

UCSF

UC San Francisco Previously Published Works

Title

The Serious Illness Population: Ascertainment via Electronic Health Record or Claims Data

Permalink

<https://escholarship.org/uc/item/6jj5w0jz>

Journal

Journal of Pain and Symptom Management, 62(3)

ISSN

0885-3924

Authors

Kelley, Amy S
Hanson, Laura C
Ast, Katherine
et al.

Publication Date

2021-09-01

DOI

10.1016/j.jpainsymman.2021.04.012

Peer reviewed



HHS Public Access

Author manuscript

J Pain Symptom Manage. Author manuscript; available in PMC 2022 September 01.

Published in final edited form as:

J Pain Symptom Manage. 2021 September ; 62(3): e148–e155. doi:10.1016/j.jpainsymman.2021.04.012.

The Serious Illness Population: Ascertainment via Electronic Health Record or Claims Data

Amy S. Kelley, MD MSHS¹, Laura C. Hanson, MD, MPH², Katherine Ast, MSW, LCSW³, Elizabeth L. Ciemins, PhD, MPH, MA⁴, Stephan C. Dunning, MBA⁵, Chris Meskow⁴, Christine S Ritchie, MD, MSPH⁶

¹Brookdale Department of Geriatrics and Palliative Medicine, Icahn School of Medicine at Mount Sinai. New York, NY.

²Division of Geriatric Medicine, Palliative Care Program, University of North Carolina School of Medicine, Chapel Hill, NC.

³American Academy of Hospice and Palliative Medicine, Chicago, IL.

⁴AMGA (American Medical Group Association), Alexandria, VA.

⁵OptumLabs, Eden Prairie, MN

⁶Division of Palliative Care and Geriatric Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA

Abstract

Context: Palliative care can improve the lives of people with serious illness, yet clear operational definitions of this population do not exist. Prior efforts to identify this population have not focused on Medicare Advantage (MA) and commercial healthplan enrollees.

Objectives: We aimed to operationalize our conceptual definition of serious illness to identify those with serious medical conditions (SMC) among commercial insurance and MA enrollees, and to compare the populations identified through electronic health record (EHR) or claims data sources.

Methods: We used de-identified claims and EHR data from the Optum Labs Data Warehouse (2016–2017), to identify adults age ≥18 with SMC and examine their utilization and mortality. Within the subset found in both data sources, we compared the performance of claims and EHR data.

Results: Within claims, SMC was identified among 10% of those aged ≥18 (5.4% ages 18–64, 27% age ≥65). Within EHR, SMC was identified among 9% of those aged ≥18 (5.6% ages 18–64,

Corresponding Author: Amy S. Kelley, MD, MSHS, Amy.kelley@mssm.edu, 1 Gustave L. Levy Place, Box 1070, New York, NY 10029.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclosures: CR is a consultant to the American Academy of Hospice and Palliative Medicine; All other authors report no conflicts of interest or disclosures.

21% ages 65). Hospital, emergency department and mortality rates were similar between the EHR and claims-based groups. Only 50% of people identified as having SMC were recognized by both data sources.

Conclusion: These results demonstrate the feasibility of identifying adults with SMC in a commercially insured population, including MA enrollees; yet separate use of EHR or claims result in populations that differ. Future research should examine methods to combine these data sources to optimize identification and support population management, quality measurement, and research to improve the care of those living with serious illness.

Keywords

Health services research; Palliative Care; Medicare Advantage

Introduction

“Palliative care is specialized medical care for people with serious illnesses.”(1) Conceptually, serious illness is defined as a health condition that carries a high risk of mortality and either negatively impacts a person’s daily function or quality of life, or excessively strains their caregivers.(2–4) Palliative care services provided earlier in the trajectory of illness have been shown to improve quality of life, control symptoms, support patients and families, and lower costs;(5–10) however, many of the patients who could benefit most from palliative care never receive it.

A gold standard operational definition of this vulnerable group simply does not exist, because core elements of the conceptual definition are not measured and reliably reported in existing data, particularly in claims data most commonly used for this purpose. While clinicians understand the concept, systematic identification of people living with serious illness is rare. While there is room for debate about who is seriously ill, consistency of an operational definition is essential for population health to gain insight into gaps in access to palliative care. A shared operational definition also permits comparisons across organizations. To date, efforts to target palliative care and other specialized clinical interventions efficiently have fallen short due to our inability to identify those seriously ill people at greatest risk for high cost, low quality care. This inability to identify this important patient population has also impeded quality measurement work due to the challenges associated with an unclear “denominator” for any quality measure in domains such as advance care planning or symptom management. Conversely, without the ability to identify prospectively and exclude those who are seriously ill, healthcare systems broadly apply preventive quality metrics, such as those for cancer screening or functional improvement that may be inappropriate. In some cases, this particular failure may directly harm patients by subjecting them to tests and treatments dictated by misapplied metrics.(11–17)

Building upon our conceptual definition of serious illness, prior work has tested operational descriptors of serious illness, specifically serious medical conditions, among national samples of Medicare fee-for-service beneficiaries.(3, 4, 18–22) This basic definition performed well in research cohorts identifying 18% of Medicare beneficiaries, over 80% of whom had 2 indicators of high care needs, and who during 6 months of follow-up incurred

on average nearly \$10,000 in Medicare costs while 20% were hospitalized, 25% received emergency department services and over 7% died.(4, 20) But this approach has not yet been translated to real-world health systems or commercial payer populations, which include a broader age range and the Medicare Advantage (MA) population. Testing this approach within the MA population, specifically, is critical as MA enrollment has almost doubled over the past 10 years and MA enrollees now comprise 42% of all Medicare beneficiaries.(23, 24) In addition, changes in MA policy may impact data accuracy and care patterns over time.(25, 26)

In this paper, we operationalize our conceptual definition of serious illness to specifically identify those with serious medical conditions using the Optum Labs Data Warehouse (OLDW). The OLDW is a data resource with claims and electronic health record (EHR) data for commercial insurance beneficiaries and a large MA population. We aimed to identify and compare the populations of adults with serious medical conditions, using both EHR and claims data sources.

Methods

Data Source:

OptumLabs® is an open, collaborative research and innovation center founded in 2013 as a partnership between Optum and Mayo Clinic. The OLDW includes de-identified administrative claims data for commercially insured and Medicare Advantage enrollees and de-identified electronic health record (EHR) data from a nationwide network of provider groups. This study used de-identified administrative claims and EHR data from the OLDW, calendar years 2016–2017. The database contains longitudinal health information on enrollees and patients, representing a diverse mixture of ages, ethnicities and geographical regions across the United States. The majority of commercially insured adults 65 years of age are Medicare Advantage enrollees, while a smaller portion of this population is comprised of those who, despite their Medicare eligibility, receive their health insurance through commercial plans. These enrollees are typically retirees with insurance benefits from their former employer as part of a pension or retirement package or individuals who continue to work, or who are covered under a spouse who works. Medicare fee-for-service beneficiaries are not included in this analysis.(27)

The database includes socioeconomic information, such as race/ethnicity, household income, and education level, for approximately 73% of enrollees. This information is derived from a nationally recognized supplier of consumer marketing data and is a compilation of public data and derived predictive data. The claims data in OLDW includes medical and pharmacy claims, laboratory results and enrollment records for commercial and MA enrollees. The EHR-derived data includes a subset of EHR data that has been normalized and standardized into a single database. A subset of individuals in OLDW have records within both the claims and EHR data and are linkable using a unique identifier within OLDW. The study was exempt from Institutional Review Board approval.

Eligible Populations:

First, we identified the eligible populations separately, using similar logic, in both the claims and EHR data. We defined the claims-based study population as individuals age 18 or older with at least 12 months of continuous enrollment, with no enrollment gaps >45 days, in a commercial or MA health plan with medical coverage in calendar year 2016. The EHR-based study population was limited to individuals age 18 or older with 2 or more qualifying encounters (as a proxy for enrollment) within an integrated delivery network (IDN) during calendar year 2016. Qualifying in-person encounters based upon place of service included outpatient, inpatient, emergency department, observation, urgent care, ambulatory surgery, home visits, hospice, long term acute care hospital, nursing home, assisted living, and skilled nursing facility. Finally, we separately assessed the subset of individuals appearing in both the claims and EHR cohorts and compared the performance of claims and EHR-derived data in identifying those with serious medical conditions within this overlapping population.

Serious Illness Operational Definition:

Claims and EHR data do not routinely include data to evaluate function, quality of life, or caregiver strain, key conceptual components of serious illness (2–4); so we focused on serious medical conditions (SMC) that are associated with high mortality risk or with functional impairments affecting daily living. Building upon prior work, we applied an operational definition of serious medical conditions based upon ICD-10.(20) Medical diagnoses, utilization, and criteria to indicate severity of ongoing illness included: advanced cancer; end stage or stage 5 renal disease; dementia; advanced lung disease and advanced congestive heart failure only if using home oxygen or hospitalized for the condition; advanced liver disease; diabetes with severe complications; advanced Parkinson's disease with indicator of dementia or selected durable medical equipment (DME); other neurodegenerative diseases; hip fracture, over age 70; stroke requiring hospital admission; and HIV with AIDS complications (full specification in Appendix 1).

Baseline and Measurement Periods:

A baseline year was established for all eligible patients as calendar year 2016, during which serious medical conditions were identified. We then measured health care utilization, including emergency department (ED) visits, inpatient admissions, and hospice admissions during calendar year 2017.

Analyses:

We first assessed the prevalence of SMC overall and of each disease group in the baseline year within the eligible claims and EHR populations. Descriptive statistics were assessed in the baseline year, including age, sex, race, ethnicity, education and income level, insurance type, number of chronic conditions, and use of DME. Health care utilization was assessed in the measurement year and compared across the following groups: all adults age 18 and over, adults age 18–64, and adults age 65 and over without serious medical conditions; and all adults age 18 and over, adults age 18–64, and adults age 65 and over with serious medical conditions. Lastly, we examined the subset of the population found in both data sources and

compared the performance of claims and EHR data in identifying those with serious medical conditions within this overlapping population.

Results

The eligible population identified within OLDW 2016 claims included 11,516,072 individuals, while the EHR data revealed 17,367,524 eligible individuals. Within the claims-based sample, serious medical conditions (SMC) were identified among 10% of those aged 18 or older (5.4% ages 18–64, 27% ages 65 and older). Within the EHR-based sample, serious medical conditions were identified among 9% of those aged 18 or older (5.6% ages 18–64, 21% ages 65 and older).

The demographic and health characteristics of the samples are described in Table 1. In both data sources, about one-third of the adults identified with SMC also had 2 or more chronic conditions. The claims-based sample revealed a higher prevalence of DME use, which was observed rarely within the EHR-based sample. Figure 1 displays the prevalence of SMC disease groups by age and data source (Figure 1). While similar in most disease groups, the claims-based sample revealed a higher percentage of adults with advanced lung disease and Parkinson's disease, while the EHR-based sample revealed a higher percentage of adults with advanced heart failure. Measures of utilization across the SMC populations identified by each data source are similar in many respects (Figures 2 and 3). The EHR reveals a higher percentage of ED admissions, particularly among the age 18–64 groups.

Among the subset of individuals appearing in both data sources, 172,540 people with SMC were identified by at least one source, while 86,460 (50%) were identified by both data sources (Figure 4). Among the overlapping subset's claims SMC population, 52% were also identified in EHR, while 48% were not. Among the overlapping subset's EHR SMC population, 87% were also identified in claims, while 13% were not.

Discussion

These results demonstrate the feasibility of identifying adults with serious illness, operationalized as serious medical conditions, in a commercially insured population, including Medicare Advantage enrollees. The high rates of hospital admissions in the study populations confirm that these individuals are suffering with a significant burden of illness. The percentage with ED admissions, in particular, indicate both a need and an opportunity to improve care for these seriously ill individuals. In comparison to other studies, the SMC population over age 65 identified by claims and by EHR experienced a level of hospital admissions comparable to those with traditional fee-for-service Medicare, about 20% over 6 months.⁽²⁰⁾ Overall, though, utilization among those with SMC in this study was lower than utilization among people with SMC identified as having additional characteristics of serious illness, including functional impairment, among whom 44–46% were hospitalized over 1 year.^(4, 19) This finding provides further evidence to support addition of standardized data on function in healthcare encounters, either captured by claims that link enhanced payment to the greater vulnerability and healthcare needs of those with functional dependency, or within the EHR, in order to better characterize SMC

populations, enhance prognostication and augment quality of care measures and outcomes in serious illness populations.(28) While existing claims-based algorithms may attempt to proxy measures of frailty or functional decline, direct measurement would add precision.(22, 29)

When seeking to improve quality of care for people with serious illness, researchers and administrators will have varied access to claims or EHR data sources. The differences in the SMC populations identified by claims and EHR-derived data expose important considerations for leaders of health care systems seeking to identify a seriously ill population for purposes of program enrollment or quality measurement. First, DME, such as home oxygen or hospital bed, is easily identified in claims data, yet poorly identified in the EHR. Thus, conditions requiring the use of DME as an indicator of disease severity, such as advanced lung disease and Parkinson's disease, will be more readily identified in claims than in EHR, as we found in this study. Furthermore, our examination of the subset population found in both data sources, revealed that only 50% of people were identified as having SMC in both data assets. Claims data provide a more comprehensive view of the services billed for the care of an individual, but may lack more clinically detailed information that could be needed to capture illness severity, confirm service eligibility, or assess a quality measure.(30) In contrast, the EHR may provide richer clinical data, but will not capture services provided outside of that healthcare system's network, a phenomenon described as "leakage." In addition, the small portion of individuals identified by EMR and not by claims data, may be due to diagnoses present on the EHR patient problem list but not submitted on a particular claim, or reflect an unpaid claim. This group could also include people who switched into FFS Medicare or from one primary insurer to another. Overall, EHR data alone would detect 56.1% of patients with SMC and Claims data alone would detect 92.8% of patients with SMC, compared to the combination of EHR and Claims data. Thus, use of both data sources is recommended when possible, and we anticipate that in the future the incorporation of novel techniques including natural language processing may support the use of these combined data as the gold standard.(31) For now, however, frequently only one data source may be available and either can be used. In these cases, users of claims data will likely achieve greater sensitivity yet be unable to assess measures requiring greater clinical detail, while those using only EHR data will have a greater depth of information available, but will miss appropriate people due to "leakage" or inability to observe DME, which could disproportionately impact specific disease groups. In addition, a more sensitive definition is optimal when the goal is to examine access; however, a more specific denominator population may be optimal when defining quality measures.

In addition to the data limitations already noted, utilization and outcome measures have notable limitations in both datasets. Neither the claims nor EHR data provide adequate information on hospice enrollment or death. The rates of hospice use observed are low and likely underestimated, and death cannot be distinguished from attrition or disenrollment in most cases. Furthermore, ongoing changes in CMS MA payment policy, including how risk scores are calculated, will surely impact and likely improve the accuracy of encounter data over time.(26) While current assessments reveal the potential risk of inflated diagnoses within MA, the data appear comparable to FFS in accuracy.(25) Lastly, we found significant differences by race/ethnicity and substantial proportions of unknown and missing data. The

source of this information varies and often is not based upon self-report. Incomplete or misattributed race and ethnicity data in both claims and EHRs can contribute to systemic racism and ongoing health disparities, a particularly important consideration for providers of palliative care seeking to improve equity.

Despite these limitations, defining a consistent operational approach to identifying the seriously ill population allows population-based planning and benchmarking across settings and providers, including proportion of those with access to palliative care service, hospice, or with use of high-intensity treatments close to death. Thus, health system leaders and researchers can still use these imperfect approaches to improve the quality of care of those living with serious illness in multiple ways. In California, for example, the state legislation (SB1004) requires that Medicaid beneficiaries with serious illness have access to palliative care. The approach described is being used to assess population size, service need and capacity. Similarly, any healthcare system could apply this approach within its EHR data to identify those with SMC, and plan to strategies to improve access to palliative care. By defining the population within claims data, an ACO could proactively target expanded supportive and palliative care services to high risk patients. This approach can also promote accountability for high quality care. For example, an outpatient primary care network could identify their SMC population and track frequency of advance care planning as a quality metric.

National efforts to improve care for the seriously ill can be advanced by building upon this work. First, functional status and other patient-centered characteristics of illness burden and severity must become a mandatory part of standardized documentation in order to achieve greater specificity in identifying those with serious illness. Some diagnoses, including dementia, require more consistent documentation of severity of illness so that individuals with these conditions are recognized and provided necessary services in a timely manner. Finally, reaching agreement on quality indicators and the appropriate denominators for palliative care, even if imperfect, will help to assure higher quality care for those living with serious illness.

In conclusion, this work offers the first glimpse of the seriously ill populations within a large cohort served by MA and commercial payers. It provides insights to the strengths and limitations of EHR and claims data and gives researchers, policy makers and health system leaders important information about what information can and cannot be easily gleaned from different types of readily available data, as well as several ways this approach can be applied today to begin to improve the care of those living with serious illness.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding:

This work was supported by the 2017 AARP Quality Measure Innovation Program with OptumLabs® and the National Quality Forum. CR and LCH receive funding from National Institute of Nursing Research U2CNR014637. AK receives funding from National Institute on Aging K24AG062785.

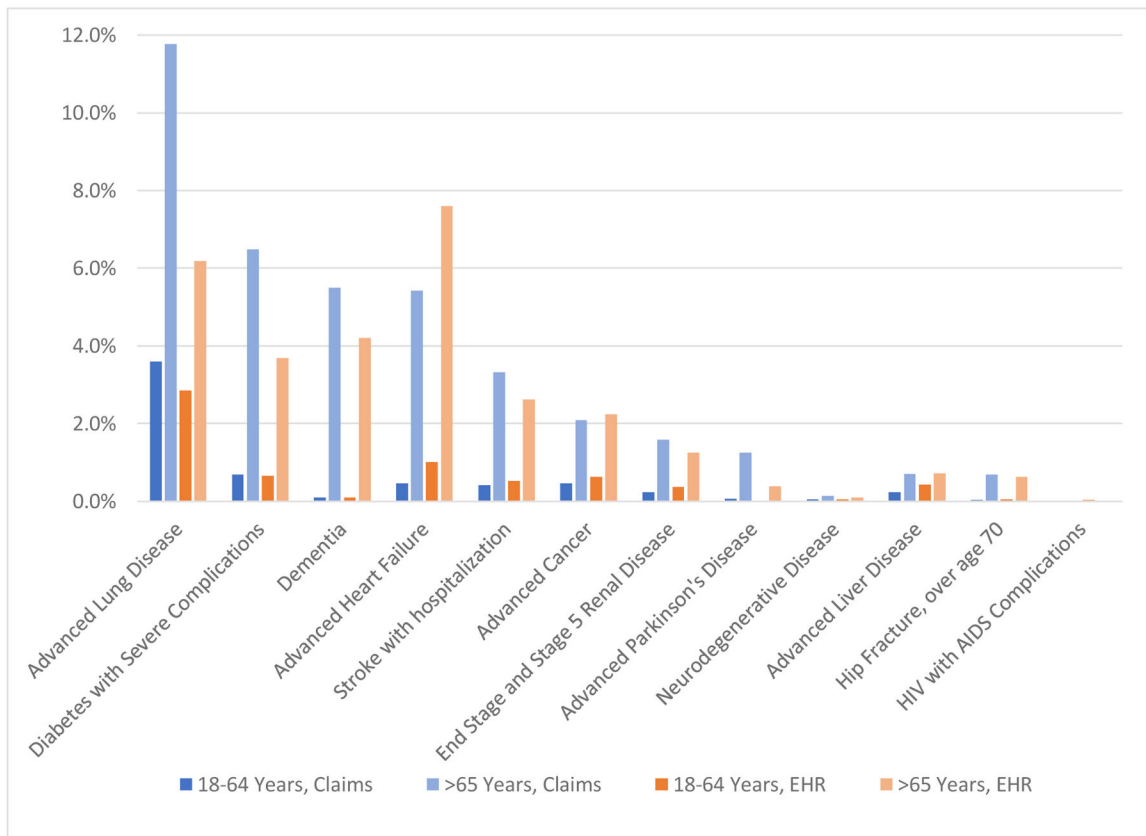
References:

1. Center to Advance Palliative Care: About Palliative Care [Available from: <https://www.capc.org/about/palliative-care/>].
2. National Consensus Panel Clinical Practice Guidelines for Quality Palliative Care October 2018 [updated 4th.].
3. Kelley AS. Defining “serious illness”. *J Palliat Med.* 2014;17(9):985. [PubMed: 25115302]
4. Kelley AS, Bollens-Lund E. Identifying the Population with Serious Illness: The “Denominator” Challenge. *Journal of palliative medicine.* 2018;21:S7–S16. [PubMed: 29125784]
5. Temel JS, Greer JA, Muzikansky A, Gallagher ER, Admane S, Jackson VA, et al. Early palliative care for patients with metastatic non-small-cell lung cancer. *N Engl J Med.* 2010;363(8):733–42. [PubMed: 20818875]
6. Brumley R, Enguidanos S, Jamison P, Seitz R, Morgenstern N, Saito S, et al. Increased satisfaction with care and lower costs: results of a randomized trial of in-home palliative care. *J Am Geriatr Soc.* 2007;55(7):993–1000. [PubMed: 17608870]
7. Rabow M, Kvale E, Barbour L, Cassel JB, Cohen S, Jackson V, et al. Moving upstream: a review of the evidence of the impact of outpatient palliative care. *Journal of palliative medicine.* 2013;16(12):1540–9. [PubMed: 24225013]
8. Bakitas M, Lyons K, Hegel M, Balan S, Brokaw F, Seville J, et al. Effects of a palliative care intervention on clinical outcomes in patients with advanced cancer: the Project ENABLE II randomized controlled trial. *JAMA.* 2009;302(7):741–9. [PubMed: 19690306]
9. Jordhoy MS, Fayers P, Loge JH, Ahlner-Elmqvist M, Kaasa S. Quality of life in palliative cancer care: results from a cluster randomized trial. *J Clin Oncol.* 2001;19(18):3884–94. [PubMed: 11559726]
10. Smith S, Brick A, O’Hara S, Normand C. Evidence on the cost and cost-effectiveness of palliative care: a literature review. *Palliative medicine.* 2014;28(2):130–50. [PubMed: 23838378]
11. Pogach L, Aron D. The other side of quality improvement in diabetes for seniors: a proposal for an overtreatment glycemic measure. *Archives of internal medicine.* 2012;172(19):1510–2. [PubMed: 22965316]
12. Schonberg MA, McCarthy EP. Mammography screening among women age 80 years and older: consider the risks. *Journal of Clinical Oncology.* 2009;27(4):640–1. [PubMed: 19075269]
13. Walter LC, Lewis CL, Barton MB. Screening for colorectal, breast, and cervical cancer in the elderly: a review of the evidence. *The American journal of medicine.* 2005;118(10):1078–86. [PubMed: 16194635]
14. Kelley AS. Epidemiology of care for patients with serious illness. *J Palliat Med.* 2013;16(7):730–3. [PubMed: 23718871]
15. Leff B, Weston CM, Garrigues S, Patel K, Ritchie C, Care NHBP, et al. Home-based primary care practices in the United States: Current state and quality improvement approaches. *Journal of the American Geriatrics Society.* 2015;63(5):963–9. [PubMed: 25940131]
16. Ritchie CS, Editor A, Rabow MW. Fostering a Quality of Care Culture in Community-Based Palliative Care: What Will It Take? *Journal of palliative medicine.* 2015;18(7):568–9. [PubMed: 26053462]
17. Dy SM, Kiley KB, Ast K, Lupu D, Norton SA, McMillan SC, et al. Measuring what matters: top-ranked quality indicators for hospice and palliative care from the American Academy of Hospice and Palliative Medicine and Hospice and Palliative Nurses Association. *Journal of pain and symptom management.* 2015;49(4):773–81. [PubMed: 25697097]
18. Kelley AS, Bollens-Lund E, Covinsky KE, Skinner JS, Morrison RS. Prospective Identification of Patients at Risk for Unwarranted Variation in Treatment. *J Palliat Med.* 2018;21(1):44–54. [PubMed: 28772096]
19. Kelley AS, Covinsky KE, Gorges RJ, McKendrick K, Bollens-Lund E, Morrison RS, et al. Identifying Older Adults with Serious Illness: A Critical Step toward Improving the Value of Health Care. *Health Services Research.* 2016;52(1):113–31. [PubMed: 26990009]

20. Kelley AS, Ferreira KB, Bollens-Lund E, Mather H, Hanson LC, Ritchie CS. Identifying older adults with serious illness: transitioning from ICD-9 to ICD-10. *Journal of pain and symptom management*. 2019;57(6):1137–42. [PubMed: 30876955]
21. Faurot KR, Jonsson Funk M, Pate V, Brookhart MA, Patrick A, Hanson LC, et al. Using claims data to predict dependency in activities of daily living as a proxy for frailty. *Pharmacoepidemiology and drug safety*. 2015;24(1):59–66. [PubMed: 25335470]
22. Mather HL, Bollens-Lund E, Husain M, Kelley AS. Candidate Claims-Based Indicators of Functional Impairment: An Exploration in a Sample of Medicare Beneficiaries. *Journal of pain and symptom management*. 2019;58(3):e8–e11.
23. Jacobson G, Damico A, Neuman T. A dozen facts about Medicare Advantage. 2018. The Henry J Kaiser Family Foundation. 2019.
24. Meyers DJ, Johnston KJ. The Growing Importance of Medicare Advantage in Health Policy and Health Services Research. *JAMA Health Forum*. 2021;2(3):e210235–e.
25. MEDPAC. Ensuring the accuracy and completeness of Medicare Advantage encounter data. 2019.
26. J CSD-TRM. Health Affairs Blog [Internet]. Available from: <https://www.healthaffairs.org/doi/10.1377/hblog20190221.696651/full/>.
27. OptumLabs and OptumLabs Data Warehouse (OLDW) Descriptions and Citation. PDF. Reproduced with permission from OptumLabsEden Prairie, MN; 202072020.
28. Hanson LC SA, Burstin H. . Quality and Outcome Measures. In: DM Amy Kelley, eds., editor. *Meeting the Needs of Older Adults with Serious Illness: Challenges and Opportunities in the Age of Health Care Reform*. New York, NY: Springer Publications; 2103.
29. Kim DH, Schneeweiss S, Glynn RJ, Lipsitz LA, Rockwood K, Avorn J. Measuring frailty in Medicare data: development and validation of a claims-based frailty index. *The Journals of Gerontology: Series A*. 2018;73(7):980–7.
30. Enguidanos S, Rahman A, Fields T, Mack W, Brumley R, Rabow M, et al. Challenges in Using Insurance Claims Data to Identify Palliative Care Patients for a Research Trial. *Journal of Pain and Symptom Management*. 2020;60(5):1012–8. [PubMed: 32592736]
31. Guo A, Foraker R, White P, Chivers C, Courtright K, Moore N. Using electronic health records and claims data to identify high-risk patients likely to benefit from palliative care. *The American Journal of Managed Care*. 2021;27(1):e7–e15. [PubMed: 33471463]

Key Message:

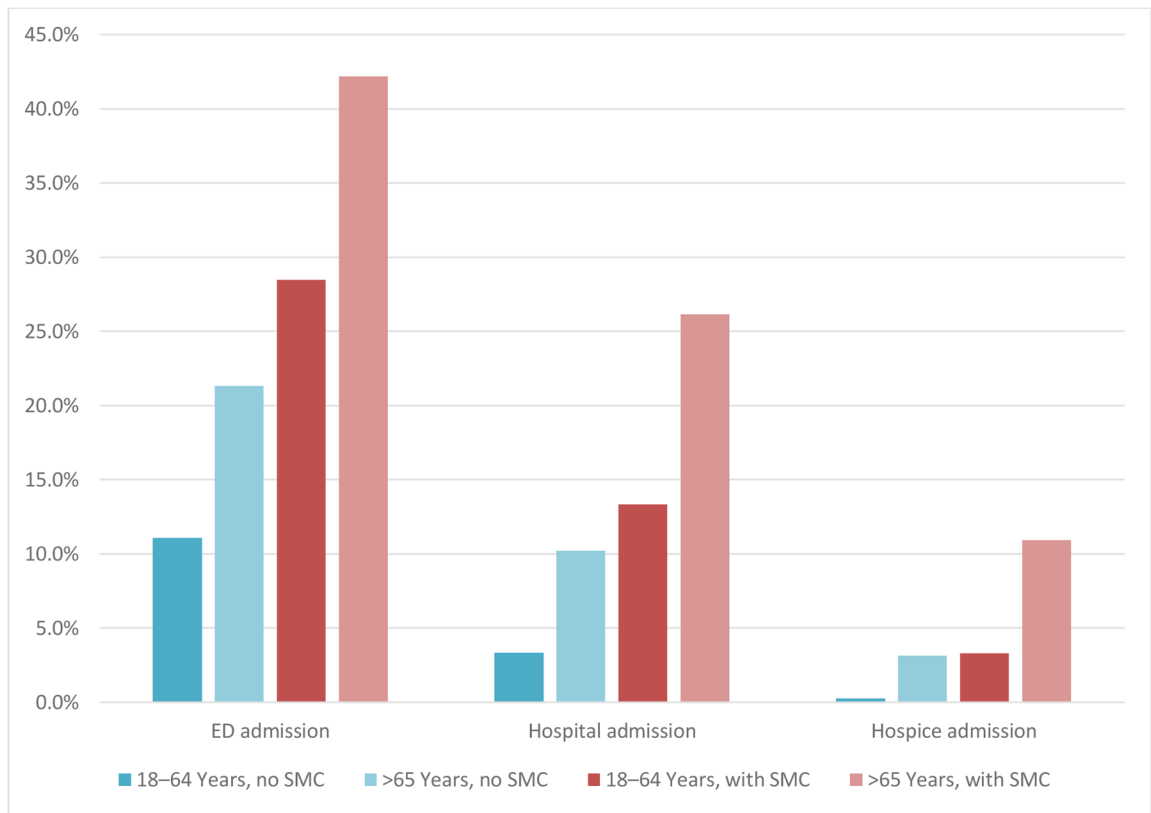
This article describes a retrospective cohort study that aims to identify a population of adults with serious medical conditions using claims and electronic health record data from a commercially insured and Medicare Advantage population. The results describe the performance and strengths and limitations of each data source.



EHR = Electronic Health Record

Figure 1. Percent of Population with Serious Medical Conditions by Disease Groups, Age and Data Source

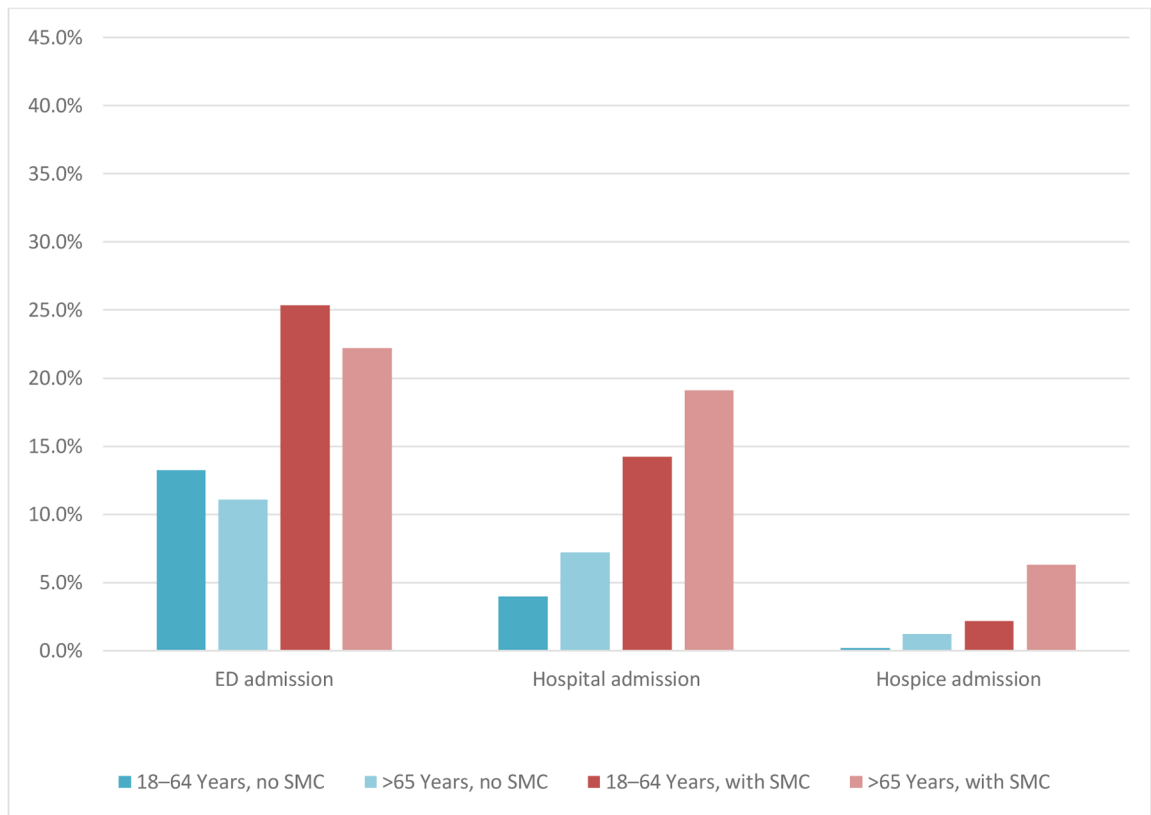
EHR = Electronic Health Record



SMC= Serious Medical Condition

Figure 2. Utilization among Adults with and without Serious Medical Condition (SMC), by Age in Claims

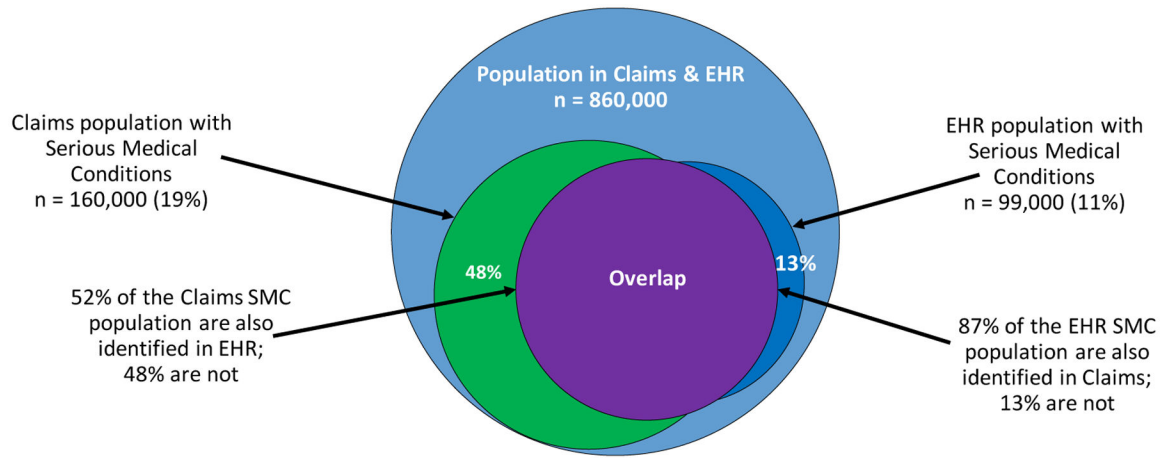
SMC= Serious Medical Condition



SMC = Serious Medical Condition

Figure 3. Utilization among Adults with and without Serious Medical Condition (SMC), by Age in EHR

SMC = Serious Medical Condition



EHR = Electronic Health Record; SMC = Serious Medical Conditions

Figure 4. Identification of People with Serious Medical Conditions Present in Both Claims and EHR Data

EHR = Electronic Health Record; SMC = Serious Medical Conditions

Table 1.

Characteristics of Population with Serious Medical Conditions by Disease Groups, Age and Data Source

	Claims data			EHR data		
	Full Sample	Seriously Ill Population		Full Sample	Seriously Ill Population	
		Age 18–64	Age 65+		Age 18–64	Age 65+
<i>N</i>	11,516,072	475,469	723,011	17,367,524	737,838	895,304
Age, mean years	49.2	49.4	76.5	50.2	49.9	76.9
Age 65 years or over (%)	23%	0%	100%	24%	0%	100%
Female (%)	52%	54.9%	56.2%	58%	52.3%	54.1%
Race/ethnicity						
Non-Hispanic White (%)	44%	43.5%	57.4%	75%	73.0%	83.3%
African American (%)	11%	14.7%	17.0%	10%	15.7%	8.8%
Hispanic (%)	14%	8.5%	7.5%	6%	6.5%	3.3%
Asian (%)	5%	3.9%	2.8%	2%	1.6%	1.2%
Unknown (%)	13%	13.9%	11.2%	12%	9.7%	6.7%
Missing (%)	17%	16.5%	4.2%	0%	0%	0%
Education (% with college)	NA	NA	NA	24.2	23.6%	23.3%
Income, % in Bottom Quartile	NA	NA	NA	22%	22.2%	22.9%
2 or more additional chronic conditions (%)	6.0%	21.3%	44.3%	6.2%	24.3%	40.8%
Durable Medical Equipment claim (%)	1.2%	2.8%	7.1%	0.1%	0.4%	0.8%
Medicare (%)				13%	10.2%	51.1%
Medicaid (%)				8%	17.5%	4.4%
Commercial (%)				55%	56.5%	33.5%

EHR = Electronic Health Record; NA = Not Available in this data source

* All differences between Claims and EHR populations, by age group, are statistically significant at $P < 0.0001$.