

3. Cost Analysis Data and Methods

To estimate the annual cost burden associated with each of the six selected conditions in 2000 among individuals 65 and older, we separately estimated direct and indirect costs for each illness. In this section, we describe the data and methods used to construct these cost estimates.

3.1 Direct Costs

To make best use of the available national data, we separated the estimation of direct costs into three mutually exclusive components:

- medical costs,
- self-administered prescription drug costs, and
- nursing home costs.

Each of these components is discussed below.

3.1.1 Medical Costs

For individuals 65 and older, Medicare claims provide the most comprehensive source of data on costs for medical services. National samples of claims data are accessible in different ways. After reviewing the alternatives, we selected the Consumer Assessment of Health Plans Surveys (CAHPS) for Medicare Fee-for-Service (MFFS) beneficiaries as the primary data source for our analysis.

3.1.1.1 Data Description

CAHPS-MFFS was selected for this part of the analysis because it was relatively easy to access and is well suited for this analysis. These data include detailed Medicare claims data and other potentially relevant demographic and health-related information for a large and nationally representative sample of older Americans.

The sampling frame for CAHPS-MFFS was drawn from CMS's Enrollment Data Base (EDB), which comprises approximately 30 million persons enrolled in fee-for-service Medicare and residing in the United States or Puerto Rico. Beneficiaries with less than 6 months' continuous enrollment in MFFS and those with a representative payee or living in an institution were excluded. To select the CAHPS-MFFS sample, a total of 276 geographic primary sampling areas were constructed, each consisting of one or more counties, and the sample was allocated to these areas to achieve a minimum of 300 respondents in each area.

The CAHPS-MFFS sample for this study included 145,875 noninstitutionalized Medicare beneficiaries, age 65 years or older, who were enrolled in MFFS and eligible for Part A and Part B benefits for at least part of 2000. Beneficiaries with proxy respondents were excluded from the analysis. It should be noted that, because the study was based on a retrospective analysis of self-reported survey data, the study sample has a disproportionate share of nondecedents. The disproportionate share of nondecedents in the sample may slightly bias the cost estimates from our analysis. However, because health care costs tend to escalate during the last few months of a person's life and drop to zero thereafter, the overall direction of this bias is not known.

3.1.1.2 Analysis Methods

To estimate medical costs associated with the selected illness categories, we began by extracting Medicare claims information for the year 2000 for all Medicare-covered medical services and for all survey respondents and nonrespondents. Medicare-covered services include

- acute and nonacute inpatient services,
- hospital outpatient and other ambulatory services,
- professional fees,
- home health services, and
- durable medical equipment.

Medicare does not cover self-injected medications administered in a home or outpatient setting and nursing home or other long-term care services; therefore, costs for these services are not included in this claims-based analysis.

Costs reported in the Medicare claims data include

- payments made by Medicare,
- beneficiary payments in the form of deductibles and copayments, and
- any third-party payments.

Total medical payments are calculated as the sum of the three payments types.

We then identified medical costs associated with each of the illness categories by matching the ICD-9 codes recorded for each claim with the condition-specific ICD-9 codes listed in Table 2-1. For inpatient services, each claim consolidates payments at an admissions level. Consequently, each of these claims may be associated with multiple ICD-9 codes—one for the primary diagnosis and potentially several more codes for related secondary diagnoses. Similarly, claims for outpatient and home health services report payments that may be associated with multiple ICD-9 codes. In contrast, claims for professional fees and durable medical equipment are all condition specific and are therefore associated with a single ICD-9 code.

For this analysis, we selected all claims that were recorded with *at least one* of the ICD-9 codes listed in Table 2-1, including both primary and secondary diagnostic codes for inpatient, outpatient, and home health services. Because payments for these types of services are consolidated, they may include costs that are attributable to conditions other than the selected ICD-9 codes. As a result, for inpatient, outpatient, and home health services, the cost estimates developed in this analysis are likely to overstate the costs that are specifically attributable to the six conditions of interest. The extent of this overestimation is not known. For professional fees and durable medical equipment, no similar overestimation of costs is expected because they are all based on condition-specific claims.

Using the CAHPS-MFFS (with appropriate geographic weights for the stratified sample), we were therefore able to obtain unbiased nationally representative estimates for the roughly 26 million noninstitutionalized Medicare beneficiaries over the age of 64 enrolled in fee-for-service in 2000. Based on the sample of 145,875 respondents, we are able to estimate the number and percentage (prevalence rate) of beneficiaries, 65 and older, who were diagnosed with conditions in each of the six illness categories in 2000. It should be noted these prevalence estimates are based on claims for services in 2000; therefore, they do not include individuals who had the underlying condition in 2000 but did not for whatever reason seek care for the condition during that period. Consequently, these claims based prevalence estimates are likely to underestimate by a small amount the overall prevalence of each condition in the noninstitutionalized fee-for-service elderly population.

For these beneficiaries, we are also able to estimate the average (per capita) and aggregate medical costs associated with these illnesses in 2000. With the information in CAHPS-MFFS, it is also possible to disaggregate these estimates by

- type of medical service,
- age/gender categories, and
- Census region and state.

The results for each illness category are summarized in Section 4, with additional detail provided in Appendix C.

To extrapolate these direct medical cost estimates to the entire non-institutionalized population of individuals 65 and older in 2000, we assumed that the prevalence and average costs for each condition based on the CAHPS-FFS data were also applicable to those not participating in fee for service Medicare and not in nursing homes. According to U.S. Census data, the total 65 and older population in 2000 was 35 million, and according to the National Center of Health Statistics (NCHS, 2003) the number of nursing home residents was 1.5 million. Therefore, we applied the prevalence and average cost estimates for each condition to an assumed total population of 33.5 million. This extrapolation is like to slightly overestimate prevalence and costs for those not in fee-for service Medicare (and not institutionalized) because they are on average somewhat younger and in somewhat better health than those in Medicare fee for service (Brown et al., 1993).

3.1.2 Self-Administered Prescription Drug Costs

One of the inherent limitations in using Medicare claims data (including CAHPS-MFFS as discussed in Section 3.1.1) to assess direct costs of illness is that expenditures for self-administered prescription drugs are not reimbursable by Medicare and are therefore excluded from the claims-based cost estimates. To estimate these prescription drug costs for each of the six conditions, we used data from the 2000 Medical Expenditures Panel Survey (MEPS).

3.1.2.1 Data Description

The MEPS is a nationally representative subsample of the National Health Interview Survey (NHIS, described in some detail below in Section 3.2). One important feature of the MEPS is that it allows for the linkage of demographic data on individuals to their health care utilization, spending, and sources of payment.

The 2000 survey requested information about the number of prescription medication purchases (including refills) and total expenditures for prescription drugs by sources of payment. The sources of payment for which information was collected include self or family, Medicare, Medicaid, private insurance, TRICARE, the Veterans' Administration, workers' compensation, and other sources. The survey also contains a conditions file that includes information about all conditions reported for household residents.

Health conditions included in MEPS are those reported by a household respondent and those listed as the reason for a medical provider visit or prescribed medication. Medical conditions were recorded by the interviewer as verbatim text; professional coders later assigned three-digit ICD-9 codes.

3.1.2.2 Analysis Methods

To estimate the prescription drug costs associated with each of the six conditions, we first created a sample of individuals 65 years and older from the 2000 MEPS and then calculated mean prescription drug expenditures for each condition by age group (65 to 74, 75 to 84, and 85 years and older) and by sex. Mean expenditures and standard errors were estimated for total prescription drug spending, for Medicare spending, for total spending except Medicare, and for total spending except Medicare and self- or family-paid costs. To avoid double-counting, our reported estimates focus on the cost estimates that exclude prescription drugs where Medicare was reported as the source of payment. However, cost estimates from

our analysis of Medicare claims may also include co-pays or deductibles that represent out-of-pocket expenses. Consequently, our prescription drug cost estimates may include some out-of-pocket expenses that are also included in our cost estimates from Medicare claims analysis.

Because estimates of mean total prescription drug costs cannot be taken as attributable cost estimates, we also estimated the incremental, or additional, costs of prescription drugs for individuals with each condition as compared to persons without the condition, by age group and by sex. Incremental cost estimates represent the additional costs for those with the condition as compared to those without the condition. We chose to consider incremental costs because it allows us to interpret our estimates as representing excess costs for people with the conditions of interest. However, because it is not uncommon for individuals in the 65 and older age group to have high medical care expenses, incremental cost estimates in an older population are often close to zero or even negative. In other words, individuals in this age group who do not have the condition of interest have a fairly high likelihood of having another costly condition or disability. In comparison, total cost estimates are limited in that they represent total costs for individuals with a condition and do not account for the high degree of comorbidity in the over-65 population.

3.1.3 Nursing Home Costs

Another gap in our cost estimates from Medicare claims data (using CAHPS-MFFS) is the cost associated with care in a nursing home. We address this gap by analyzing nursing home charges using the 1999 National Nursing Home Survey (NNHS).

3.1.3.1 Data Description

The 1999 NNHS sample consisted of 1,496 of the approximately 18,400 nursing homes in the United States that had at least three beds and were either certified by Medicare or Medicaid or had a state license to operate as a nursing home. A sample of up to six current residents was selected from each participating nursing home for inclusion in the current resident file. Interviews were then conducted with nursing staff familiar with residents' medical records to collect demographic information, diagnoses (ICD-9 codes assigned at admission and based on current records), total charges for a specified period, and indicators for a range of different sources of payment (e.g., self pay, Medicare, Medicaid). The responding nurse referred to the resident's medical record when answering the interview questions.

3.1.3.2 Analysis Methods

To develop an estimate of the nursing home costs associated with each of the six conditions, we used the NNHS current resident file and limited our analysis to those people 65 years and older. We then calculated mean total charges by sex, age group (65 to 74, 75 to 84, and 85 years and older), and condition. These estimated charges were converted to costs by using the Medicare cost-to-charge ratio for skilled nursing facilities, which was 0.78777. Estimated costs were updated to 2000 dollars using the consumer price index for all urban consumers.

Incremental cost estimates were then generated for each condition by taking the difference of mean total costs by age group and sex for those with and without the condition of interest. We also examined frequencies of Medicare as a source of payment and the inclusion of drugs and medical supplies in the total charges reported.

Applying estimates of *mean* incremental costs for each health condition to calculate *aggregate* nursing home costs also requires estimates of the number of elderly in nursing homes with each condition. Data on nursing home trends (NCHS, 2003) indicates that there were approximately 1.5 million nursing home residents in 2000. To estimate the portion of this population that experienced each condition, we applied the prevalence rates derived from the CAPHS-MFFS data. These claims based prevalence estimates are likely to somewhat underestimate prevalence among nursing home residents because rates

of illness are typically higher in the institutionalized population and because the claims based estimates do not include those who did not seek care for their condition in 2000.

3.2 Indirect Costs

Indirect costs are those for which no payment is made, but for which an economic effect is still observed. In our analysis and in most COI studies, a human capital approach is used to value the indirect cost of time lost from productive activities due to excess morbidity or premature mortality. The human capital approach values productivity losses based on market earnings and an imputed value for household production (Drummond et al., 1997). Advantages of the human capital approach are that it assumes a societal perspective and relies on data that are readily available. A disadvantage is that estimates of the value of life and health among the retired elderly will be lower than comparable estimates for working-age adults because the estimates are based largely on measures of market productivity. The human capital approach also ignores the costs of pain and suffering associated with an illness and excludes the value of time lost from consumption activities (Tolley, Kenkel, and Fabian, 1994).

The concept of willingness to pay (WTP) is useful for addressing the limitations of the human capital approach. However, the implementation of stated or revealed preference approaches to estimate the value of health effects, or WTP, requires far more time and resources than does a human capital approach using existing data. For this reason, we used the human capital approach to value indirect costs associated with the six conditions. These estimates may be thought of as lower-bound estimates for the full indirect costs of each condition.

3.2.1 Increased Morbidity

We used data from the 2001 National Health Interview Survey (NHIS) to estimate the impact of each condition on labor force participation rates, number of work loss days for those who were working during the survey year, and number of bed days. These estimates were combined with age- and sex-specific earnings and household productivity values for 2000 to generate estimates of the increased morbidity costs associated with each condition.

3.2.1.1 Data Description

The NHIS is an annual household interview survey designed primarily to collect data about the health status, health conditions, and health care utilization of household members. The survey also requests information about basic demographics and about days lost from work, bed days, and functional limitations. Data are collected annually from about 43,000 households and about 106,000 persons within the households. Information about diseases and other health conditions among people was also collected. However, the NHIS did not contain information about pneumonia or gastrointestinal illness; as a result, these conditions are excluded from our morbidity cost estimates.

Data from the NHIS were combined with productivity estimates from Grosse (2003). Grosse provides estimates of annual earnings for those in the workforce and household productivity for all adults by age group (65 to 74 and 75 and older) and by sex.

3.2.1.2 Analysis Methods

We used the 2001 NHIS to estimate two components of labor productivity losses: the losses associated with being completely unable to work (out of the labor force) and the losses associated with missing work days. For the first component, we used estimates of the percentage of those who report that they are not currently working but worked previously as a proxy for the probability of being unable to work due to poor health. We estimated the probability of being completely unable to work because of one of the six health conditions as the excess percentage of those with the condition as compared to those

without the condition who are not currently working. Separate estimates of the probability of being unable to work were developed for all those 65 years and older for each condition.

We then estimated the per-person morbidity cost associated with being completely unable to work as the estimated probability of being unable to work because of the condition multiplied by annual earnings. Annual labor market earnings estimates were by sex for the 65 to 74 and 75 years and older age groups.

For the second component of labor productivity losses, we limited our analysis to those over 65 years of age who reported that they currently work. We then used a regression approach to estimate the impact of each condition on missed work days, controlling for a number of variables, including education, poverty status, self-reported health status, occupation, and smoking status.

Results from our regression analysis are shown in Appendix B. Using the regression results, we generated predictions of the number of missed work days attributable to each of the six conditions. These estimates were generated by condition and by sex for all age groups combined. We then estimated labor productivity losses by multiplying predicted work-loss days by sex-specific average daily earnings for those 65 to 74 years and 75 years and older in the workforce (Grosse, 2003). Per-person estimates for all those with the condition were generated by multiplying estimated productivity losses due to missed work days by the percentage with each condition that currently works. Our methods for estimating the household productivity losses associated with being sick in bed were similar.

Our bed days analysis included all those 65 years and older. We used a regression approach to estimate the impact of each condition on number of bed days, controlling for education, poverty status, self-reported health status, labor force participation status, and smoking status. Based on these results, we predicted the number of bed days attributable to each condition by sex. Household productivity losses were estimated by multiplying predicted bed days for each condition by sex-specific average household productivity for those 65 to 74 years and 75 years and older.

The estimates of labor productivity losses shown in Section 4 represent the sum of the expected costs due to being unable to work and the expected costs due to missing work for people with each condition. Household productivity losses represent the value of lost household work attributable to each condition.

Our estimates for lung cancer are limited by the small number of observations in the 65 years and older age group in the 2001 NHIS. Only 26 males and 15 females were identified with lung cancer. The number with lung cancer who reported being in the labor force was even smaller and forced us to exclude lung cancer from our work days-loss analysis. We also excluded lung cancer from our analysis of bed days attributable to lung cancer. The only morbidity costs for lung cancer included in our analysis are the costs associated with being completely unable to work.

We do not provide measures of dispersion for our indirect cost estimates, in large part because our source for the earnings and household productivity estimates did not provide standard errors or other measures of dispersion. Because the variation in earnings among older individuals who work is much greater than the variation in the percentage of older Americans who work, we chose not to estimate standard errors that would necessarily treat earnings as having no variation.

3.2.2 Premature Mortality

We used data from the 1998 National Vital Statistics Report (NVSR) to determine the number of annual deaths with one of the six conditions listed as the cause of death. These data were combined with estimates of the present value of earnings and household productivity from Grosse (2003) to estimate the cost of mortality due to each condition.

3.2.2.2 Data Description

The 1998 NVSR provides calculations from all death certificates filed in the 50 states and the District of Columbia in 1996 (Peters, Kockanek, and Murphy, 1998). We relied on the 1998 report primarily because it uses ICD-9 codes to categorize causes of death. Later reports began using the revised ICD-10 classification scheme. The NVSR provides statistics on deaths and death rates by age group (65 to 74, 75 to 84, and 85 years and older) for groupings of ICD-9 codes that closely match the disease groupings we selected for COI analysis. It also provides death rates for the overall population by age group and sex.

Grosse (2003) provides estimates of the present value of earnings and household productivity by age for several different discount rates. In our analysis, we use the Grosse (2003) present value estimates for earnings only and for earnings and household production combined. For both measures, we assume that the value of future production is discounted at an annual rate of 5 percent.

3.2.2.3 Analysis Methods

To estimate the indirect costs resulting from premature mortality among those 65 years and older, we multiplied the cause- and age group-specific number of deaths for each condition by the age group-specific present value of earnings (earnings and household production). Because Grosse (2003) does not provide an estimate for expected productivity among those 85 and older, we were not able to estimate losses (labor or household) for deaths in the 85 and older age group. The present value of productivity losses in this age group is expected to be lower than for the younger age groups.

For ischemic heart disease and pneumonia, the health condition categories provided in the NVSR exactly matched our characterization of the condition using ICD-9 codes. However, for chronic lung disease (CLD), stroke, and lung cancer, the NVSR condition categories included some ICD-9 codes that were not used in our assessment of direct costs. For gastrointestinal illness, the NVSR condition categories excluded some of the ICD-9 codes used in our direct cost analyses. Because of these discrepancies, our estimates of mortality costs for CLD, stroke, and lung cancer are slight overestimates of mortality costs, while our estimate for gastrointestinal illness is an underestimate of actual mortality costs for the conditions of interest.