Age and Sex Estimates in the National Health Expenditure Accounts: Definitions, Sources, and Methods, 2020

Introduction

Health spending is estimated by age and sex and presented by type of service (and health care good) and by source of funds for males and females in various age groups for the even years from 2002 - 2020. These estimates are linked to the National Health Expenditure Accounts and use the same categories and definitions.[1] Thus, the age and sex estimates are based on data collected on an establishment basis, grouping services together according to place of service, rather than according to type of service. For example, hospital-based nursing homes are shown in our hospital category, while freestanding nursing homes are shown in the nursing home category.

Age and sex estimates are shown for personal health care (PHC) expenditures, which include spending for hospital care, physician and clinical services, dental care, other professional services, home health care, nursing care facilities and continuing care retirement communities, other health residential and personal care, and retail sales of medical products (such as prescription drugs or over-the-counter medicines sold in pharmacies or eyeglasses sold in optical goods stores). Age and sex estimates are not available for total national health expenditures (NHE) because data is not available to break out the non-PHC categories including government public health programs, government administration and the net cost of private health insurance, and investment.

We disaggregate PHC by sex into the following five age categories: 0-18, 19-44, 45-64, 65-84, and 85 and over. Analysis is also provided for three broader age groupings: children (age 0-18), working-age adults (age 19-64), and the elderly (age 65 and over). We produced estimates of health spending by age and sex for selected years including 2002, 2004, 2006, 2008, 2010, 2012, 2014, 2016, 2018, and 2020.[2]

Data Sources

Since no single source of comprehensive health spending by age and sex exists, we used several sources and methods to develop these estimates. The table below lists the data sources that we used to create these estimates.

[1] For a complete methodology of the historical national health expenditure accounts, http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html

[2] Previous estimates were also published for 1987, 1996, 1999, 2002, and 2004 but were only by age-group with the exception of 2004 where the estimates are also available by sex. However, the previous estimates are not directly comparable with the current estimates, as they have not been controlled to the recent historical national health expenditure estimates and do not reflect the statistical and methodological improvements that are included in the later estimates.

Data Source	Years Used	Reference
U.S. Census Bureau, Population Division: National Population by Characteristics: 2010-2020: Annual Estimates of the Resident Population by Single Year of Age and Sex for the United States: April 1, 2010 to July 1, 2020	2010-2020	https://www.census.gov/programs-surveys/popest/technical-documentation/research/evaluation-estimates/2020-evaluation-estimates/2010s-national-detail.html
Medicaid Analytic eXtract (MAX) data	2002-2014	http://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAXGeneralInformation.html
Medical Expenditure Panel Survey (MEPS)	2000-2020	http://meps.ahrq.gov/
Medicare Current Beneficiary Survey (MCBS)	2002-2020	https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/MCBS-Public-Use-File/index.html
Monthly Membership Report (MMR)	2007-2014	http://www.cms.gov/Research-Statistics-Data-and-Systems/CMS-Information-Technology/mapdhelpdesk/downloads/PCUG_v5_3_111710_Appendices_With_Cover_Final.pdf
National Ambulatory Medical Care Survey (NAMCS)	2002-2014	http://www.cdc.gov/nchs/ahcd.htm
National Claims History Files (NCH)	2002-2020	http://www.cms.gov/FilesForOrderGenInfo/
National Hospital Ambulatory Medical Care Survey (NHAMCS)	2001-2014	http://www.cdc.gov/nchs/ahcd.htm
National Hospital Discharge Survey (NHDS)	2001-2010	http://www.cdc.gov/nchs/nhds.htm
National Intercensal Estimates	2002-2020	https://www.census.gov/programs-surveys/popest/data/data-sets.All.html
Prescription Drug Event File	2006-2020	http://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovGenIn/downloads/GuidePartDDataRequests.pdf
Transformed Medicaid Statistical Information System (T-MSIS)	2014-2020	https://www.medicaid.gov/medicaid/data-systems/macbis/transformed-medicaid-statistical-information-system-t-msis/index.html
Merative - Marketscan Commercial Claims Databases	2002-2020	https://www.merative.com/real-world-evidence

General Methods

For Medicare, Medicaid, and the Children's Health Insurance Program (CHIP) we use CMS administrative data to develop health spending estimates by age and sex. The Medicare estimates are based on data from the National Claims History Files, Monthly Membership Report, and Prescription Drug Event Files while the Medicaid and CHIP estimates are based on data from the Transformed Medicaid Statistical Information System (T-MSIS) and the Medicaid Analytic eXtract System (MAX).

We develop all other age/sex-based spending estimates for the remaining payers of PHC goods and services using one of two methods. The first method derives spending by age and sex by multiplying (i) cost per use data for the non-institutional population from the Office of the Actuary's Enhanced Medical Expenditure Panel Survey (eMEPS)[3] by the (ii) utilization counts by age and sex from provider surveys, such as the National Ambulatory Medical Care Survey, which include both the institutional and non-institutional population. In the MEPS, source of funds data are aligned by primary payer and include all secondary sources of payments. For example, if an eMEPS respondent had a doctor visit that cost $50, private health insurance may cover $40 with the remainder out-of-pocket. In this example, we recorded $40 as private health insurance (the primary payer) and $10 as out-of-pocket spending (secondary payer). In this manner we aggregated all spending data for each payer, while also recording the number of visits/use for each primary payer. Recording payment data separately for primary and secondary payers allows us to calculate the average payment per visit/use for each payer -- total payments by source of funds (primary and secondary) divided by the number of visits (on a primary payer basis). We then multiply the average payment per use by age, sex and source of funds with the utilization data from provider surveys which include use by both institutionalized and non-institutionalized individuals. This method implicitly assumes that the cost per use/visit for an institutionalized individual is the same as for a non-institutionalized individual, but captures the higher utilization expected from the institutional population.

When no provider survey was available, a second method was used to calculate health spending by age and sex. For home health, other professional services, dental services, non-durables, and durable medical equipment, we computed total health spending by age and sex for 2002 – 2020 by combining eMEPS data for the non-institutionalized population with independently derived estimates of spending for the institutional population, and then controlled these estimates to the overall PHC totals. This method was also used to estimate hospital and physician and clinical expenditures for 2016 – 2020. For the estimates of spending for the institutional population we identified four distinct populations: Medicare beneficiaries enrolled in an institutional setting, Non-Dual Medicaid beneficiaries enrolled in an institutional setting, federal prisoners, and the disabled population in an institutional setting not yet eligible to enroll in Medicare. We then used the population counts and multiplied them by spending estimates created from MCBS or MEPS.

---

[3] The eMEPS dataset includes cost adjustments for certain employer-sponsored insurance (ESI) claims. This dataset was created by matching MEPS event level claims with those in the Marketscan database, and imputing Marketscan spending values into MEPS claims if the MEPS values fell above or below certain thresholds. Matching was completed using nine levels of matching precision based on demographic variables. These adjustments were performed to mitigate known limitations within the MEPS data, including MEPS' underestimation of PHI spending and volatility in year-over-year growth rates. The adjusted eMEPS data more closely matches historical NHE trends.

For the nursing care facilities and continuing care retirement communities category, we supplemented the Medicare, Medicaid, and CHIP program spending estimates by age and sex with data from the MCBS for the over-65 population for all remaining payers. For the under-65 population, we relied on Marketscan data from Truven for the remaining payers. These distributions were controlled to the PHC expenditure levels.

In some instances, source data and methodological constraints required us to average the resulting health spending estimates by age or sex over several years to reflect more reasonable trends over time and within age and sex groups. Where possible, we compared our estimates to other service specific health spending by age and/or sex estimates to ensure reasonableness.

Medicare Current Beneficiary Survey (MCBS) data was used to supplement the MEPS data for the population age 65 and over; and Truven data was used to supplement MEPS data for the population under 65. These data sets were particularly useful for the inpatient hospital and prescription drug components, where the higher sample sizes in MCBS enabled us to improve upon the reliability of the MEPS data. We use MEPS as a primary data source rather than MCBS because we need to have a consistent data source for all age groups, not just the elderly. The final step for both methods was to scale aggregate spending levels to match the control totals in the national health expenditure accounts by type of service and source of funding.

The Provider Relief Fund program and the Paycheck Protection Program (PPP) Loans were created in 2020 as a response to the COVID 19 pandemic. The Provider Relief Fund was designed to offset health care providers losses in revenue from the pandemic as well as give them resources to improve their ability to fight the COVID 19 pandemic. The PPP Loan program was designed to help small businesses maintain employees and cover other eligible expenses by providing forgivable loans. While not directly targeted at healthcare providers, healthcare providers that met eligibility requirements took part in the program. Historical National Health Expenditures for the Provider Relief Fund were estimated using data from the Health Resources and Services Administration (HRSA). Expenditures for the PPP Loan program were estimated using data from the Small Business Administration. The provider relief/PPP loans are included with the Other Payer category and were distributed by service and age/sex using the 2020 shares from Medicare, Medicaid, Private Health Insurance, and out-of-pocket (OOP).

As a final check for reasonableness, we compared the age and sex distributions in the NHE with the age and sex distributions provided in the published MEPS results. Definitional differences, most specifically the exclusion of the institutionalized population in MEPS, accounted for a part of the discrepancies in these series.[4]

[4] For a more complete description of these definitional differences, see Bernard, D. et al.: "Reconciling Medical Expenditure Estimates from the MEPS and the NHE, 2012." Rockville (MD): Agency for Healthcare Research and Quality (Web Only) Available from: https://meps.ahrq.gov/data_files/publications/workingpapers/wp_17003.pdf.