used (RTI International, Research Triangle Park, N.C.). All analyses include sample adult weights for NHIS 2005. To account for possible differences between adults who were eligible for the linkage and those who were not, sample weights were further adjusted for nonlinkage eligibility by using the SUDAAN procedure PROC WTADJUST. With PROC WTADJUST, a model-based calibration approach was used to adjust the NHIS sample adult weights for race and sex (28).

Estimates presented have a relative standard error (RSE) less than or equal to 30%. The RSE is equal to the standard error of the estimate divided by the estimate itself. The RSE is often multiplied by 100 and expressed as a percentage. RSEs greater than 30% are considered unreliable and are not presented in the report.

Differences in prevalence by CC Summary file compared with self-report were not tested for statistical significance because the two sources were not measured on independent populations. Therefore, statements comparing overall diabetes estimates from Medicare records with those from self-report should be considered as observations, and not as

conclusions of a statistical test. For comparisons among subgroups within either the CC Summary file or NHIS, statements such as “more likely” and “less likely” indicate statistical significance and are chi-squared tests, or are pairwise tests with Bonferroni corrections when chi-squared tests indicated statistical significance ($p = 0.05$) and more than two categories were present in a variable.

Results

Characteristics of those self-reporting diabetes in 2005 NHIS

- A total of 93.1% of NHIS diabetes self-reports had a 2005 CC Summary indicator for diabetes.
- Men were less likely than women to have a CC Summary indicator (Figure 2, Table 2).
- Medicare beneficiaries aged 65–74 were less likely to have a CC Summary indicator than beneficiaries aged 75 and over (Figure 2, Table 2).
- No significant differences were seen in the percentage of beneficiaries who self-reported diabetes with a CC

Summary indicator by race or ethnicity, marital status, education, poverty level, geography, or health indicators (Figures 2–5, Tables 2–5).

Characteristics of those with CC Summary indicator

- A total of 67.0% having a CC Summary indicator self-reported diabetes in the 2005 NHIS.
- Beneficiaries aged 65–74 were more likely to self-report diabetes than beneficiaries aged 75 and over (Figure 6, Table 6).
- No significant differences were observed in the percentage who self-reported diabetes by sex, race and ethnicity or marital status, education or poverty status, BMI, or geography (Figures 6–9, Tables 6–9).
- Beneficiaries in the Northeast were less likely to self-report diabetes than beneficiaries in any other region (Figure 8, Table 8).
- Beneficiaries who self-reported fair or poor health were more likely to self-report diabetes than beneficiaries in good or excellent health (Figure 9, Table 9).
- Beneficiaries who self-reported two or more conditions were more likely

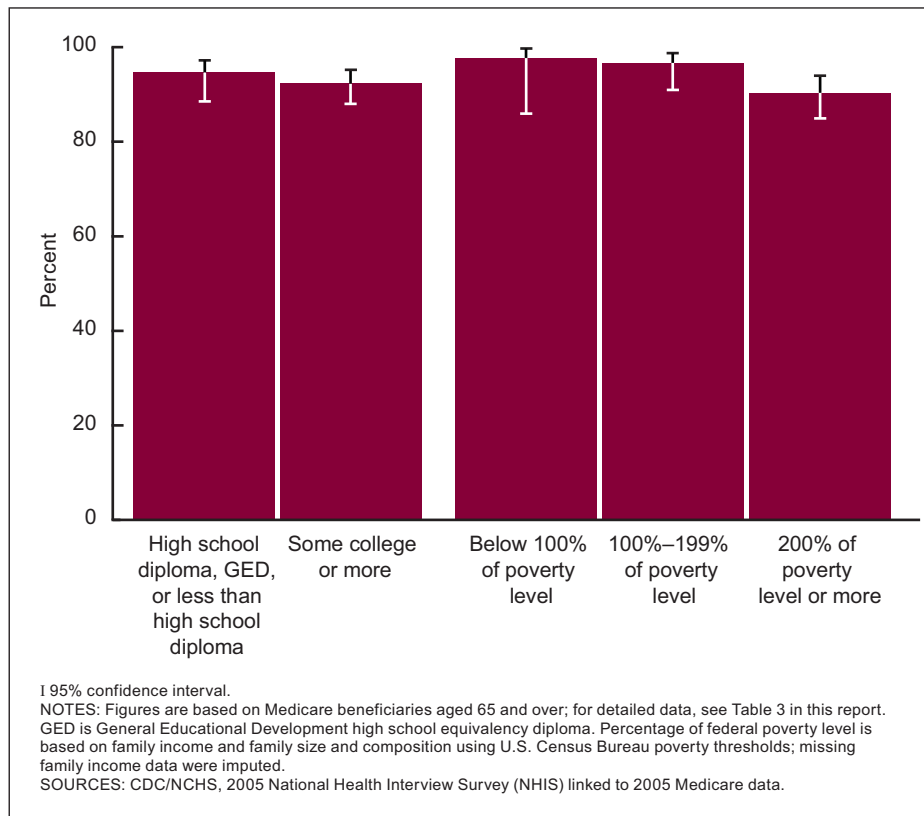


Figure 3. NHIS Medicare beneficiaries with diabetes in Chronic Conditions Summary file among those who self-reported diabetes, by education and poverty level

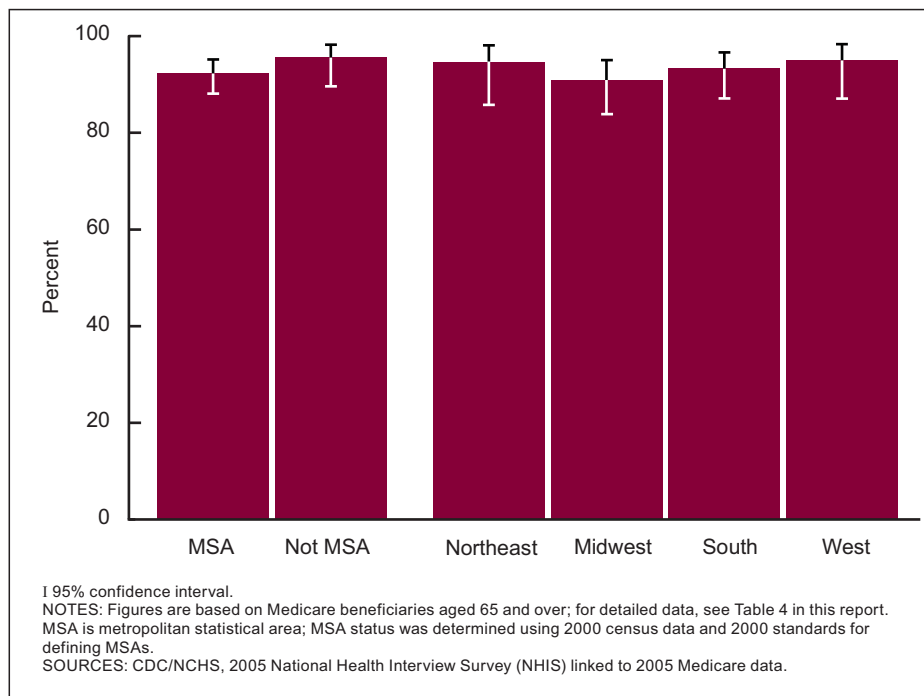


Figure 4. NHIS Medicare beneficiaries with diabetes in Chronic Conditions Summary file among those who self-reported diabetes, by metropolitan status and region

to self-report diabetes than beneficiaries with fewer conditions (Figure 9, Table 9).

Sensitivity analyses

A sensitivity analysis examining only beneficiaries with 12 months of FFS in 2005 reduced the sample size from 2,109 to 1,955; however, results remained similar. Of the beneficiaries who self-reported diabetes and had 12 months of FFS, 93.9% of the 1,955 beneficiaries had a CC Summary indicator, compared with 93.1% of the 2,109 in the full sample who reported any FFS in 2005. Of the beneficiaries with a CC Summary indicator and 12 months of FFS, 67.4% of 1,957 persons self-reported diabetes compared with 67.0% of 2,109 persons with any FFS in 2005.

A separate sensitivity analysis combined those who answered “yes” to the diabetes question with those who responded they had “borderline” diabetes. Of the 478 sample adults who responded “yes” or “borderline” to the diabetes question, 429 (90.0%) had diabetes indicated in the CC Summary file. Of the sample adults who had a CC Summary indicator for diabetes, 68.8% responded “yes” or “borderline” in NHIS. Of the 58 sample adults aged 65 and over who responded they had borderline diabetes, 39 had diabetes indicated in Medicare (65.2%) and 19 did not (34.8%).

A final sensitivity analysis revealed that only 2.7% of the 2,109 beneficiaries in the full sample had a first recorded date in the database of 2005, the survey year. Of those who had a first recorded date of 2005, 39% self-reported diabetes.

In addition to examining beneficiaries with a CC Summary indicator among those who self-reported diabetes, and beneficiaries who self-reported diabetes among those with a CC Summary indicator, beneficiaries who had a CC Summary indicator but did not self-report diabetes were also examined. The rationale was to determine if beneficiaries who had a CC Summary indicator but did not self-

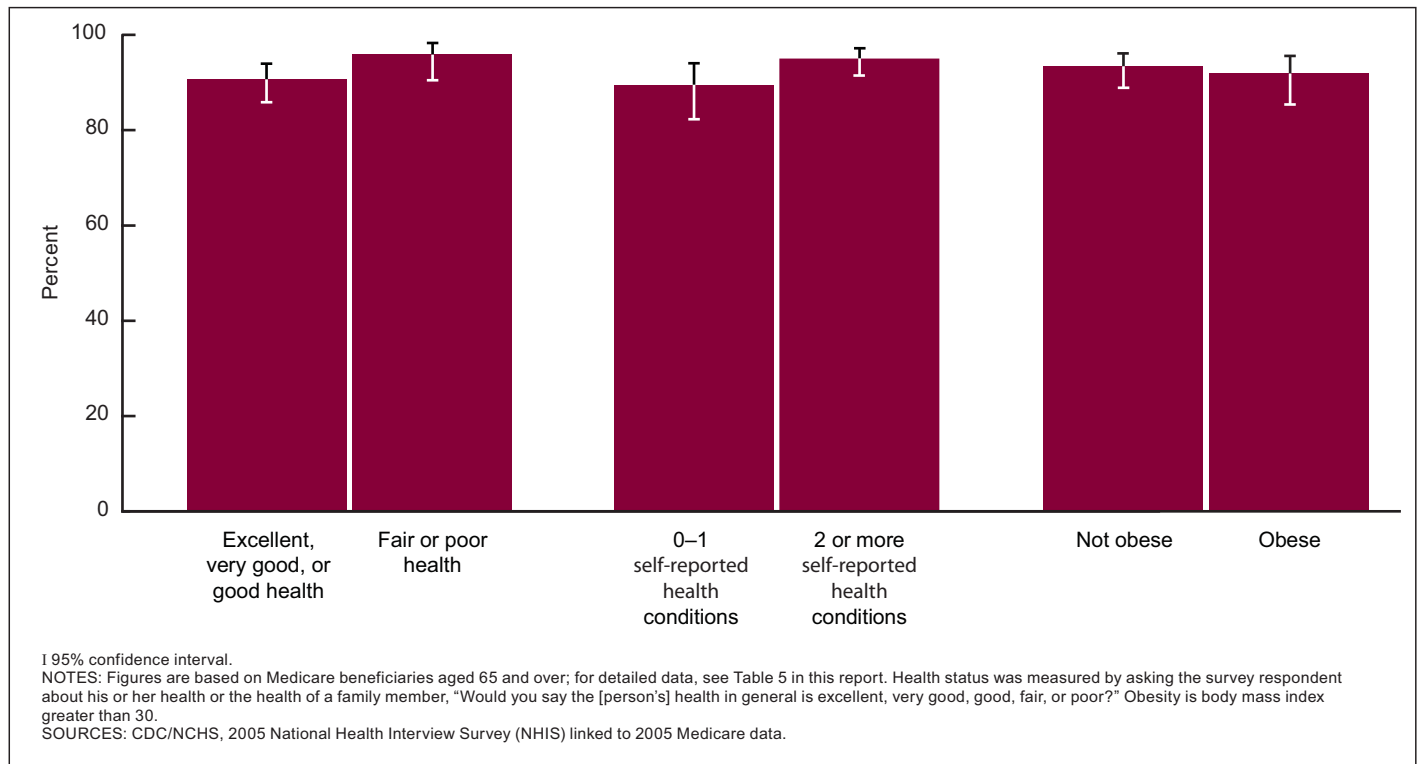


Figure 5. NHIS Medicare beneficiaries with diabetes in Chronic Conditions Summary file among those who self-reported diabetes, by health status

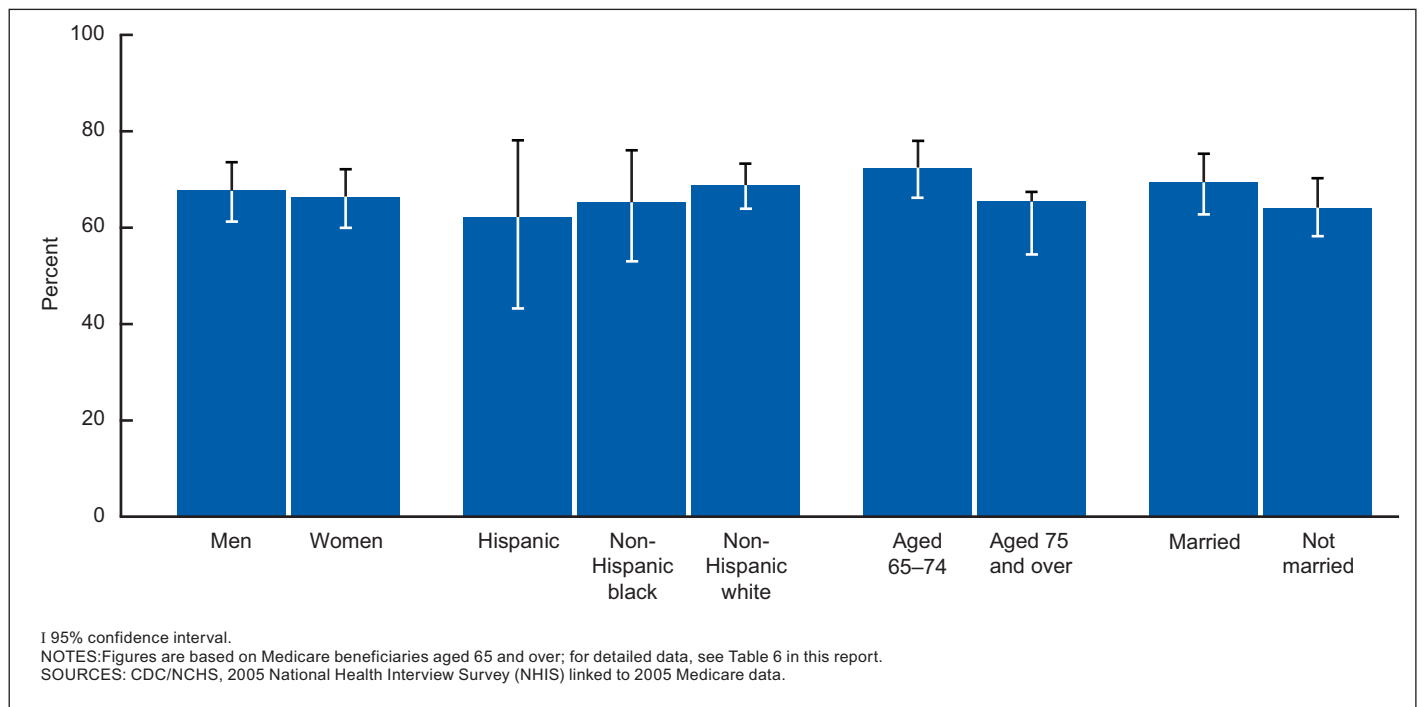


Figure 6. NHIS Medicare beneficiaries who self-reported diabetes, among those with diabetes in Chronic Conditions Summary file, by demographic characteristics

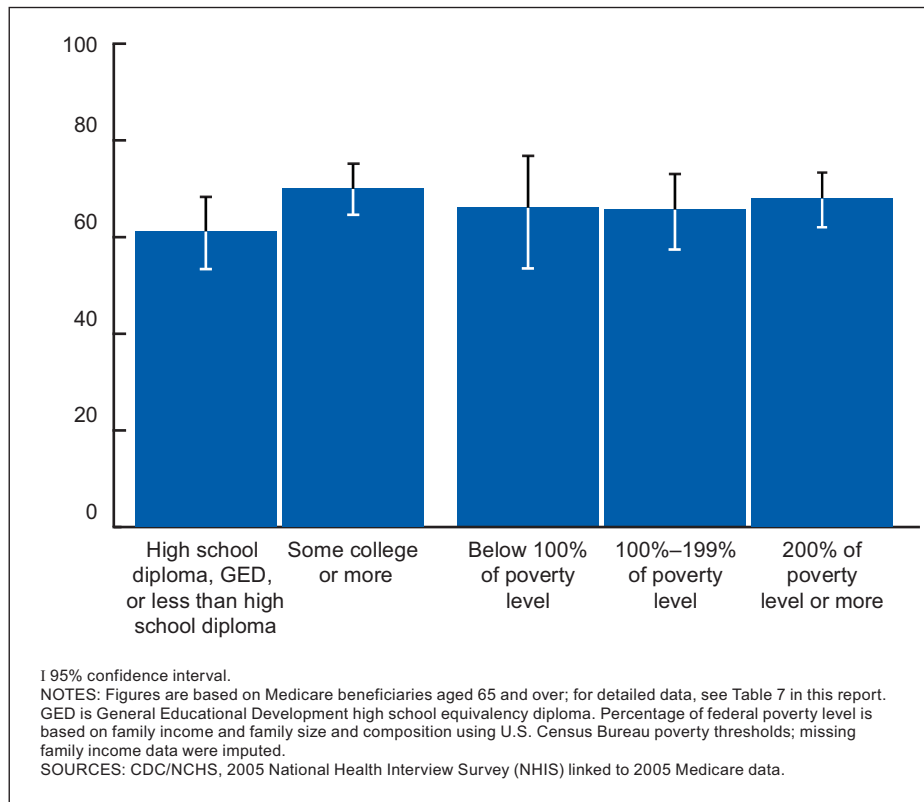


Figure 7. NHIS Medicare beneficiaries who self-reported diabetes, among those with diabetes in Chronic Conditions Summary file, by education and poverty level

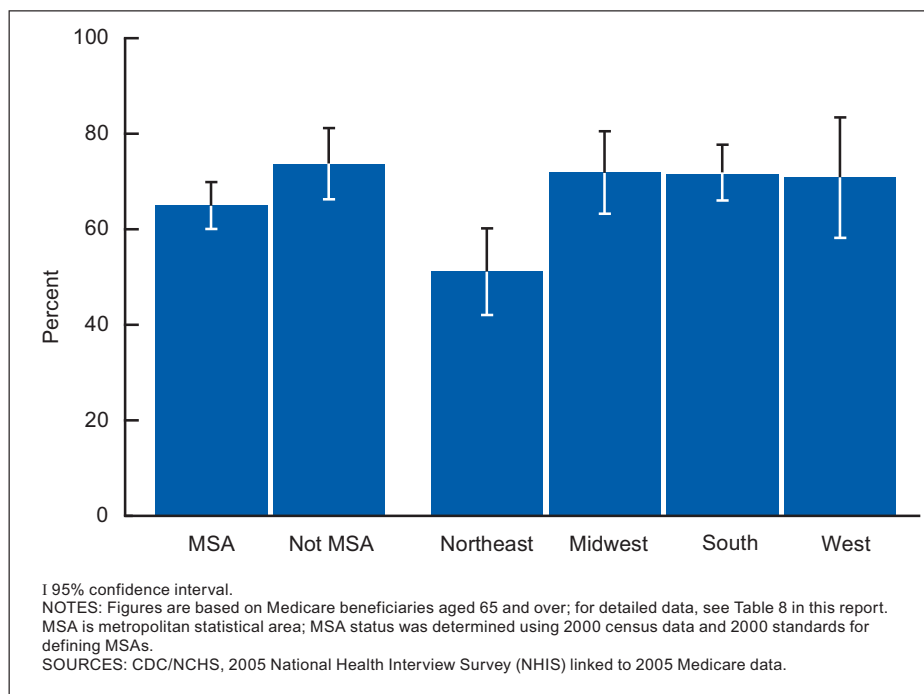


Figure 8. NHIS Medicare beneficiaries who self-reported diabetes, among those with diabetes in Chronic Conditions Summary file, by metropolitan status and region

report were at increased risk of diabetes. Beneficiaries who had a CC Summary indicator but did not self-report diabetes were more likely to be obese than beneficiaries who did not have a CC Summary indicator and did not self-report diabetes (32.0% compared with 17.5%, $p < 0.05$).

Discussion

Among Medicare beneficiaries aged 65 and over, the prevalence of diabetes was 27.8% using the CC Summary file. In comparison, the prevalence in this age group based on self-report from NHIS was 20.0%. Other research based on surveys report prevalence for this age group of 17%–22% (6,9,10). The Agency for Healthcare Research and Quality (AHRQ) estimated the prevalence of diabetes in the Medicare population in 2008 to be 28.0%, based on data from the Medicare Enrollment Database for beneficiaries aged 65 and over with at least 12 months of continuous enrollment in Medicare Parts A and B FFS (11); other recent Medicare estimates range from 24% to 28% for 2005–2008 (11,12). Other authors have found administrative database estimates of diabetes prevalence to be higher than self-reported estimates (7,8).

Overall, 93.0% of beneficiaries who self-reported diabetes had a CC Summary indicator for diabetes. This number is higher than the estimate reported from a recent study by Gorina and Kramarow (29) based on records from the NHANES I Epidemiologic Followup study (1971–1992) linked to Medicare claims from 1991 to 2000. The Gorina and Kramarow (29) study found that 68.6%–76.5% of previously diagnosed and self-reported (based on medical records and multiple years of self-report) diabetes cases were identified by Medicare claims (29).

However, the present report uses 2005 survey and administrative data, and changes in diagnostic criteria, increased awareness, and changes in

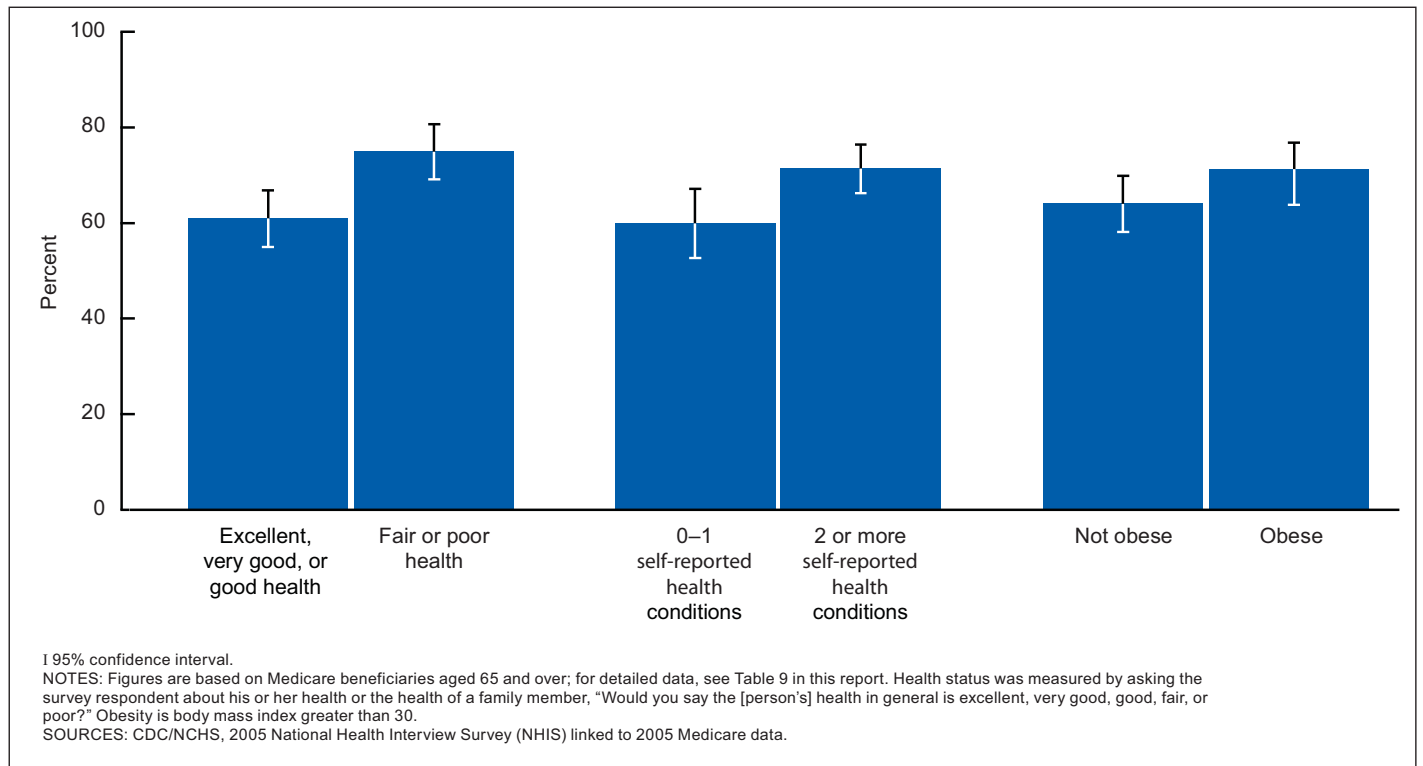


Figure 9. NHIS Medicare beneficiaries who self-reported diabetes, among those with diabetes in the Chronic Conditions Summary file, by health status

policies may have contributed to differences in results. As prevalence has increased and awareness has grown, CMS policies have expanded covered treatments for diabetes (30). As a result, diabetes treatments are more likely to generate Medicare claims and to be picked up by the CC Summary algorithm. Differences in the methods and populations studied between the current 2005 NHIS evaluation and the Gorina and Kramarow study also exist. The current study is based on a nationally representative cross-sectional sample of community-dwelling adults aged 65 and over who were enrolled in FFS in 2005. The Gorina and Kramarow study was a longitudinal study with a baseline of 1971–1975. That study sample included participants born in 1935 or earlier who survived to 1991, and did not exclude persons who had been institutionalized during the follow-up period.

In the current report, of those who self-reported diabetes in NHIS, more

women than men and a higher percentage of beneficiaries aged 75 and over were identified as having diabetes in the CC Summary file. Claims data are for a specific time period (2005 in this report), and diagnosis may have occurred at any point prior to the interview. Therefore, the presence of diagnosis in the claim is subject to utilization patterns. Women are more likely than men to use health care services (31). Research in adults younger than age 65 has found that women with diabetes have greater health care utilization than men (32), and women are more likely than men to use inpatient and physician services in their last years of life, although the relationship may be confounded by more women than men surviving to later years of life (33). Differences by age may be due to differences in health care use, because a beneficiary must utilize Medicare services to have claims count toward the CC Summary algorithm, by

definition. For example, based on the 2010 NHIS, 13.6% of adults aged 65–74 had a hospitalization, compared with 18.3% of those aged 75–84, and 20.8% of those aged 85 and over (34). When looking at certain subpopulations of survey respondents, especially men with self-reported diabetes or respondents younger than age 75 with self-reported diabetes, consider that these groups are less likely to have a CC Summary indicator for diabetes despite the fact that they have self-reported diabetes. Outcomes based on administrative data, such as cost estimates, also may underestimate the burden of disease in this subpopulation. This is especially true if different patterns of health care use are responsible for the lack of a CC Summary indicator for those who self-report diabetes.

Of the beneficiaries with an indicator for diabetes in the CC Summary file, only 67.0% self-reported diabetes at the interview. The self-report

percentage was higher for beneficiaries under age 75 with a CC Summary indicator, adults in regions other than the Northeast with a CC Summary indicator, adults in self-assessed poor health, and adults with two or more self-reported conditions with a CC Summary indicator. Previous studies have found that the presence of diabetes in medical records tends to be higher than self-report (7,9). This difference may be due to respondents misunderstanding survey questions or their doctors' diagnoses (8,35). In addition, the respondent may not be aware of his or her diagnosis at the time of the survey (35). Although a medical diagnosis of diabetes may have occurred after the survey date, only 2.7% of the beneficiaries in this study with an indicator for diabetes in the CC Summary file had a first recorded date for diabetes in 2005. Alternatively, physicians may code for diabetes when ordering tests for diabetes; however, this possibility could not be assessed (7). When looking at certain subpopulations of Medicare beneficiaries—such as respondents from the Northeast, those aged 65–74, and those with fewer self-reported conditions or better self-rated health—note that these populations were less likely to self-report diabetes, despite the fact that administrative data identifies them with diabetes. In studying these subpopulations, researchers should evaluate how potential overestimation could affect their analyses and consider using codes for conditions or treatments related to diabetes in addition to CC Summary identification.

Obesity is a risk factor for diabetes. In this study, beneficiaries who had a CC Summary indicator but did not self-report diabetes were more likely to also report a high BMI. The data do not show whether the CC Summary indicator is picking up at-risk individuals who are being tested for diabetes or high-risk individuals who either do not know they have diabetes or develop diabetes after the survey. However, differences by obesity indicate that the CC Summary indicator may be an accurate representation of high-risk

individuals for diabetes, even if those persons do not always report diabetes.

One potential limitation of this study is that CMS does not receive claims for MA enrollees (11). Although, historically, MA beneficiaries have been reported to be healthier and to have lower mortality than traditional FFS Medicare beneficiaries (15–17), recent studies have found differences between the two groups lessening (18,19). For example, a recent analysis of NHANES data linked to Medicare claims found no differences in diagnosed diabetes between FFS and MA enrollees (36).

Using the 2005 NHIS showed that beneficiaries who were linked to Medicare in 2005 but had no FFS claims for diabetes during 1999–2005 were slightly older than those who had FFS claims, but no difference was seen in the number of comorbidities. A sensitivity analysis examining only beneficiaries with 12 months of FFS in 2005 reduced the sample size from 2,109 to 1,955; however, the percentage of beneficiaries who self-reported diabetes among those with a CC Summary indicator, and the percentage of those with a CC Summary indicator among those who self-reported diabetes, remained similar to the results presented in this report. Another possible limitation is that NHIS contains only self-reported diagnosis of disease, and it has been shown in other surveys that as much as 46.2% of diabetes in the population aged 65 and over is undiagnosed (9). Other studies have also reported that diabetes is more often underreported than overreported (37).

The CC Summary indicator identified more cases of diabetes than self-report using the 2005 linked NHIS–CMS data. Most beneficiaries who self-reported diabetes in NHIS had a CC Summary indicator, but not as many people with a CC Summary indicator self-reported diabetes. Some participant characteristics were associated with the concordance between self-report and medical records. Those interested in utilizing the NHIS–Medicare linked files for research, or in examining factors associated with diabetes prevalence in national data sources, should be aware

of the differences between self-report in surveys and the Medicare CC Summary algorithm based on encounter data, and choose to use either self-report or CC Summary file to represent diabetes depending on their question of interest. Linked data can be used to compare different sources of information, or the information from the two data sources can be combined for a more complete picture of diabetes from both administrative and survey data. As more survey years and new linkages with Medicare become available, it will be possible to compare the two sources in greater detail.

References

1. Fradkin JE. Confronting the urgent challenge of diabetes: An overview. *Health Aff* 31(1):12–9. 2012.
2. Bertoni AG, Krop JS, Anderson GF, Brancati FL. Diabetes-related morbidity and mortality in a national sample of U.S. elders. *Diabetes Care* 25(3):471–5. 2002.
3. Sloan FA, Bethel MA, Ruiz D, Shea AM, Feinglos MN. The growing burden of diabetes mellitus in the U.S. elderly population. *Arch Intern Med* 168(2):192–9. 2008.
4. Thorpe KE, Ogden LL, Galactionova K. Chronic conditions account for rise in Medicare spending from 1987 to 2006. *Health Aff* 29(4):718–24. 2010.
5. American Diabetes Association. Economic costs of diabetes in the U.S. in 2012. *Diabetes Care* 36(4):1033–46. 2013.
6. Adler GS. Diabetes in the Medicare aged population, 2004. *Health Care Financ Rev* 29(3):69–79. 2008.
7. Hebert PL, Geiss LS, Tierney EF, Engelgau MM, Yawn BP, McBean AM. Identifying persons with diabetes using Medicare claims data. *Am J Med Qual* 14(6):270–7. 1999.
8. Edwards WS, Winn DM, Kurlantzick V, et al. Evaluation of National Health Interview Survey diagnostic reporting. *National Center for Health Statistics. Vital Health Stat* 2(120). 1994.
9. Cowie CC, Rust KF, Ford ES, Eberhardt MS, Byrd-Holt DD, Li C, et al. Full accounting of diabetes and pre-diabetes in the U.S. population in 1988–1994 and 2005–2006. *Diabetes Care* 32(2):287–94. 2009.

10. Barnes PM, Ward BW, Freeman G, Schiller JS. Early release of selected estimates based on data from the January–September 2011 National Health Interview Survey. National Center for Health Statistics. 2012.
11. Margolis DJ, Malay DS, Hoffstad OJ, Leonard CE, MaCurdy T, López de Nava K, et al. Prevalence of diabetes, diabetic foot ulcer, and lower extremity amputation among Medicare beneficiaries, 2006 to 2008. Diabetic Foot Ulcers, Data Points #1; AHRQ Publication No. 10 (11)–EHC009–EF. Rockville, MD: Agency for Healthcare Research and Quality. 2011.
12. Schneider KM, O'Donnell BE, Dean D. Prevalence of multiple chronic conditions in the United States' Medicare population. *Health Qual Life Outcomes* 7(1):82–92. 2009.
13. Riley GF. Administrative and claims records as sources of health care cost data. *Med Care* 47(7 Suppl 1):S51–5. 2009.
14. Medicare & Medicaid statistical supplement. 2011 ed, Table 2.1. Medicare & Medicaid Research Review. Available from: http://www.cms.gov/Medicare/MedicaidStatSupp/08_2011.asp.
15. Riley GF. Impact of continued biased disenrollment from the Medicare Advantage Program to fee-for-service. *Medicare & Medicaid Research Review* 2(4):E1–17. 2012.
16. Murgolo MS. Comparison of Medicare risk HMO and FFS enrollees. *Health Care Financ Rev* 24(1):177–85. 2002.
17. Nicholas LH. Who joins Medicare Managed Care? Voluntary enrollment and positive selection. Population Studies Center Research Report 09–670. Ann Arbor, MI: Population Studies Center. 2009.
18. McWilliams JM, Hsu J, Newhouse JP. New risk-adjustment system was associated with reduced favorable selection in Medicare Advantage. *Health Aff* 31(12):2630–40. 2012.
19. Newhouse JP, Price M, Huang J, McWilliams JM, Hsu J. Steps to reduce favorable risk selection in Medicare Advantage largely succeeded, boding well for health insurance exchanges. *Health Aff* 31(12):2618–28. 2012.
20. Ahluwalia SC, Gross CP, Chaudhry SI, Leo-Summers L, Van Ness PH, Fried TR. Change in comorbidity prevalence with advancing age among persons with heart failure. *J Gen Intern Med* 26(10):1145–51. 2011.
21. Madigan EA, Gordon N, Fortinsky RH, Koroukian SM, Piña I, Riggs JS. Predictors of functional capacity changes in a U.S. population of Medicare home health care (HHC) patients with heart failure (HF). *Arch Gerontol Geriatr* 54(3):e300–6. 2012.
22. Curtis JR, Arora T, Narongroeknawin P, Taylor A, Bingham CO, Cush J, et al. The delivery of evidence-based preventive care for older Americans with arthritis. *Arthritis Res Ther* 12(4):R144–52. 2010.
23. Simoni-Wastila L, Wei Y, Qian J, Zuckerman IH, Stuart B, Shaffer T, et al. Association of chronic obstructive pulmonary disease maintenance medication adherence with all-cause hospitalization and spending in a Medicare population. *Am J Geriatr Pharmacother* 10(3):201–10. 2012.
24. Schenker N, Raghunathan TE, Chiu P, Makuc DM, Zhang G, Cohen AJ. Multiple imputation of missing income data in the National Health Interview Survey. *J Am Stat Assoc* 101(475):924–33. 2006.
25. Ingram DD, Franco SJ. NCHS urban–rural classification scheme for counties. National Center for Health Statistics. *Vital Health Stat* 2(154). 2012.
26. National Center for Health Statistics, Office of Analysis and Epidemiology. Linkages between survey data from the National Center for Health Statistics and Medicare program data from the Centers for Medicare and Medicaid Services. 2012. Hyattsville, MD. Available from: http://www.cdc.gov/nchs/data/datalinkage/cms_medicare_methods_report_final.pdf. Accessed February 6, 2013.
27. Buccaneer Computer Systems & Service. Chronic condition data warehouse: Medicare administrative data user guide. 2011. Available from: http://www.ccwdata.org/cs/groups/public/documents/document/ccw_userguide.pdf [Accessed February 2013].
28. National Center for Health Statistics, Office of Analysis and Epidemiology. Use of survey weights for linked data files—Preliminary guidance. Hyattsville, MD. 2011. Available from: http://www.cdc.gov/nchs/data/datalinkage/use_of_survey_weights_for_linked_data_files.pdf [Accessed May 2013].
29. Gorina Y, Kramarow EA. Identifying chronic conditions in Medicare claims data: Evaluating the Chronic Condition Data Warehouse algorithm. *Health Serv Res* 46(5):1610–27. 2011.
30. Ashkenazy R, Abrahamson MJ. Medicare coverage for patients with diabetes. *J Gen Intern Med* 21(4):386–92. 2006.
31. Bertakis KD, Azari R, Helms LJ, Callahan EJ, Robbins JA. Gender differences in the utilization of health care services. *J Fam Pract* 49(2):147–52. 2000.
32. Shalev V, Chodick G, Heymann AD, Kokia E. Gender differences in healthcare utilization and medical indicators among patients with diabetes. *Public Health* 119(1):45–9. 2005.
33. Bird CE, Shugarman LR, Lynn J. Age and gender differences in health care utilization and spending for Medicare beneficiaries in their last years of life. *J Palliat Med* 5(5):705–12. 2002.
34. National Center for Health Statistics. Health, United States, 2011: With special feature on socioeconomic status and health. Table 102. Hyattsville, MD. 2012.
35. Miller DR, Rogers WH, Kazis LE, Spiro A, Ren XS, Haffer SC. Patients' self-report of diseases in the Medicare Health Outcomes Survey based on comparisons with linked survey and medical data from the Veterans Health Administration. *J Ambul Care Manage* 31(2):161–77. 2008.
36. Mirel LB, Wheatcroft G, Parker JD, Makuc DM. Health characteristics of Medicare traditional fee-for-service and Medicare Advantage enrollees: 1999–2004 National Health and Nutrition Examination Survey linked to 2007 Medicare data. National health statistics reports; no 53. Hyattsville, MD: National Center for Health Statistics. 2012.
37. Kriegsman DM, Penninx BW, van Eijk JT, Boeke AJ, Deeg DJ. Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly: A study on the accuracy of patients' self-reports and on determinants of inaccuracy. *J Clin Epidemiol* 49(12):1407–17. 1996.

Table 1. Percentage with diabetes among Medicare beneficiaries aged 65 and over at NHIS interview, by self-report or Chronic Condition Summary indicator

Category with diabetes	<i>n</i>	Weighted percent	CI
Self-reported diabetes in 2005 NHIS	420	20.0	18.0–22.2
2005 CC Summary indicator for those who self-reported diabetes in 2005 NHIS.	390	93.1	89.9–95.4
CC Summary indicator for diabetes in 2005 Medicare records	582	27.9	25.7–30.1
Self-reported diabetes in 2005 NHIS among those with a 2005 Medicare CC Summary indicator.	390	67.0	62.6–71.1

NOTES: Categories are not mutually exclusive. NHIS is National Health Interview Survey, and CC Summary is Medicare Chronic Condition Summary file. CI is 95% confidence interval.

SOURCE: CDC/NCHS, 2005 National Health Interview Survey linked to 2005 Medicare data.

Table 2. Unweighted sample sizes and weighted percentages for Medicare beneficiaries with diabetes in the CC Summary file, among beneficiaries aged 65 and over at NHIS interview who self-reported diabetes, by demographic characteristics

Characteristic	<i>n</i>	Percent	SE	CI
Sex ¹				
Men	168	87.4	2.6	81.3–91.7
Women	222	98.4	0.9	95.2–99.5
Age ¹ (years)				
65–74	212	90.5	2.2	85.2–94.0
75 and over.	178	96.8	1.3	92.8–98.6
Race and ethnicity				
Hispanic.	40	96.0	2.6	86.6–98.9
Non-Hispanic black	55	88.5	4.9	74.8–95.2
Non-Hispanic white	287	93.5	1.6	89.7–96.0
Marital status				
Widowed, divorced, or not married	234	94.2	1.5	90.3–96.6
Married or living with partner	154	92.3	2.2	86.7–95.6

¹Probability ($p < 0.05$) based on chi-squared statistic for association.

NOTES: CC is chronic condition, and NHIS is National Health Interview Survey. SE is standard error. CI is 95% confidence interval.

SOURCE: CDC/NCHS, 2005 National Health Interview Survey linked to 2005 Medicare data.

Table 3. Unweighted sample sizes and weighted percentages for Medicare beneficiaries with diabetes in the CC Summary file, among beneficiaries aged 65 and over at NHIS interview who self-reported diabetes, by education and poverty level

Characteristic	<i>n</i>	Percent	SE	CI
Education				
High school diploma, GED, or less than a high school diploma	142	94.6	2.1	88.8–97.5
Some college or more	246	92.3	1.8	88.0–95.2
Poverty level				
Below 100%	71	97.7	2.1	86.0–99.7
100%–199%	133	96.6	1.6	91.7–98.7
200% or more	186	90.3	2.2	85.0–93.9

NOTES: CC is chronic condition, and NHIS is National Health Interview Survey. SE is standard error. CI is 95% confidence interval. GED is General Educational Development high school equivalency diploma. Percentage of federal poverty level is based on family income and family size and composition using U.S. Census Bureau poverty thresholds; missing family income data were imputed.

SOURCE: CDC/NCHS, 2005 National Health Interview Survey linked to 2005 Medicare data.

Table 4. Unweighted sample sizes and weighted percentages for Medicare beneficiaries with diabetes in the CC Summary file, among beneficiaries aged 65 and over at NHIS interview who self-reported diabetes, by metropolitan residence and region

Characteristic	<i>n</i>	Percent	SE	CI
Location of residence				
MSA	283	92.3	1.7	88.1–95.1
Not MSA	103	95.6	2.0	89.6–98.2
Region				
Northeast	76	94.7	2.8	85.8–98.1
Midwest	86	90.9	2.7	83.9–95.0
South	179	93.3	2.3	87.2–96.6
West	49	95.0	2.5	87.1–98.2

NOTES: CC is chronic condition, and NHIS is National Health Interview Survey. SE is standard error. CI is 95% confidence interval. MSA is metropolitan statistical area; MSA status was determined using 2000 census data and 2000 standards for defining MSAs.

SOURCE: CDC/NCHS, 2005 National Health Interview Survey linked to 2005 Medicare data.

Table 5. Unweighted sample sizes and weighted percentages for Medicare beneficiaries with diabetes in the CC Summary file, among beneficiaries aged 65 and over at NHIS interview who self-reported diabetes, by health status

Characteristic	<i>n</i>	Percent	SE	CI
Self-rated health				
Good, very good, or excellent	204	90.7	2.0	85.9–94.0
Fair or poor	186	95.9	1.8	90.5–98.3
Number of self-reported health conditions				
0–1	146	89.5	2.9	82.3–94.0
2 or more	244	95.1	1.4	91.5–97.2
Obesity				
Not obese	217	93.4	1.8	89.0–96.1
Obese	152	91.9	2.5	85.4–95.6

NOTES: CC is chronic condition, and NHIS is National Health Interview Survey. SE is standard error. CI is 95% confidence interval. Health status was measured by asking the survey respondent about his or her health or the health of a family member, "Would you say the [person's] health in general is excellent, very good, good, fair, or poor?" Obesity is body mass index greater than 30.

SOURCE: CDC/NCHS, 2005 National Health Interview Survey linked to 2005 Medicare data.

Table 6. Unweighted sample sizes and weighted percentages for Medicare beneficiaries who self-reported diabetes, among beneficiaries aged 65 and over at NHIS interview with a CC Summary indicator for diabetes, by demographic characteristics

Characteristic	<i>n</i>	Percent	SE	CI
Sex				
Men	168	67.7	3.1	61.3–73.5
Women	222	66.4	3.0	60.1–72.1
Age ¹ (years)				
65–74	212	72.4	2.9	66.3–77.8
75 and over	178	61.1	3.3	54.5–67.4
Race				
Hispanic	40	62.2	9.2	43.2–78.1
Non-Hispanic black	55	65.4	6.0	53.0–76.0
Non-Hispanic white	287	68.8	2.4	64.0–73.2
Marital status				
Widowed, divorced, or not married	234	64.0	3.0	57.9–69.8
Married or living with partner	154	69.4	3.2	62.8–75.3

¹Probability ($p < 0.05$) based on chi-squared statistic for association.

NOTES: NHIS is National Health Interview Survey, and CC is chronic condition. SE is standard error. CI is 95% confidence interval.

SOURCE: CDC/NCHS, 2005 National Health Interview Survey linked to 2005 Medicare data.

Table 7. Unweighted sample sizes and weighted percentages for beneficiaries who self-reported diabetes, among Medicare beneficiaries aged 65 and over at NHIS interview with a CC Summary indicator for diabetes, by demographic characteristics

Characteristic	<i>n</i>	Percent	SE	CI
Education				
High school diploma, GED, or less than a high school diploma	142	61.1	3.8	53.4–68.3
Some college or more	246	70.0	2.7	64.6–75.1
Poverty level				
Below 100%	81	66.1	6.0	53.5–76.8
100%–200%	139	65.6	4.0	57.4–73.0
200% or more	194	67.9	2.9	62.0–73.3

NOTES: NHIS is National Health Interview Survey, and CC is chronic condition. SE is standard error. CI is 95% confidence interval. GED is General Educational Development high school equivalency diploma. Percentage of federal poverty level is based on family income and family size and composition using U.S. Census Bureau poverty thresholds; missing family income data were imputed.

SOURCE: CDC/NCHS, 2005 National Health Interview Survey linked to 2005 Medicare data.

Table 8. Unweighted sample sizes and weighted percentages for those who self-reported diabetes, among Medicare beneficiaries aged 65 and over at NHIS interview with a CC Summary indicator for diabetes, by metropolitan residence and region

Characteristic	<i>n</i>	Percent	SE	CI
Location of residence				
MSA	283	64.9	2.5	59.8–69.6
Not MSA	103	73.7	3.8	65.5–80.5
Region ¹				
Northeast	76	51.1	4.7	42.0–60.1
Midwest	86	71.9	4.4	62.5–79.7
South	179	71.5	3.0	65.3–77.0
West	49	70.8	6.5	56.6–81.9

¹Probability ($p < 0.05$) based on chi-squared statistic for association. The percentage for the Northeast was significantly different from the percentages for all other regions ($p < 0.01$). There were no significant differences between other regions.

NOTES: NHIS is National Health Interview Survey, and CC is chronic condition. SE is standard error. CI is 95% confidence interval. MSA is metropolitan statistical area; MSA status was determined using 2000 census data and 2000 standards for defining MSAs.

SOURCE: CDC/NCHS, 2005 National Health Interview Survey linked to 2005 Medicare data.

Table 9. Unweighted sample sizes and weighted percentages for those who self-reported diabetes, among Medicare beneficiaries aged 65 and over at NHIS interview with a CC Summary indicator for diabetes, by health status

Characteristic	<i>n</i>	Percent	SE	CI
Self-rated health ¹				
Good, very good, or excellent	204	60.9	3.0	54.8–66.6
Fair or poor	186	74.9	2.9	68.7–80.2
Number of self-reported health conditions ¹				
0–1	146	59.9	3.7	52.5–66.9
2 or more	244	71.2	2.6	65.8–76.0
Obesity				
Not obese	217	64.0	2.9	58.1–69.6
Obese	152	71.3	3.3	64.3–77.3

¹Probability ($p < 0.05$).

NOTES: NHIS is National Health Interview Survey, and CC is chronic condition. SE is standard error. CI is 95% confidence interval. Health status was measured by asking the survey respondent about his or her health or the health of a family member, "Would you say the [person's] health in general is excellent, very good, good, fair, or poor?" Obesity is body mass index greater than 30.

SOURCE: CDC/NCHS, 2005 National Health Interview Survey linked to 2005 Medicare data.

**U.S. DEPARTMENT OF
HEALTH & HUMAN SERVICES**

Centers for Disease Control and Prevention
National Center for Health Statistics
3311 Toledo Road, Room 5419
Hyattsville, MD 20782

FIRST CLASS MAIL
POSTAGE & FEES PAID
CDC/NCHS
PERMIT NO. G-284

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, \$300

National Health Statistics Reports ■ Number 69 ■ November 1, 2013

Suggested citation

Day HR, Parker JD. Self-report of diabetes and claims-based identification of diabetes among Medicare beneficiaries. National health statistics reports; no 69. Hyattsville, MD: National Center for Health Statistics. 2013.

Copyright information

All material appearing in this report is in the public domain and may be reproduced or copied without permission; citation as to source, however, is appreciated.

National Center for Health Statistics

Charles J. Rothwell, M.S., *Acting Director*
Jennifer H. Madans, Ph.D., *Associate Director
for Science*

Office of Analysis and Epidemiology

Irma E. Arispe, Ph.D., *Director*

For e-mail updates on NCHS publication releases, subscribe online at: <http://www.cdc.gov/nchs/govdelivery.htm>.
For questions or general information about NCHS: Tel: 1-800-CDC-INFO (1-800-232-4636) • TTY: 1-888-232-6348
Internet: <http://www.cdc.gov/nchs> • Online request form: <http://www.cdc.gov/cdc-info/requestform.html>

DHHS Publication No. 2014-1250 • CS241740