High-Need Phenotypes in Medicare Beneficiaries: Drivers of Variation in Utilization and Outcomes

Tamra Keeney, PhD,[*][†] *Emmanuelle Belanger, PhD,*[*][†] *Rich N. Jones, ScD,*[‡]
Nina R. Joyce, PhD,[*][†] *David J. Meyers, MPH,*[†] *and Vincent Mor, PhD*[*][†]

OBJECTIVES: High-need (HN) Medicare beneficiaries heavily use healthcare services at a high cost. This population is heterogeneous, composed of individuals with varying degrees of medical complexity and healthcare needs. To improve healthcare delivery and decrease costs, it is critical to identify the subpopulations present within this population. We aimed to (1) identify distinct clinical phenotypes present within HN Medicare beneficiaries, and (2) examine differences in outcomes between phenotypes.

DESIGN: Latent class analysis was used to identify phenotypes within a sample of HN fee-for-service (FFS) Medicare beneficiaries aged 65 years and older using Medicare claims and post-acute assessment data.

SETTING: Not applicable.

PARTICIPANTS: Two cross-sectional cohorts were used to identify phenotypes. Cohorts included FFS Medicare beneficiaries aged 65 and older who survived through 2014 (n = 415 659) and 2015 (n = 416 643).

MEASUREMENTS: The following variables were used to identify phenotypes: acute and post-acute care use, functional dependency in one or more activities of daily living, presence of six or more chronic conditions, and complex chronic conditions. Mortality, hospitalizations, healthcare expenditures, and days in the community were compared between phenotypes.

RESULTS: Five phenotypes were identified: (1) comorbid ischemic heart disease with hospitalization and skilled nursing facility use (22% of the HN sample), (2) comorbid ischemic heart disease with home care use (23%), (3) home care use (12%), (4) high comorbidity with hospitalization (32%), and (5) Alzheimer's disease/related dementias with functional dependency and nursing home use (11%). Mortality was highest in phenotypes 1 and 2; hospitalizations and expenditures were highest in phenotypes 1, 3, and 4.

CONCLUSIONS: Our findings represent a first step toward classifying the heterogeneity among HN Medicare beneficiaries. Further work is needed to identify modifiable utilization patterns between phenotypes to improve the value of healthcare provided to these subpopulations. **J Am Geriatr Soc 68:70-77, 2020.**

Key words: health services; outcomes; multimorbidity; latent class analysis

In recent years, substantial effort has been devoted to identifying individuals who are high users of healthcare; the impetus for this work stems from the fact that approximately 5% of the US population accounts for 50% of healthcare costs.[1] Early studies identified the salient characteristics of "high need, high cost" (HNHC) individuals that include various combinations of high medical comorbidity, functional limitations, and age.[2-4] Investigation of administrative claims data has improved our understanding of this HNHC population by providing insight into utilization patterns among the entire population of Medicare beneficiaries. Our research team recently proposed a claims-based definition of high need (HN) that aimed to identify HN Medicare beneficiaries nationally. This HN definition incorporated a complex array of beneficiary characteristics including complexity of chronic conditions, medical comorbidity, overall healthcare utilization, and functional limitations.[5] This innovative definition outperformed previous definitions in its ability to identify large numbers of HN beneficiaries in the entire Medicare population while also being predictive of hospitalization and mortality.[5]

Although recent research has defined the predominant characteristics of the HN population and improved our

global understanding of this population, much remains unexplored. The HN population is incredibly heterogeneous in composition, with significant variation seen between HN individuals with regard to age, medical diagnoses, complexity of comorbidity, and so on. Many of the identifying characteristics that increase the risk of hospitalization and mortality tend to co-occur along specific patterns that remain underexplored. This heterogeneity can complicate attempts to intervene in this population because a one-size-fits-all approach may not work for all HN patients. Improved understanding of the different subgroups that exist within the HN population can be used to explore variation in utilization patterns among the subgroups, some of which may be modifiable.[6,7] Identification of HN subgroups with modifiable utilization patterns represents a powerful tool for health systems, allowing them to better classify and manage complex patients at risk of poor outcomes and to develop targeted intervention programs for the HN population. As a result, improved understanding of the subpopulations present within the HN population was declared a research priority.[7,8]

Scholars have identified subpopulations among high-risk populations including high-cost[9-11] and high-risk[12] cohorts using data from a variety of sources including managed care plans,[10] administrative claims,[11] and health systems.[9,12] Prior work into HN subpopulations classified subgroups using clinical conditions, risk scores, hospital procedures, and acute utilization. Although valuable, past work is limited by a lack of focus on post-acute care that accounts for approximately 73% of the regional variation in Medicare spending.[13] Because most HNHC patients use post-acute care, due to high rates of functional limitations and disability,[3,5] inclusion of post-acute care utilization is important if we are to better understand clinical subgroups. Additionally, subpopulations among high-cost Medicare beneficiaries were qualitatively defined using expert opinion.[2,11] Although these qualitative designations are based on conceptual theory of HNHC characteristics and refined using clinical expertise, an empirically based definition of HN subpopulations is needed.

In this study we expand our HN definition by modeling heterogeneous "phenotypes" in HN Medicare beneficiaries. HN phenotypes were derived using a combination of clinical characteristics, post-acute care utilization, and functional impairment. We aimed to (1) identify these phenotypes present among HN Medicare beneficiaries, and (2) examine variation in the outcomes that the HN phenotypes experience in the following year.

METHODS

We conducted a retrospective study of a national cohort of fee-for-service (FFS) Medicare beneficiaries using administrative data. Latent class analysis (LCA) was used to detect heterogeneity among HN beneficiaries to potentially classify them into clinically meaningful HN phenotypes.

Data Sources

We use data from the Medicare Master Beneficiary Summary File (MBSF), the Chronic Conditions Warehouse (CCW), Cost and Use file, MedPAR hospitalization file,

and post-acute care assessment files. The MBSF was used to capture demographic data; the CCW codes were used to establish the presence of 26 chronic conditions in the study year. The MedPAR and Cost and Use files were used to identify acute and post-acute care stays and healthcare expenditures. Post-acute care assessments included the Minimum Data Set (MDS), the Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI), and the Home Health Outcome and Assessment Information Set (OASIS).

Study Population: High-Need Beneficiaries

We used our claims-based definition to identify HN FFS beneficiaries aged 65 and older in 2014, conditional on survival through the end of 2014. The HN definition is hierarchical, and its predictive validity for hospitalization and mortality was established in previous work.[5] Briefly, we first identified beneficiaries with medical complexity or comorbidities (presence of two or more complex chronic conditions/presence of six or more chronic conditions) and any acute or post-acute healthcare utilization. Then we added any beneficiaries with complete dependency in activities of daily living or mobility who had not already been retrieved.[5] A total of 4 156 594 HN beneficiaries were identified as HN. To improve computing efficiency, we used a random 10% sample from the 2014 national HN cohort (n = 415 659) to identify phenotypes. We then replicated and confirmed phenotypes in a random 10% sample from a 2015 national FFS HN cohort of individuals who survived until the end of 2015 (n = 416 643).

Clinical and Functional Predictors of Phenotypes

To characterize heterogeneity among the HN population, we used variables from our HN definition to examine the how these clinical and functional predictors contribute to phenotype classification: skilled nursing facility (SNF) or nursing home (NH) use, home care service use, one or more hospitalizations, functional limitation, having six or more chronic conditions, and the presence of specific complex chronic conditions (ischemic heart disease, congestive heart failure, atrial fibrillation, Alzheimer's disease/related dementias [ADRD], chronic obstruction pulmonary disease (COPD)/asthma, and chronic kidney disease). Complex chronic conditions were examined separately to identify conditions that have the largest impact on phenotype classification. Use of SNF or NH was indicated by the presence of MDS assessments, and use of home care services was indicated by the presence of OASIS assessments. We characterized functional limitation as dependency in one or more mobility or activities of daily living items (ADLs) obtained in the beneficiary's last post-acute care assessment in 2014.[5]

Outcome Measures

We examined differences between phenotypes assigned in 2014 for the following outcomes in 2015: mortality, hospitalizations, healthcare expenditures per day alive, and percentage of days in the community. Hospitalizations were characterized as the proportion of beneficiaries who were hospitalized as least once in 2015. Healthcare expenditures were calculated as the total Medicare expenditures per days

alive in 2015. We used data from MedPAR, IRF-PAI, MDS, and OASIS files to quantify the total days spent in the community per days alive in 2015.[5]

Statistical Analysis

To ensure that the 10% analytic samples were representative of the overall HN cohort, we calculated descriptive statistics to identify potential differences in demographic and clinical characteristics across cohorts. We then used LCA to identify clinical phenotypes within the HN population. The LCA model is an analytic approach that can be used to identify subgroups present within a heterogeneous population.[14,15] In an LCA model, manifest variables (eg, clinical and functional characteristics) are indicators of the hypothesized underlying, or "latent," categorical variable (eg, distinct HN phenotypes).[16] Response patterns for observed variables are used to classify individuals into mutually exclusive latent classes.[15]

Models were constructed using an iterative approach. All clinical and functional predictor (observed) variables were included in the initial LCA model. The initial models were run sequentially, starting with one class and then estimating models with up to an eight-class solution. We examined a variety of fit indices across the models to determine the optimal class solution. The following indices were assessed together to determine best fit: log likelihood (LL), Bayesian information criterion (BIC), and Lo-Mendell-Rubin likelihood ratio test (LMR-LRT).[17] First, we examined trends in LL and BIC values across classes because the lowest value is used, in part, to identify the final model.[17] We then we compared LMR-LRT results among the values with the lowest BIC. Significant P values from the LMR-LRT test indicate the superiority of the current model compared with the previous simpler model with fewer classes.[15,16]

Finally, we used relative entropy (≥.70) and mean posterior class probability (≥.70) to assess for classification accuracy. Typically, an entropy greater than .70 is an additional marker of how well clinical and functional predictors discriminated between classes. Similarly, mean posterior class probabilities were used to indicate the likelihood of class assignment for each class in the model.[16] As in other latent variable approaches, we first determined the optimal model solution through interpretation of the statistical model fit indices outlined and then refined the solution based on clinical interpretability and judgment. Beneficiaries were assigned to the class they were most likely to belong to, as indicated by posterior membership probabilities.

Conditional item probabilities can be used to indicate the degree of class separation in a sample. When using binary indicator variables, strong separation can be indicated with high (.70) or low (.30) conditional item probabilities.[16] Therefore, we retained indicator variables only if they met these two criteria. To arrive at our final indicator variable set, we used a stepwise procedure to identify the most parsimonious model. After an initial optimal class

Table 1. Demographic Characteristics of High-Need Beneficiary Cohorts

	2014 sample n = 4 156 594		2014 10% sample n = 415 659		2015 10% sample n = 416 643	
	n	%	n	%	n	%
Age, y						
65-74	1 309 330	31.5	131 021	31.5	134 849	32.3
75-84	1 497 267	36.0	149 787	36.0	148 672	35.7
85+	1 349 997	32.5	134 851	32.5	133 122	32.0
Sex						
Female	2 523 115	60.7	252 713	60.8	250 370	60.1
Male	1 633 424	39.3	162 941	39.2	166 273	39.9
Dual eligible	1 278 608	30.8	127 815	30.8	120 328	28.9
Race/Ethnicity						
Other	148 483	3.6	14 634	3.5	15 323	3.7
White	3 500 664	84.2	350 182	84.3	352 660	84.6
Black	417 269	10.0	41 705	10.0	40 478	9.7
Hispanic	90 178	2.2	9 138	2.2	8 182	2.0
SNF or NH utilization	1 505 427	36.2	150 669	36.3	147 985	35.5
Home health utilization	1 968 342	47.4	197 160	47.4	198 978	47.8
Acute care utilization	3 068 777	73.8	307 034	73.9	310 395	74.5
Functional dependence in ADLs	1 587 112	38.2	158 938	38.2	156 123	37.5
6 or more chronic conditions	3 154 782	76.0	315 717	76.0	320 479	77.0
2 or more complex chronic conditions	3 323 828	80.0	332 340	80.0	340 938	82.0
Ischemic heart disease	2 509 247	60.4	251 261	60.5	247 995	59.5
Alzheimer's/Dementia	1 563 158	37.6	155 954	37.5	154 698	37.1
Atrial fibrillation	1 067 631	25.7	106 742	25.7	108 218	26.0
Congestive heart failure	1 921 096	46.2	192 592	46.3	189 052	45.4
Chronic kidney disease	1 936 915	46.6	193 843	46.6	206 455	49.6
COPD/Asthma	1 589 828	38.3	159 215	38.3	163 681	39.3

Abbreviations: ADLs, activities of daily living; COPD, chronic obstructive pulmonary disease; HN, high need; NH, nursing home; SNF, skilled nursing facility.

solution was determined, we dropped observed variables with maximum class-specific conditional item probabilities of less than .70 because this indicates the variable does not discriminate classes well and is not useful for classifying beneficiaries. We then restarted the analysis anew and continued until all indicator variables were strong indicators of class membership.

Following the determination of the optimal number of classes, we named the HN phenotypes based on conditional item probabilities. We examined conditional item probabilities for each class to identify the clinical and functional characteristics that best predicted class assignment (probability ≥.70). We calculated descriptive statistics to identify differences in demographic characteristics in 2014 and outcomes in the subsequent year between HN phenotypes. Mplus v.8 was used for LCA.[18] All other analyses were performed using SAS v.9.4. Access to the data was obtained through a data use agreement (RSCH-2017-51007). This study received approval from the Brown University institutional review board.

RESULTS

Table 1 describes the characteristics of the overall 2014 HN sample and 10% HN samples from 2014 and 2015. Across all cohorts, the sample was predominantly female (61%), white (84%), and had high rates of comorbidity. Most of the population was hospitalized in 2014 (74%) and received subsequent post-acute care services in the home (47%) or in an SNF or NH setting (36.3%).

Based on fit indices and clinical interpretability, we identified a five-class model as the optimal solution (Supplementary Table S1 shows the model fit indices). Observed LL and BIC values decreased monotonically as the number of classes increased. Although the seven-class model had the lowest LL and BIC values and a significant LMR-LRT, this solution had as many variables as classes and was not selected due to potential model over fit. Ultimately, the five-class model was chosen given increased entropy and higher smallest mean class posterior probability when compared with the six-class model.

Table 2. Five-Class Model Conditional Item Probabilities for 10% Sample of Fee-for-Service High-Need Beneficiaries Aged 65 and Older

	Class 1	Class 2	Class 3	Class 4	Class 5
	Comorbid ischemic heart disease with hospitalization and SNF use	Comorbid ischemic heart disease with home care use	Home care use	High comorbidity and hospitalization	ADRD with functional dependency and NH use
2014 cohort (n = 415 659)					
SNF or NH use	.91	.00	.25	.03	*1.00*
Home care use	.63	*.94*	*.75*	.15	.00
Acute care use	.92	.57	.46	*1.00*	.23
Functional dependence	.60	.36	.68	.02	*.73*
Six or more chronic conditions	*.98*	*.98*	.03	*.84*	.51
Ischemic heart disease	*.73*	*.75*	.20	.67	.36
Alzheimer's/Dementia	.51	.37	.37	.14	*.82*
Classification probability	.88	.89	.86	.86	.84
% of HN in class	22	23	12	32	11
n	92 781	93 898	51 455	131 817	45 708
2015 cohort (n = 416 643)					
SNF or NH use	*.88*	.00	.27	.03	*1.00*
Home case use	.65	*.95*	*.72*	.16	.00
Acute care use	.93	.54	.47	*1.00*	.23
Functional dependence	.60	.36	.67	.02	*.73*
Six or more chronic conditions	*.99*	*.96*	.03	*.84*	.55
Ischemic heart disease	*.72*	*.73*	.17	.65	.38
Alzheimer's/Dementia	.51	.37	.38	.13	*.83*
Classification probability	.85	.89	.84	.86	.84
% of HN in class	22	23	12	33	10
n	92 525	95 365	49 422	135 505	43 826

Note. Italicized values indicate conditional item probabilities greater than or equal to 0.70. Abbreviations: ADRD, Alzheimer's disease/related dementias; HN, high need; SNF, skilled nursing facility.
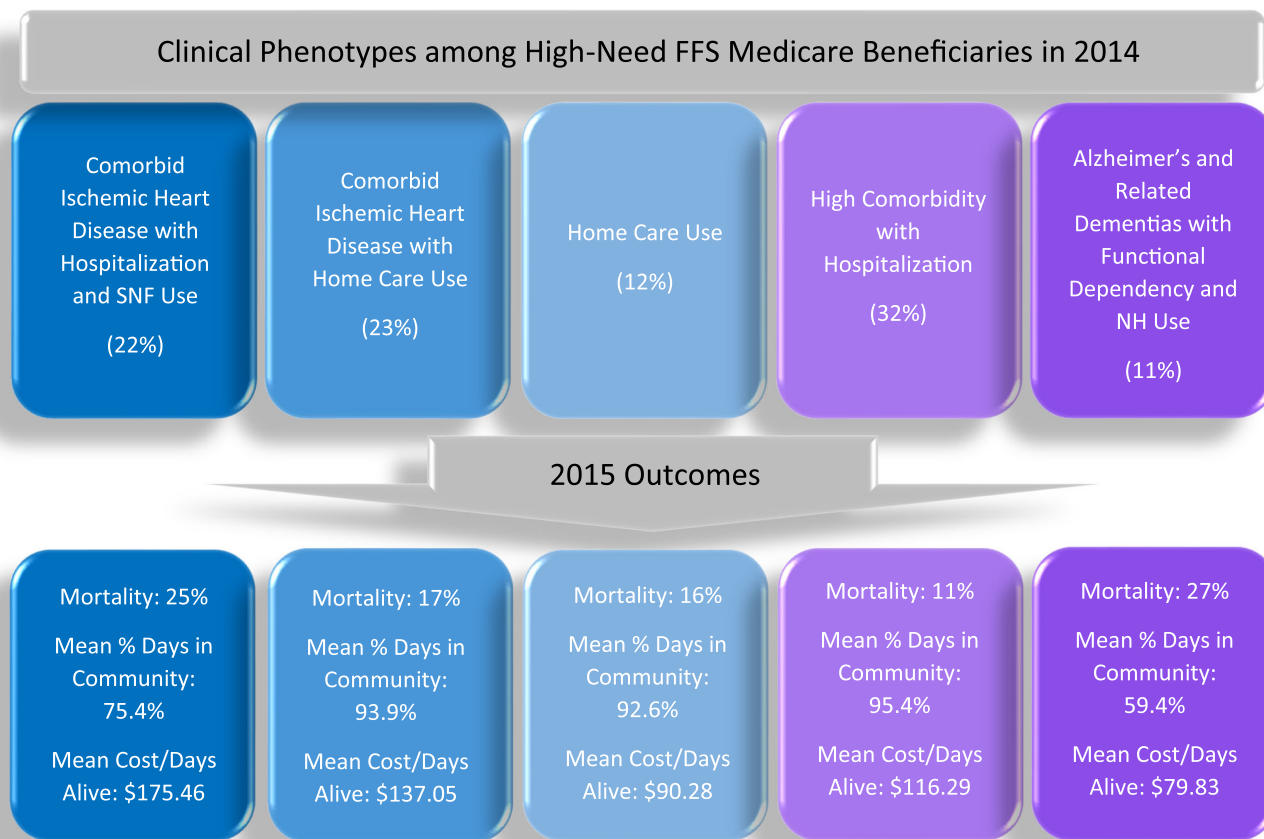
Clinical Phenotypes among High-Need FFS Medicare Beneficiaries in 2014

| Comorbid Ischemic Heart Disease with Hospitalization and SNF Use (22%) | Comorbid Ischemic Heart Disease with Home Care Use (23%) | Home Care Use (12%) | High Comorbidity with Hospitalization (32%) | Alzheimer's and Related Dementias with Functional Dependency and NH Use (11%) |

2015 Outcomes

| Mortality: 25% Mean % Days in Community: 75.4% Mean Cost/Days Alive: $175.46 | Mortality: 17% Mean % Days in Community: 93.9% Mean Cost/Days Alive: $137.05 | Mortality: 16% Mean % Days in Community: 92.6% Mean Cost/Days Alive: $90.28 | Mortality: 11% Mean % Days in Community: 95.4% Mean Cost/Days Alive: $116.29 | Mortality: 27% Mean % Days in Community: 59.4% Mean Cost/Days Alive: $79.83 |

Figure 1. Clinical phenotypes and outcome variation among high-need fee-for-service (FFS) Medicare beneficiaries. NH, nursing home.

Table 3. Demographic Characteristics by High-Need Phenotype in 2014

| | Class 1 | Class 2 | Class 3 | Class 4 | Class 5 |
	Comorbid ischemic heart disease with hospitalization and SNF use (n = 92 781)	Comorbid ischemic heart disease with home care use (n = 93 898)	Home care use (n = 51 455)	High comorbidity with hospitalization (n = 131 817)	ADRD with functional dependency and NH use (n = 45 708)
Age, y, n (%)					
65-74	21 805 (24)	26 710 (28)	15 049 (29)	59 528 (45)	7929 (17)
75-84	34 429 (37)	36 424 (39)	16,516 (32)	48,915 (37)	13 503 (30)
85+	36 547 (39)	30764 (33)	19 890 (39)	23 374 (18)	24 276 (53)
Sex, n (%)					
Female	59 146 (64)	56 651 (60)	35 772 (69)	67 876 (51)	33 268 (73)
Male	33 634 (36)	37 246 (40)	15 683 (31)	63 939 (49)	12 439 (27)
Dual eligible, n (%)	32 143 (35)	27 745 (30)	12 645 (25)	20 811 (16)	34 471 (75)
Race/Ethnicity, n (%)					
White	79 868 (86)	74 966 (80)	42 814 (83)	114 220 (87)	38 314 (84)
Black	8868 (9)	11 214 (12)	5547 (11)	10 928 (8)	5148 (11)
Hispanic	1482 (2)	3764 (4)	1226 (2)	1797 (1)	869 (2)
Other	2563 (3)	3954 (4)	1868 (4)	4872 (4)	1377 (3)

Abbreviations: ADRD, Alzheimer's disease/related dementias; NH, nursing home.

After initial modeling, clinical and functional characteristics that discriminated well among the identified classes were SNF or NH use, home health care use, hospitalization, functional dependency, having six or more chronic conditions, ischemic heart disease, and ADRD. Variables that did not discriminate well among the identified classes were congestive heart failure, atrial fibrillation, chronic obstruction pulmonary disease/asthma, and chronic kidney disease. Based on

Table 4. Outcomes in 2015 by High-Need Phenotype in 2014

	Class 1	Class 2	Class 3	Class 4	Class 5
	Comorbid ischemic heart disease with hospitalization and SNF use (n = 92 781)	Comorbid ischemic heart disease with home care use (n = 93 898)	Home care use (n = 51 455)	High comorbidity with hospitalization (n = 131 817)	ADRD with functional dependency and NH use (n = 45 708)
Mortality, n (%)					
Death in 2015	22 718 (25)	15 636 (17)	8314 (16)	14 555 (11)	12 473 (27)
Hospitalization, n (%)					
Any hospitalization in 2015	46 132 (50)	42 817 (46)	18 232 (35)	54 401 (41)	11 501 (25)
% days in the community per days alive, mean (SD)	75.4 (37.5)	93.9 (15.7)	92.6 (19.8)	95.4 (14.8)	59.4 (47.3)
Cost per days alive, mean (SD)	175.46 (630.00)	137.05 (539.20)	90.28 (234.48)	116.29 (765.06)	79.83 (181.20)

Abbreviations: ADRD, Alzheimer's disease/related dementias; NH, nursing home; SD, standard deviation; SNF, skilled nursing facility.

conditional item probabilities, we named the classes (Table 2 and Figure 1): (1) comorbid ischemic heart disease with hospitalization and SNF use (22% of HN sample), (2) comorbid ischemic heart disease with home care use (23%), (3) home care use (12%), (4) high comorbidity with hospitalization (32%), and (5) ADRD with functional dependency and NH use (11%). Replication in the 2015 validation cohort demonstrated good reproducibility of HN phenotypes. A five-class solution emerged with similar results seen between classes with respect to all indices of model fit (Supplementary Table S1) and conditional item probabilities values within .05 or less compared with those in 2014 (Table 2).

Demographic characteristics across HN phenotypes varied significantly (Table 3). The ADRD with functional dependency and NH use phenotype had the largest proportion of beneficiaries who were dual eligible (75%), aged 85 years and older (53%), and female (73%). High comorbidity with hospitalization was predominantly composed of beneficiaries 65 to 74 years of age (45%) and had the lowest proportion of dual-eligible individuals (16%).

Patient-centered outcomes in 2015 were also found to vary significantly by 2014 HN phenotype (Table 4). Mortality rates were highest among the ADRD phenotype (27%) and the comorbid ischemic heart disease with hospitalization and SNF use phenotype (25%). The proportion of beneficiaries hospitalized at least once in 2015 was highest among individuals in the comorbid ischemic heart disease with hospitalization and SNF use (50%), comorbid ischemic disease with home care use (46%), and high comorbidity with hospitalization (41%). These phenotypes were also found to have the highest cost per days alive in 2015, with mean costs ranging from $116.29/day to $175.46/day. Beneficiaries in the home care use, comorbid ischemic heart disease with home care use, and high comorbidity with hospitalization phenotypes experienced the highest percentage of days spent in the community, with 92% of days in 2015 or greater.

DISCUSSION

This investigation revealed that post-acute care utilization, functional dependency, and medical comorbidity can be used to categorize the heterogeneity present within HN FFS Medicare beneficiaries. Findings from our analyses were robust as demonstrated by replicability of LCA results using 2015 data. Our work improves on previous literature by identifying phenotypes among a representative sample of HN FFS beneficiaries that may provide more actionable insights than investigating the top decile of cost or utilizers alone.[2,11] This is an important distinction when identifying areas in which health systems could reduce unnecessary utilization because targeting a small group of high-cost individuals may not accurately reflect the true amount of wasteful utilization at the health system level.[19] Therefore, systems-focused approaches based on aggregate data may be more effective in decreasing wasteful utilization for all patients, resulting in greater overall savings.[19] In addition to focusing on populations with high healthcare expenditures, previous studies relied on expert panels to guide the segmentation of high-cost subpopulations.[2,11] Although these panels used a qualitative approach to begin exploring heterogeneity among high-cost beneficiaries, our results contribute to the identification of empirically derived subpopulations among HN beneficiaries. We expand on the previous qualitative definition by demonstrating the importance of functional limitation and post-acute care utilization in HN phenotype classification. Thus the analyses presented in this article provide an empirical approach for identifying subpopulations ("phenotypes") among HN beneficiaries based on acute and post-acute care utilization and other clinical characteristics.

Programs aimed at improving utilization and outcomes for HNHC individuals have traditionally focused on patient-level intervention strategies but overall have failed to yield significant benefits and cost reductions.[8] Many of these early programs proposed to improve outcomes for the HNHC population by identifying at-risk patient populations (eg, individuals at risk for hospitalizations, readmissions, etc). Although these programs were created using clinically informed designations of HNHC, evidence suggests that the absence of system-level efficacy in reducing utilization and costs for these programs is partly related to a lack of empirical classification of HNHC cohorts. In fact,

many researchers have acknowledged the need to identify subgroups that go beyond individual medical diagnoses within the HNHC population if health systems are to adopt interventions targeted to the specific needs for each subgroup with improved accuracy.[7,8]

Our investigation provides a method for identifying phenotypes among HN Medicare FFS beneficiaries using administrative claims data. Healthcare utilization and cost varied among the five phenotypes, illustrating that different patterns of utilization exist among HN older adults. Additionally, demographic characteristics and outcomes varied by phenotype, further supporting the conclusion that phenotypes will require different interventions to improve value. Importantly, health systems have access to administrative claims data that would allow providers to leverage these phenotypes within their system to identify the HN subgroups within their organization. Health systems and clinicians may then be more empowered to tailor interventions toward HN phenotypes with modifiable utilization patterns, thereby improving costs and maximizing impact.

Limitations to this study warrant discussion. Data for this investigation came from administrative claims that are collected for billing purposes. Therefore, claims data are unable to reflect measures of clinical severity and have limited ability to account for complexity of conditions, all of which could impact phenotype classification. As a result, our analysis may have been limited with regard to identifying more granular contributions of conditions to HN phenotype classification because chronic conditions tend to cluster together in older adults and are crudely represented in claims. Our definition of HN and HN phenotypes relies on claims-based measures of use of acute and post-acute care. This makes it difficult to differentiate whether HN individuals by our definition are actually HN or are perhaps captured as such because they are lacking proper medical management and therefore are classified as HN as a result of inappropriate utilization.

Additionally, there may be a population of older adults with functional limitations who are unable to be classified as HN using our definition, due to lack of home health or NH assessments. Although these individuals are not high utilizers of healthcare, additional primary care based data may be useful for identifying community-dwelling older adults who may be HN based on the criteria of functional limitations and who could be better served to prevent future increases in utilization and expenditures. When identifying our HN cohort for this investigation, 2014 decedents were excluded from analyses. A subset of those decedents may have potentially been classified as HN in 2014 and, if analyzed, could have led to variation in phenotype classification in our analyses.

Two additional limitations related to our use of LCA include identification of future outcomes based on latent class membership (known as distal outcomes) and the assumption of conditional independence. Concern is growing regarding the potential for overestimation of associations between latent class membership and distal outcomes. As a result, one- and three-step methods for estimating distal outcomes in LCA were developed.[20] However, these methods were tested using simulation data with the acknowledgment that practical application of these methods can lead to model convergence issues and potential changes in classification of persons into latent groups.[20,21] Therefore, in this study, we descriptively compared outcomes across HN phenotypes rather than using regression analysis, which can result in biased estimates of outcome effect sizes and standard errors.[21]

Moreover, the assumption of conditional independence stipulates that predictor variables within a latent class should be independent of one another, or uncorrelated. However, this assumption was difficult to test formally in our data because statistical tests of association are systematically significant in large samples, regardless of effect size. Conditional dependence between predictor variables may lead to the inclusion of one more class than is necessary in the final model.[22] Although we defined model fit criteria a priori and conducted a variety of sensitivity analyses (eg, comparing subsequent models with differing combinations of indicator variables including leaving in or taking out variables thought to be closely related), our latent class groupings may be impacted by conditional dependence.

In conclusion, this study identifies the presence of five distinct phenotypes among HN Medicare beneficiaries using administrative claims data. Phenotypes varied with regard to demographic characteristics and outcomes including mortality, hospitalizations, cost, and days spent in the community. Future studies should continue to refine phenotype classification and to examine outcomes over time in a variety of health systems. Identification of potential regional or market-level differences in outcomes and utilization among phenotypes is also needed.

ACKNOWLEDGMENTS

REFERENCES

1. Cohen SB. The concentration of health care expenditures and related expenses for costly medical conditions, 2012. Rockville, MD: Agency for Healthcare Research and Quality; 2001. Statistical Brief 455. https://www.ncbi.nlm.nih.gov/books/NBK470837/. Accessed October 20, 2018.

2. Joynt KE, Figueroa JF, Beaulieu N, Wild RC, Orav EJ, Jha AK. Segmenting high-cost Medicare patients into potentially actionable cohorts. Healthc (Amst). 2017;5(1-2):62-67.

3. Hayes SL, Salzberg CA, McCarthy D, et al. High-need, high-cost patients: who are they and how do they use health care? Issue Brief (Commonw Fund). 2016;26(1897):14.

4. Colla CH, Lewis VA, Kao LS, O'Malley AJ, Chang CH, Fisher ES. Association between Medicare accountable care organization implementation and spending among clinically vulnerable beneficiaries. JAMA Intern Med. 2016; 176(8):1167-1175.

5. Belanger E, Silver B, Meyers DJ, et al. A retrospective study of administrative data to identify high-need Medicare beneficiaries at risk of dying and being hospitalized. J Gen Intern Med. 2019;34(3):405-411.

6. Blumenthal D, Chernof B, Fulmer T, Lumpkin J, Selberg J. Caring for high-need, high-cost patients—an urgent priority. N Engl J Med. 2016;375(10):909-911.

7. Blumenthal DM, Anderson G, Burke SP, Fulmer T, Jha AK, Long P. Tailoring complex-care management, coordination, and integration for high-need, high-cost patients: a vital direction for health and health care. NAM Perspective. Discussion paper. Washington, DC: National Academy of Medicine; 2016 doi:10.31478/201609q.

8. Anderson GF, Ballreich J, Bleich S, et al. Attributes common to programs that successfully treat high-need, high-cost individuals. Am J Manag Care. 2015;21(11):e597-e600.

9. Davis AC, Shen E, Shah NR, et al. Segmentation of high-cost adults in an integrated healthcare system based on empirical clustering of acute and chronic conditions. J Gen Intern Med. 2018;33(12):2171-2179.

10. Powers BW, Yan J, Zhu J, et al. Subgroups of high-cost Medicare advantage patients: an observational study. J Gen Intern Med. 2019;34(2):218-225.

11. Figueroa JF, Joynt Maddox KE, Beaulieu N, Wild RC, Jha AK. Concentration of potentially preventable spending among high-cost Medicare subpopulations: an observational study. Ann Intern Med. 2017;167(10):706-713.

12. Wong ES, Yoon J, Piegari RI, Rosland AM, Fihn SD, Chang ET. Identifying latent subgroups of high-risk patients using risk score trajectories. J Gen Intern Med. 2018;33(12):2120-2126.

13. Institute of Medicine. Variation in Health Care Spending: Target Decision Making, Not Geography. Washington, DC: National Academies Press; 2013.

14. Lanza ST, Rhoades BL. Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. Prev Sci. 2013;14(2):157-168.

15. Collins LM, Lanza ST. Latent Class and Latent Transition Analysis: With Applications in the Social Behavioral, and Health Sciences. Hoboken, NJ: Wiley; 2010.

16. Masyn KE. Latent class analysis and finite mixture modeling. In: Little TD, ed. The Oxford Handbook of Quantitative Methods in Psychology. Vol 2. UK, Oxford: Oxford Handbooks Online; 2013:1-120.

17. Nylund KL, Asparaoutiov T, Muthen BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Struct Equ Model Multidiscip J. 2007;14(4):535-569.

18. Muthén LK, Muthén BO. Mplus User's Guide. 8th ed. Muthén & Muthén: Los Angeles, CA; 2017.

19. McWilliams JM, Schwartz AL. Focusing on high-cost patients—the key to addressing high costs? N Engl J Med. 2017;376(9):807-809.

20. Asparouhov T, Muthen B. Auxiliary variables in mixture modeling: three-step approaches using Mplus. Struct Equ Model Multidiscip J. 2014;21(3):329-341.

21. Kamata A, Kara Y, Patarapichayatham C, Lan P. Evaluation of analysis approaches for latent class analysis with auxiliary linear growth model. Front Psychol. 2018;9:130.

22. Uebersax JS. Probit latent class analysis with dichotomous or ordered category measures: conditional independence/dependence models. Appl Psychol Measur. 1999;23(4):283-297.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Supplementary Table S1: Model Fit Indices for Model with MDS, Home, Acute Care Use, Functional Dependency, Six or More Chronic Conditions, Ischemic Heart Disease, and Alzheimer's Dementia