

**Brief Report**

# Novel Data Linkages to Characterize Palliative and End-Of-Life Care: Challenges and Considerations



Cara L. McDermott, PhD, PharmD, Ruth A. Engelberg, PhD, Cossette Woo, MSW, Li Li, MA, MPA, Catherine Fedorenko, MMSci, Scott D. Ramsey, MD, PhD, and J. Randall Curtis, MD, MPH  
*Cambia Palliative Care Center of Excellence Department of Medicine (C.L.M., R.A.E., J.R.C.), University of Washington, Seattle, Washington; Hutchinson Institute for Cancer Outcomes Research Fred Hutchinson Cancer Research Center (C.L.M., L.L., C.F., S.D.R.), Seattle, Washington; and Department of Social Welfare University of Washington (C.W.), Seattle, Washington, USA*

**Abstract**

**Context.** Working groups have called for linkages of existing and diverse databases to improve quality measurement in palliative and end-of-life (EOL) care, but limited data are available on the challenges of using different data sources to measure such care.

**Objectives.** To assess concordance of data obtained from different sources in a novel linkage of death certificates, electronic health records (EHRs), cancer registry data, and insurance claims for patients who died with cancer.

**Methods.** We joined a database of Washington State death certificates and EHR to a data repository of commercial health plan enrollment and claims files linked to registry records from Puget Sound Cancer Surveillance System. We assessed care in the last month including hospitalizations, intensive care unit (ICU) admissions, emergency department visits, imaging scans, radiation, and hospice, plus chemotherapy in the last 14 days. We used a Chi-squared test to compare differences between health care in EHR and claims.

**Results.** Records of hospitalization, ICU use, and emergency department use were 33%, 15%, and 33% lower in EHR versus claims, respectively. Radiation, hospice, and imaging were 6%, 14%, and 28% lower, respectively, in EHR, but chemotherapy was 4% higher than that in claims. These differences were statistically different for hospice ( $P < 0.02$ ), hospitalization, ICU, ER, and imaging (all  $P < 0.01$ ) but not radiation ( $P = 0.12$ ) or chemotherapy ( $P = 0.29$ ).

**Conclusion.** We found substantial variation between EHR and claims for EOL health-care use. Reliance on EHR will miss some health-care use, while claims will not capture the complex clinical details in EHR that can help define the quality of palliative care and EOL health-care utilization. *J Pain Symptom Manage* 2019;58:851–856. © 2019 American Academy of Hospice and Palliative Medicine. Published by Elsevier Inc. All rights reserved.

**Key Words**

*Palliative care, end-of-life, cancer, databases, electronic health records, insurance claims*

**Introduction**

Optimizing end-of-life (EOL) care requires comprehensive analysis of the incidence of and reasons for unwanted or nonbeneficial high-intensity health care. Many studies rely on administrative claims databases or single-site retrospective analyses to elucidate who is likely to be hospitalized or go to the emergency

department (ED) at end of life<sup>1-6</sup> as these data sources are readily available. However, workgroups such as Measuring What Matters of the American Academy of Hospice and Palliative Medicine and Hospice and Palliative Nurses Association have called for unique data linkages to improve quality measurement in palliative and EOL care.<sup>7</sup> The utilization of widely available routine data, such as electronic health records (EHR),

*Address correspondence to:* Cara L. McDermott, PhD, PharmD, Cambia Palliative Care Center of Excellence, Pulmonary, Critical Care and Sleep Medicine, University of Washington, Seattle, WA, USA. P. O. Box 359765. E-mail: [clm2@uw.edu](mailto:clm2@uw.edu)

*Accepted for publication:* July 16, 2019.

can facilitate palliative care and EOL research while avoiding the high costs and time intensity of primary data collection.<sup>8</sup>

Reliance on one data source may not provide a complete landscape of health-care services provided and the reasons for such services. For example, Medicare claims underestimated the number of palliative care consultations delivered compared to findings from a phone survey of all Medicare-certified hospitals in Colorado to assess specialty palliative care services.<sup>9</sup> Ideally, researchers would be able to link insurance claims, which capture all billed health-care use, to clinical registries such as the Surveillance, Epidemiology, and End Results (SEER) cancer registry to confirm diagnosis and then to EHR to identify clinical and decision-making information not available in claims or registries.

In this brief report, we summarize challenges in creating unique data linkages to research palliative and EOL care. To illustrate some of these challenges, we report on our attempt to extend our previous research on EOL care for commercially insured adults.<sup>10</sup> We created a novel linkage of death certificates, a large health system EHR, cancer registry records, and insurance enrollment and claims files. Our goal was to use the linked database to identify high-intensity EOL health-care utilization from claims among people with a confirmed cancer diagnosis from SEER and then to explore the reasons for such health care per the EHR. We describe our data linkage experience and offer considerations for researchers considering similar work.

## Methods

First, we identified eligible subjects in a database of Washington State death certificates linked to UW Medicine, a health system spanning four hospitals, a comprehensive cancer center, and a network of outpatient primary and specialty care clinics. Of these subjects, we included those with a diagnosis of solid or hematologic malignancy who died in 2015–2016 and were aged 18 years or older. To ensure sufficient contact with UW Medicine to assess utilization, patients had to have at least one inpatient visit that was not for an elective procedure or two outpatient visits that were not for a second opinion in their last six months of life. We limited our years to 2015–2016 as this linkage was part of a mixed-methods study of EOL health care and we wanted proximal data to minimize recall bias in interviews with bereaved caregivers.

Then, we linked this cohort to a Hutchinson Institute for Cancer Outcomes Research (HICOR) database that includes registry data on cancer diagnoses from the Puget Sound Cancer Surveillance System, a

part of the SEER program. These registry data were linked to enrollment and claims files from two commercial nonprofit insurers, Premera Blue Cross and Regence BlueShield, covering ~50% of commercially insured patients in the region. For inclusion in the database linkage (Figure 1), subjects had to be aged 18 years or older, died in 2015–2016, continuously insured in their last 30 days of life, and incurred at least one insurance claim in their last 90 days of life.

We used Registry Plus Link Plus 2.0 to link between the HICOR and UW databases. Link Plus is a probabilistic record linkage program developed at the Centers for Disease Control and Prevention to facilitate record linkages between cancer registries. We matched patient records using Social Security Number, last name, first name, date of birth, and death date. After receiving the linkage report, an analyst (L. L.) manually reviewed matches for quality. After linking the UW and HICOR databases, we extracted EHR data for all subjects in the linked cohort. The principal investigator (C. L. M.) and a research assistant (C. W.) reviewed available EHR data for patients' inpatient and outpatient visits, telephone encounters, and documentation from other health systems that had been uploaded into the UW Medicine EHR.

From the abstracted EHR data, we created a binary variable (yes/no) as to whether we observed documentation of hospitalizations, intensive care unit (ICU) admission, ED visits, advanced imaging scans or radiation in the last 30 days of life, chemotherapy use in the last 14 days of life, and as a comparison of non-high-intensity health-care utilization, hospice use in the last 30 days.<sup>11,12</sup> We then compared our EHR findings to those from the claims database. We used a Chi-squared test to compare recorded health-care use in each database and report statistically significant differences between databases. Finally, from the EHR, we qualitatively examined documentation of: 1) goals of care discussions, 2) roles/needs of caregivers, 3) hospice discussions, and 4) palliative care needs and services.

Our research was in accordance with data-use agreements between the insurers and HICOR to use claims for research purposes. We received a waiver of consent and authorization from UW Medicine to access subjects' EHR per Washington state law regarding access to personal health information of deceased individuals. The University of Washington Institutional Review Board approved this study.

## Results

We linked 142 subjects from the initial cohort of 2025 subjects from UW Medicine and 1336 subjects from HICOR who met eligibility criteria. After manual

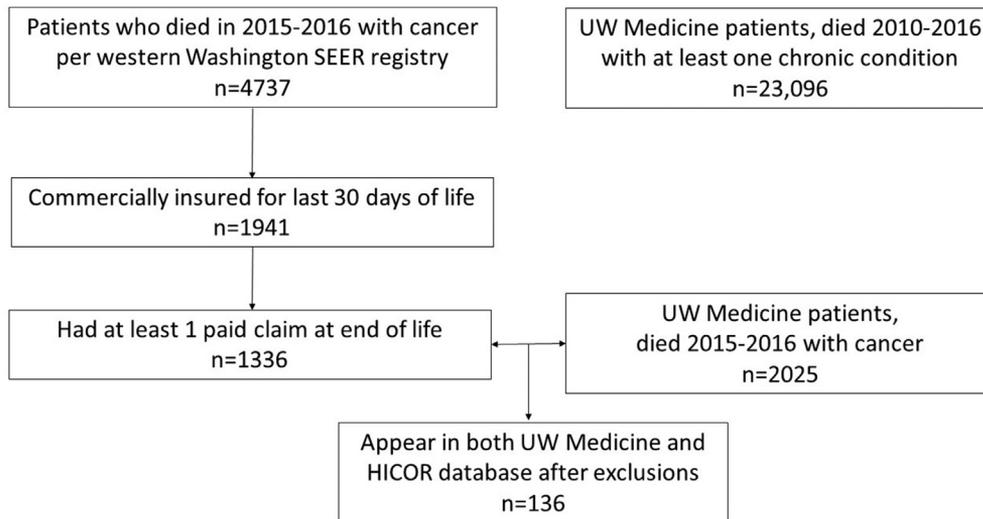


Fig. 1. Cohort diagram for population identification and subsequent linkage. SEER = Surveillance, Epidemiology, and End Results; HICOR = Hutchinson Institute for Cancer Outcomes Research.

review, we excluded six inaccurate matches, for a total of 136 subjects (7%). On average, subjects were aged 64.5 years when they received a cancer diagnosis (range: 35–90 years) and 40% were female. Most (89%) were white and married (87%) at the time of the cancer diagnosis. The most common cancer diagnosis was lung cancer (19%) followed by pancreatic cancer (12%).

Using the linked database, in the last month of life, 49% of patients ( $n = 66$ ) visited the ED, 57% ( $n = 77$ ) were hospitalized, 24% ( $n = 33$ ) had an ICU admission, 57% ( $n = 77$ ) received advanced imaging, 14% ( $n = 19$ ) underwent radiation, 7% ( $n = 10$ ) had

chemotherapy, and 46% ( $n = 63$ ) enrolled in hospice. When using the EHR to identify health-care utilization, we consistently underestimated all services except chemotherapy (Figure 2). Our estimates of hospitalization, ICU admission, and ED use were 33%, 15%, and 33% lower in the EHR, respectively, than those in the claims database. We found 6% lower radiation receipt, 14% lower hospice use, 28% less imaging, but 4% higher chemotherapy use when comparing EHR to insurance claims. When comparing claims to EHR as sources of data, we found statistically significant differences for frequency of hospitalization, ICU, ED use, or imaging (all

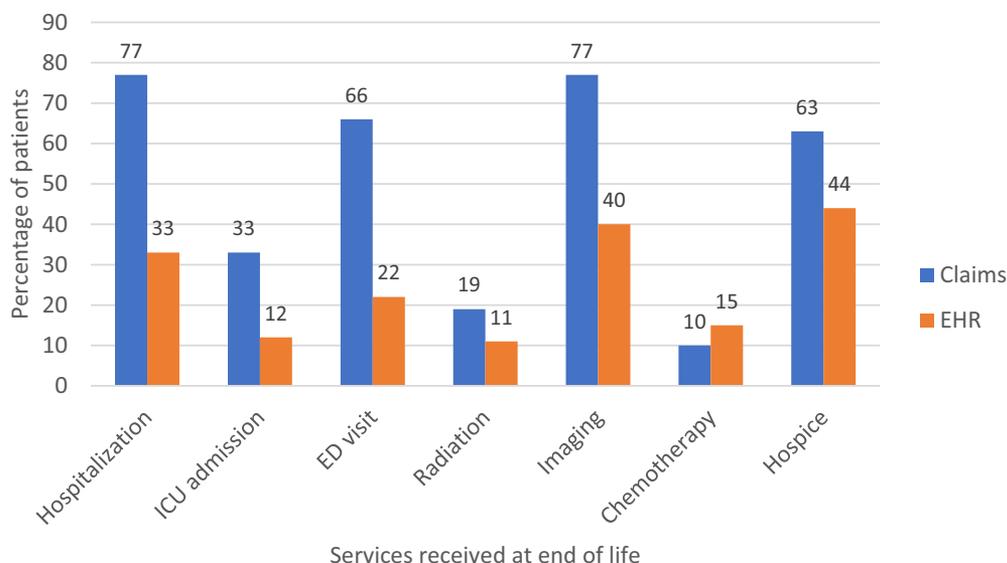


Fig. 2. Comparison of services identified by insurance claims versus EHR for the last 30 days of life for patients who died in 2015–2016 with cancer. ICU = intensive care unit; ED = emergency department; EHR = electronic health record.

$P < 0.01$ ) and hospice use ( $P < 0.02$ ). We found no significant difference between claims and EHR regarding radiation ( $P = 0.12$ ) or chemotherapy ( $P = 0.29$ ). While we attempted to locate advance care planning in claims using Current Procedural Terminology codes 99,497 and 99,498, <1% of patients had such a code, thus we did not evaluate these codes.

From the EHR, we found goals of care documentation or a discussion about goals of care for 112 of 136 patients (82%). Caregiver availability, limits of available social/caregiving support, or patient caregiving needs were noted for 95 patients (70%). We identified documented discussions about hospice for 52 patients (38%). Other narrative data we found in the EHR that are not available in claims included a description of decision-making around chemotherapy or radiation, including family preferences for palliative care consultations while patients were receiving chemotherapy or radiation ( $n = 11, 8\%$ ).

## Discussion

In this study, we created a novel linkage of death certificates, cancer registry records, insurance claims, and EHR data. We found substantial variation between what was reported in the EHR and what we found in claims regarding EOL health-care use. We note it is possible to link claims and EHR to provide a more complete picture of clinical details and decision-making, as we were able to find narrative data in the EHR for some patients regarding EOL decision-making.

Unfortunately, when we linked our cohort to insurance claims, we lost over 90% of our sample, attrition similar to that noted by West and colleagues<sup>13</sup> when they attempted to link EHR and insurance claims. Similarly, Paterno and colleagues linked EHR data and claims of only 4% of their study population.<sup>14</sup> Part of this attrition is the result of our inclusion criteria. Although we assessed a commercially insured population that included Medicare Advantage patients covered by programs offered by Premera and Regence, we did not have access to claims for the large number of patients with cancer who are Medicare fee-for-service insured.

Additionally, commercially insured patients may change health-care plans, which limits the ability to observe health-care utilization.<sup>15</sup> While such switching is unlikely at end of life, we addressed this by requiring subjects be insured in their last month and incurred at least one paid claim in the last 90 days of life, which reduced our available population for linkage. While this criterion may exclude patients with limited EOL health-care use, this criterion helps ensure we are only including patients enrolled in their last 30 days,

which is the timeframe for our study outcomes. It is important to note attrition may occur as the result of differences in coding across databases, for example, patient data in insurance claims may not match EHR data if a date of birth is entered incorrectly in one of the systems.<sup>16</sup> Finally, some commercially insured patients may seek care at another regional cancer center. UW Medicine is the largest publicly funded health system in the Puget Sound region; as such, this health system serves those who are uninsured or on Medicaid. Such patients would not be captured in the HICOR database.

In this analysis, we treated claims as a gold standard for health-care use. We acknowledge that insurance claims may be incomplete, but insurance claims often more accurately capture health care delivered if a patient uses multiple health-care providers. A previous evaluation found claims captured 83%–97% of EOL services for patients with cancer.<sup>17</sup> By relying on EHR data alone, we underestimated EOL health-care utilization because the EHR often does not contain information on health care delivered outside of the UW Medicine system unless a provider in the other health system or the patient provides documentation of such care to enter into the UW Medicine EHR.

Other studies have found high concordance between insurance claims and the EHR for services that are specific and generate a claim (e.g., influenza vaccination) compared with services that do not generate a claim (e.g., body mass index measurement)<sup>18</sup> or are performed as quality-improvement initiatives.<sup>19</sup> As we relied on one data source for hospitalization data (e.g., EHR from one system or claims data from one insurer), we would have missed out-of-system hospitalizations.<sup>20</sup> It is also possible that miscoding of services contributes to the disconnect between EHR documentation and billed services in claims data.<sup>21</sup> Finally, the EHR database includes Seattle Cancer Care Alliance consultations. We found more notations indicating chemotherapy receipt in EHR versus claims; some were for patients receiving chemotherapy alongside clinical trial enrollment. In that case, the clinical trial sponsor may pay for chemotherapy rather than the insurer. If a patient comes to UW Medicine for a second opinion, the patient may return to another health system for care, which would not be recorded in UW Medicine EHR but would be paid for by the insurer.

We encountered several limitations during our data linkage. First, the database we used for this study resulted from a novel agreement between HICOR, local commercial insurers, and the Puget Sound Cancer Surveillance System, which allowed us to link to the UW Medicine EHR. This unique collaboration allowed us to address one of the challenges of using other claims database linkages that may prevent data

linkages using personal health information.<sup>22</sup> Unfortunately, as cancer is largely a disease of older adults and most older adults are enrolled in fee-for-service Medicare, many people did not have insurance claims in our data set. Second, our EHR sample included patients receiving care at UW Medicine, using Dartmouth Atlas criteria for attribution of care to a health-care system (i.e., one nonsurgical hospitalization or two outpatient visits).<sup>23</sup> Patients referred for a single-visit second opinion would be included in the claims database but excluded from our attribution criteria for the EHR database. Similarly, patients evaluated at the cancer center for a clinical trial may be in the EHR database but not the claims database because commercial insurance is generally not billed for such visits. Finally, we used Washington State death certificates to confirm death, so patients who died elsewhere were not included in this study.

We offer multiple considerations for researchers who want to triangulate data sources to capture multiple aspects of palliative and EOL care. First, as the EHR does not note eligibility for services and is not necessarily complete across all systems where a patient might receive care, studies that use EHR alone are subject to surveillance bias.<sup>24</sup> Second, the EHR is built for clinical care rather than research, so any study using the EHR must include careful consideration of data availability and completeness.<sup>25</sup> Given the significant expense associated with updating EHRs to incorporate algorithms and metrics for research,<sup>26</sup> it is difficult for most health systems to adapt their EHR to better suit research needs. Third, while routinely collected data such as EHRs offer a potentially cost-effective, easily available way to measure palliative and EOL care, having data from one health-care system is not necessarily generalizable to larger populations. Fourth, availability of claims data may be a barrier for researchers, given the significant costs and lag time associated with accessing publicly funded claims-linked databases.<sup>27</sup> However, even with exhaustive amounts of claims data, patients may still switch commercial insurers, Medicare Part D plans, or Medicare Advantage plans, making complete capture of health-care use difficult.<sup>28</sup> Fifth, while the SEER registries are available for research on cancer, registries are not available for all diseases and conditions, which can make verifying a diagnosis difficult if using claims alone. Finally, if budget and time allow, larger population-based databases that use claims plus interviews (e.g., the Health and Retirement Study) should be considered for palliative care and EOL research,<sup>29</sup> to accommodate the large attrition typically seen when applying inclusion/exclusion criteria and linking between smaller databases.

Our findings suggest that while novel database linkages may be a rich source of data to analyze health-

care utilization, verify diagnoses, and confirm clinical outcomes (such as death) while incorporating rich clinical data from the EHR, such linkages can involve substantial loss of potentially eligible patients. When linking between different sources, researchers must consider the biases that linkage entails and the potential effect on sample size. Additionally, we found significant differences between EHR data and claims data in measuring EOL health-care use. Researchers should note that reliance on a single-source EHR will likely miss some health-care use, while reliance on claims will fail to capture the complex clinical details available in EHR that affect palliative care provision and EOL health-care utilization.

### ***Disclosures and Acknowledgments***

**Funding:** This work was supported by the Academy-Health New Investigator Small Grant Award (Cara L. McDermott) and the National Institutes of Health's National Heart, Lung, and Blood Institute Grants T32 HL125195-02 (J. Randall Curtis) and K12 HL137940-02 (J. Randall Curtis).

**Conflict of interest:** Dr. Ruth A. Engelberg reports receiving grants from the Cambia Health Foundation during the conduct of this study.

### ***References***

1. Henson LA, Higginson IJ, Gao W, BuildCare. What factors influence emergency department visits by patients with cancer at the end of life? Analysis of a 124,030 patient cohort. *Palliat Med* 2018;32:426-438.
2. Falchook AD, Dusetzina SB, Tian F, Basak R, Selvam N, Chen RC. Aggressive end-of-life care for metastatic cancer patients younger than age 65 years. *J Natl Cancer Inst* 2017;109:1-6.
3. Stutz M, Kao RL, Huard L, Grotts J, Sanz J, Ross MK. Associations between pediatric palliative care consultation and end-of-life preparation at an academic medical center: a retrospective EHR analysis. *Hosp Pediatr* 2018;8:162-167.
4. Schuler MS, Joyce NR, Huskamp HA, Lamont EB, Hatfield LA. Medicare beneficiaries with advanced lung cancer experience diverse patterns of care from diagnosis to death. *Health Aff (millwood)* 2017;36:1193-1200.
5. McDermott CL, Bansal A, Ramsey SD, Lyman GH, Sullivan SD. Depression and health care utilization at end of life among older adults with advanced non-small-cell lung cancer. *J Pain Symptom Manage* 2018;56:699-708.e1.
6. Nipp RD, Tramontano AC, Kong CY, Hur C. Patterns and predictors of end-of-life care in older patients with pancreatic cancer. *Cancer Med* 2018;7:6401-6410.
7. Dy SM, Herr K, Bernacki RE, et al. Methodological research priorities in palliative care and hospice quality measurement. *J Pain Symptom Manage* 2016;51:155-162.

8. Davies JM, Gao W, Sleeman KE, et al. Using routine data to improve palliative and end of life care. *BMJ Support Palliat Care* 2016;6:257–262.
9. Kassner CT, Bhavsar NA, Harker M, Bull J, Taylor DH Jr. Hospital-based palliative care with medicare claims: evidence from Colorado. *Am J Hosp Palliat Care* 2018;35:66–68.
10. McDermott CL, Fedorenko C, Kreizenbeck K, et al. End-of-Life services among patients with cancer: evidence from cancer registry records linked with commercial health insurance claims. *J Oncol Pract* 2017;13:e889–e899.
11. Earle CC, Neville BA, Landrum MB, et al. Evaluating claims-based indicators of the intensity of end-of-life cancer care. *Int J Qual Health Care* 2005;17:505–509.
12. Earle CC, Landrum MB, Souza JM, Neville BA, Weeks JC, Ayanian JZ. Aggressiveness of cancer care near the end of life: is it a quality-of-care issue? *J Clin Oncol* 2008;26:3860–3866.
13. West SL, Johnson W, Visscher W, Kluckman M, Qin Y, Larsen A. The challenges of linking health insurer claims with electronic medical records. *Health Inform J* 2014;20:22–34.
14. Patorno E, Gopalakrishnan C, Franklin JM, et al. Claims-based studies of oral glucose-lowering medications can achieve balance in critical clinical variables only observed in electronic health records. *Diabetes Obes Metab* 2018;20:974–984.
15. Ehrenstein V, Nielsen H, Pedersen AB, Johnsen SP, Pedersen L. Clinical epidemiology in the era of big data: new opportunities, familiar challenges. *Clin Epidemiol* 2017;9:245–250.
16. Prela CM, Baumgardner GA, Reiber GE, et al. Challenges in merging Medicaid and Medicare databases to obtain healthcare costs for dual-eligible beneficiaries: using diabetes as an example. *Pharmacoeconomics* 2009;27:167–177.
17. Earle CC, Neville BA, Landrum MB, et al. Evaluating claims-based indicators of the intensity of end-of-life cancer care. *Int J Qual Health Care* 2005;17:505–509.
18. Heintzman J, Bailey SR, Hoopes MJ, et al. Agreement of Medicaid claims and electronic health records for assessing preventive care quality among adults. *J Am Med Inform Assoc* 2014;21:720–724.
19. Laws MB, Michaud J, Shield R, McQuade W, Wilson IB. Comparison of electronic health record-based and claims-based diabetes care quality measures: causes of discrepancies. *Health Serv Res* 2018;53(Suppl 1):2988–3006.
20. Savitz ST, Stearns SC, Groves JS, Kucharska-Newton AM, Bengtson LGS, Wruck L. Mind the gap: hospitalizations from multiple sources in a longitudinal study. *Value Health* 2017;20:777–784.
21. Dacey B, Bholat MA. Health information technology: medical record documentation issues in the electronic era. *Prim Care* 2012;39:633–642.
22. Erdem E, Korda H, Haffer SC, Sennett C. Medicare claims data as public use files: a new tool for public health surveillance. *J Public Health Manag Pract* 2014;20:445–452.
23. Iezzoni LI, Heeren T, Foley SM, Daley J, Hughes J, Coffman GA. Chronic conditions and risk of in-hospital death. *Health Serv Res* 1994;29:435–460.
24. Yu TC, Zhou H. Benefits of applying a proxy eligibility period when using electronic health records for outcomes research: a simulation study. *BMC Res Notes* 2015;8:229.
25. Vemulakonda VM, Bush RA, Kahn MG. “Minimally invasive research?” Use of the electronic health record to facilitate research in pediatric urology. *J Pediatr Urol* 2018;14:374–381.
26. Sendak MP, Balu S, Schulman KA. Barriers to achieving economies of scale in analysis of EHR data. A cautionary tale. *Appl Clin Inform* 2017;8:826–831.
27. Doshi JA, Hendrick FB, Graff JS, Stuart BC. Data, data everywhere, but access remains a big issue for researchers: a review of access policies for publicly-funded patient-level health care data in the United States. *EGEMS (Wash DC)* 2016;4:1204.
28. Fung V, Brand RJ, Newhouse JP, Hsu J. Using medicare data for comparative effectiveness research: opportunities and challenges. *Am J Manag Care* 2011;17:488–496.
29. Kelley AS, Langa KM, Smith AK, et al. Leveraging the health and retirement study to advance palliative care research. *J Palliat Med* 2014;17:506–511.