

Hierarchical Cluster Analysis of Service Usage and Geographic Variation in Medicare Spending

Date
April 2013

Prepared by
Linnet Vacha, Jack Schryver, Mallikarjun Shankar and Edmon Begoli

DOCUMENT AVAILABILITY

Reports produced after January 1, 1996, are generally available free via the U.S. Department of Energy (DOE) Information Bridge.

Web site <http://www.osti.gov/bridge>

Reports produced before January 1, 1996, may be purchased by members of the public from the following source.

National Technical Information Service

5285 Port Royal Road

Springfield, VA 22161

Telephone 703-605-6000 (1-800-553-6847)

TDD 703-487-4639

Fax 703-605-6900

E-mail info@ntis.gov

Web site <http://www.ntis.gov/support/ordernowabout.htm>

Reports are available to DOE employees, DOE contractors, Energy Technology Data Exchange (ETDE) representatives, and International Nuclear Information System (INIS) representatives from the following source.

Office of Scientific and Technical Information

P.O. Box 62

Oak Ridge, TN 37831

Telephone 865-576-8401

Fax 865-576-5728

E-mail reports@osti.gov

Web site <http://www.osti.gov/contact.html>

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Computational Sciences and Engineering Division

**HIERARCHICAL CLUSTER ANALYSIS OF SERVICE USAGE AND
GEOGRAPHIC VARIATION IN MEDICARE SPENDING**

Linnet Vacha, Jack Schryver, Mallikarjun Shankar and Edmon Begoli

Date Published: April, 2013

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831-6283
managed by
UT-BATTELLE, LLC
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

	Page
CONTENTS	iii
LIST OF FIGURES	v
ACKNOWLEDGMENTS	vii
ABSTRACT	1
1. INTRODUCTION	1
2. ANALYSIS	3
2.1 DATA SOURCE	3
2.2 DATA	3
2.3 DATA TRANSFORMATION	3
2.4 HIERARCHICAL CLUSTER ANALYSIS	4
2.5 LINKAGE METHODS	6
3. CLUSTER VALIDATION	6
3.1 INTERNAL METHODS	6
3.2 EXTERNAL METHODS	8
3.2.1 Comparison to Cost	8
3.2.2 Geospatial Autocorrelation	8
4. CLUSTER PROFILE CONSTRUCTION AND COMPARISON	10
5. CONCLUSION AND FUTURE WORK	10
5.1 SUMMARY OF FINDINGS	10
5.2 POSSIBLE EXPLANATIONS	12
5.3 OUTSTANDING QUESTIONS	14
6. REFERENCES	15

LIST OF FIGURES

Figure		Page
1	Standardized risk-adjusted per capita costs for 2007	2
2	Projected federal spending on Medicare and Medicaid as a percentage of GDP	2
3	Service usage factors	4
4	Graph of the logit function.....	5
5	Clustering dendrogram	6
6	Elbow plots of dendrogram height against number of clusters.....	7
7	Map of clusters for 2007	7
8	Boxplot of 2007 clusters versus standardized risk-adjusted cost.....	8
9	Clustering and comparison to cost of non-transformed 2007 data	9
10	Histogram of service use by cluster	11
11	Degree of variation of service use	12
12	Histograms of high-variation services by clusters	13

ACKNOWLEDGMENTS

This research was supported by the U. S. Department of Energy Science Undergraduate Laboratory Internship program.

ABSTRACT

Per capita Medicare spending varies considerably throughout the United States, even after taking into account factors such as cost of living and pre-existing conditions of the Medicare populations. Additionally, per capita spending may vary up to four times as much under Medicare as it does under private insurance. Several studies have explored the Medicare-costs differences between regions with respect to how Medicare services are used. The recent public release of a large, relatively fine-grained set of Medicare data over four annual periods (2007–2010) by the Centers for Medicare and Medicaid Services (CMS), allowed us to perform a broad-scale exploratory analysis of Medicare service-usage patterns. The CMS data set contains data on the Health Care Referral Region level, which subsets the United States into 306 distinct regions. To identify potential trends in service usage over these regions, we used hierarchical agglomerative cluster analysis (HACA) to group regions into clusters based on similarity of service-usage patterns. Much previous work on the subject of cost variation starts with analysis of cost. By beginning with service-usage data, HACA effectively uncoupled patterns of service-usage from overall patterns of cost. We found that while some Medicare services were used in a fairly consistent manner over all clusters, other services varied considerably. Variation in how frequently services are used may provide insight into practices that are particularly expensive or particularly cost effective. It also may be possible to develop recommendations for more cost-efficient high-quality medical care by taking such information into account. In addition, the variance among the frequencies with which some services are used may point to inconsistent standards of medical practice that are not necessarily directly reflected in total cost, or to more deeply underlying geographic or demographic factors that are not necessarily reflected directly in total cost.

1. INTRODUCTION

Previous research has found significant geographic variation in Medicare spending.¹ Even when facts such as cost of living and the general health of the local population are standardized, an “unwarranted” amount of variation remains² as illustrated in Fig. 1. The degree of cost variation seems to some degree to be unique to Medicare, as regional Medicare spending does not correlate very well to general healthcare spending in the same regions.¹ Figure 2 shows the projected growth in Medicare and Medicaid costs, much of which is driven by the inflation of medical costs themselves. As such, bringing down the cost of medical care is an area of great interest. The hope is that by identifying practices that lead to particularly high or low costs, trends that areas with high costs can be brought more into line with the national average, while areas with particularly low costs may serve as a model that other areas can emulate.

Work by the Dartmouth Atlas Project, among others, has suggested that regional differences in service-utilization are drivers for cost variation and that some of the variations in medical practice and service usage are themselves “unwarranted”.⁴ Therefore, we targeted service usage patterns as a potential further explanation of geographic variation in cost. Rather than starting with an analysis of cost variance and seeking a link between that variance and service usage, we began an analysis of service usage itself. We elected to use hierarchical clustering of Health Care Referral Regions (HRRs) with respect to the proportion on an HRR’s population that used each service as our primary angle of attack. In essence, this method let the data on service use speak for itself, regardless of whether or not it was directly linked to geographic variation in cost. Furthermore, comparing two HRRs directly is nearly impossible due to the myriad of differences between any two regions of the country. Clustering the HRRs into groups based on their overall patterns of service usage let us identify eight general categories of HRRs that could then be compared.

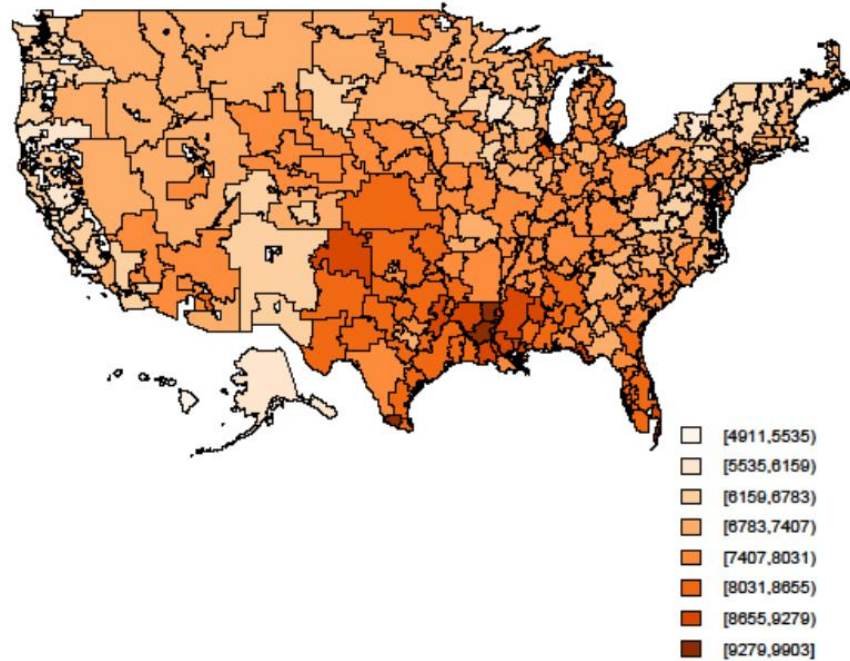


Fig. 1 Standardized risk-adjusted per capita costs for 2007. Medicare costs vary distinctly throughout the United States, even when they are adjusted based on measures of risk within the HRR population. In 2007, standardized risk-adjusted costs ranged from \$4911 in Honolulu, HI to \$9903 in Miami, FL. The regions represented in the map are Health Care Referral regions.

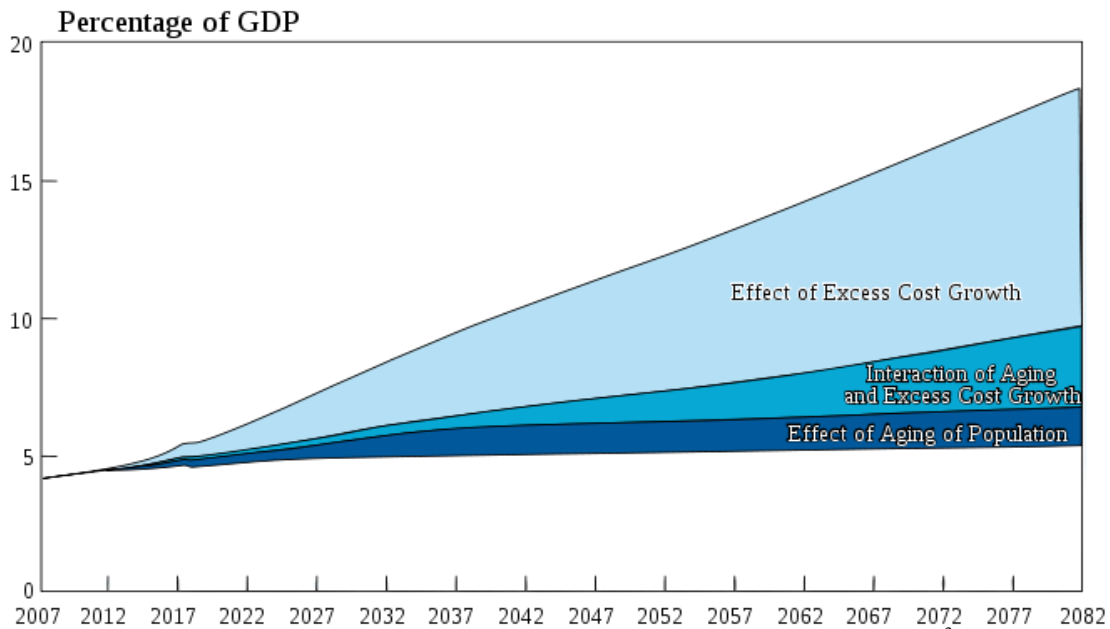


Fig 2. Projected federal spending on Medicare and Medicaid as a percentage of GDP.³ The percentage of the national GDP spent on Medicare and Medicaid is projected to grow dramatically over the coming years. The majority of that cost growth will be driven by disproportionate inflation in the cost of health care itself.

We found that while some services are used in a consistent manner over all clusters, the proportions of clusters' populations using other services varies considerably. In general, services used by a large proportion of the population are fairly consistent between clusters, while services used by a low

proportion of the population vary, sometimes considerably. Contiguous HRRs are more likely to be members of the same cluster than we would expect under a random assignment of cluster membership. Some pairs of clusters show a tendency to appear together geographically, while others show a lower than expected rate of contiguity.

Although we performed the analysis for all four years covered in the CMS data set, we have not yet developed a methodology for precise temporal comparison of clusters. Thus we focus on our analyses of the data from 2007 in the discussion of our methods and results.

2. ANALYSIS

2.1 DATA SOURCE

In response to a January 2011 request by the Institute of Medicine and Committee on Geographic Variation in Health Care Spending and Promotion of High-Value Care, the Centers for Medicare and Medicaid Services (CMS) released a set of Medicare data.⁵ The data set covers the years 2007 through 2010. Data sets are available on both the state and HRR level.

Looking at data on the HRR level is useful because American use of health care resources is generally localized. The use of HRRs attempts to reflect this behavior. The United States is subset into 306 HRRs in such a way that even if a patient is referred to another facility for treatment as severe as cardiac or brain surgery, they are likely to remain in the same HRR.¹ As illustrated in Fig. 1, the size of HRRs varies throughout the country and corresponds roughly with the population density of an area, so that an urban center will have a smaller geographic area than a sparsely populated rural area. That said, as of 2010 the total beneficiaries per HRR ranged from 13,885 in Johnstown, PA to 476,413 in Los Angeles, CA. Data on the HRR level contains much more geographic specificity than if it were on a state level and may reflect trends of rural vs. urban or suburban location that would be lost on a state level.

2.2 DATA

The CMS data set contains information on spending, average calculated risk scores of the population, beneficiary demographic information, care quality indicators, and service utilization. For the purposes of our work, we focused on the service utilization factors. The data set contains 22 service usage factors, as listed in Fig. 3. The data set contains information on the proportion of an HRR's population that uses a service each year, the number of times a service is utilized each year, and annual per capita spending per service. We elected to use the proportion of an HRR's population that utilizes a service as our measure of service use.

The data sets contain per capita cost information in both risk-adjusted and non-risk adjusted form. The risk-adjusted cost attempts to take into account the effect of the overall health of the population on Medicare costs. Because we are interested in the effects of service use on costs, we elected to focus on risk-adjusted costs in our analysis. Admittedly, the health of a population does affect what services they will use, but the risk-adjusted costs free our analysis somewhat from this factor.

2.3 DATA TRANSFORMATION

The proportion of an HRR's population using services varies greatly between services. For example, in the Colorado Springs HRR, 0.27% of the population used Critical Access Hospitals in 2007, while in the same year 88 percent of the population used Evaluation and Management Services. This difference in the scale of usage between factors presents some challenges for analysis. Namely, while a difference of 0.5 percentage points would not be particularly meaningful in the case of Evaluation and Management service usage, that same increase would represent a near tripling in the number of beneficiaries using Critical Access Hospitals.

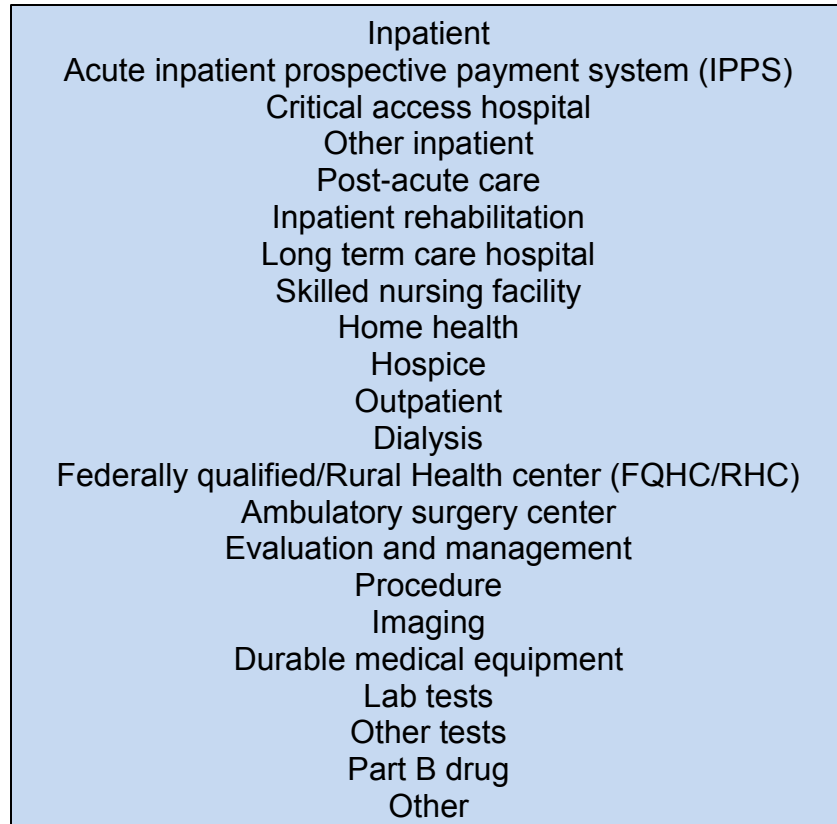


Fig. 3. Service usage factors. The CMS data set breaks health services used by the Medicare population into 22 categories. We focused our analysis on the proportion of the population in each HRR that uses these services.

A similar issue exists for values very close to one, although since none of the services are used by that high a proportion of the population, the issue is of theoretical interest. Depending on the actual values of the percentage of the population that utilize a service, a small numerical difference in that percent may or may not be noteworthy.

To address this issue, we applied the logit function to the data. As illustrated in Fig. 4, the logit function has the effect of essentially “stretching out” values near zero and one, while leaving the middle values relatively unchanged. This solves the problem of the context of small changes by exaggerating small for very low and very high proportions.

The logit of a value P , in which P may be thought of as a proportion which must be between 0 and 1, is given by $\text{logit}(P) = \log(P)/(1-P)$. In keeping with convention, we used the natural logarithm. Note that the logit function may not be applied to values of exactly 0 or 1, as this would result in either taking the logarithm of 0 or dividing by zero, both of which are impossible. For this reason, the data must be scaled before the logit function may be applied to it. To do this, we define a scaled proportion, P_0 , where $P_0 = (F + (1/2))/(N + 1)$. In this equation, F represents the number of beneficiaries in an HRR who have used a particular service, while N represents the total number of beneficiaries in the HRR.

2.4 HIERARCHICAL CLUSTER ANALYSIS

We now find ourselves confronted with a data set of 306 HRRs and 22 service use factors. One of the challenges of working with such a large data set is that it becomes difficult to identify trends over so many variables. For this reason, we elected to use hierarchical agglomerative clustering analysis.

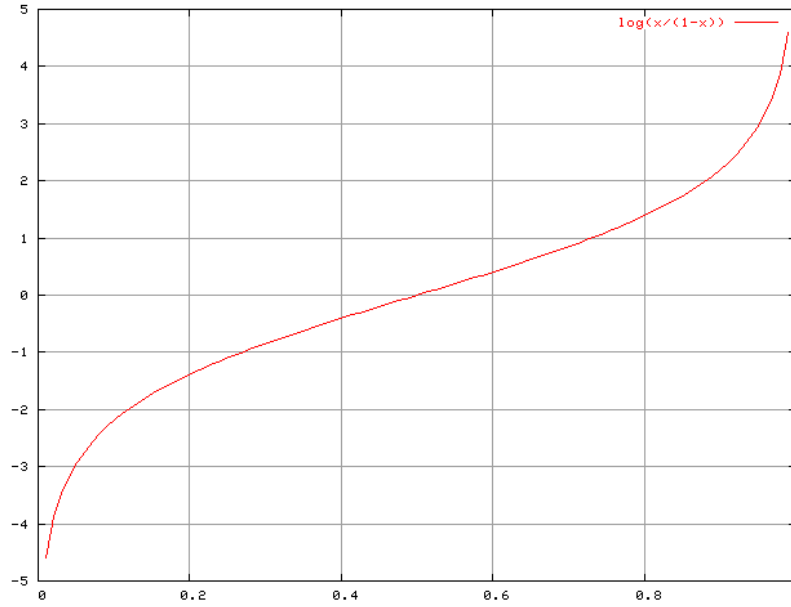


Fig. 4. Graph of the logit function.⁶ The logit function stretches out values near zero and one while leaving more central values relatively unaffected.

Clustering assigns objects to groups, or clusters, so that objects within a group are more similar to each other than to objects in other groups. Hierarchical clustering assumes that objects that are closer together are more similar than objects that are far apart. In our case, each HRR is considered an object and similarity, or distance from other objects, is determined by the patterns of service usage within the HRR over all service use factors. By clustering HRRs based on service use, we can look within the clusters to identify patterns in service usage.

Hierarchical clustering may be either agglomerative or divisive. In essence, agglomerative clustering begins by placing each object in its own cluster and then linking these clusters together based on similarity. Divisive clustering starts by placing all objects in a single cluster and separating them out based on dissimilarity. While both are calculation intensive, the agglomerative method is the less intensive of the two, and it is the one we elected to use. The actual process of clustering is represented in Figure 5. Initially, all the objects are considered to belong to their own cluster. The distance between all objects is then calculated, and the two objects that are closest together are grouped together in a cluster. This cluster is now considered its own object and its distance from the other objects is calculated. This process is then repeated until all the objects have been grouped into a single cluster. The open source statistical programming language R has a function “hclust,” which we used to perform our clustering analysis.

In a dendrogram of clustering, such as the one illustrated in Fig. 5, the y-axis can be thought of as a measure of variance of individual objects within a cluster. Thus, when each object is in its own cluster, the measure of variance within clusters is zero. Furthermore, the measure of variance within clusters will always increase or perhaps remain the same when two objects or clusters are merged together. For our purposes, we use the Euclidean measure of distance in all cases. What is less clear, however, is how to determine the distance between two clusters of objects. We refer to these measures as linkages, since they determine how clusters will be linked together.

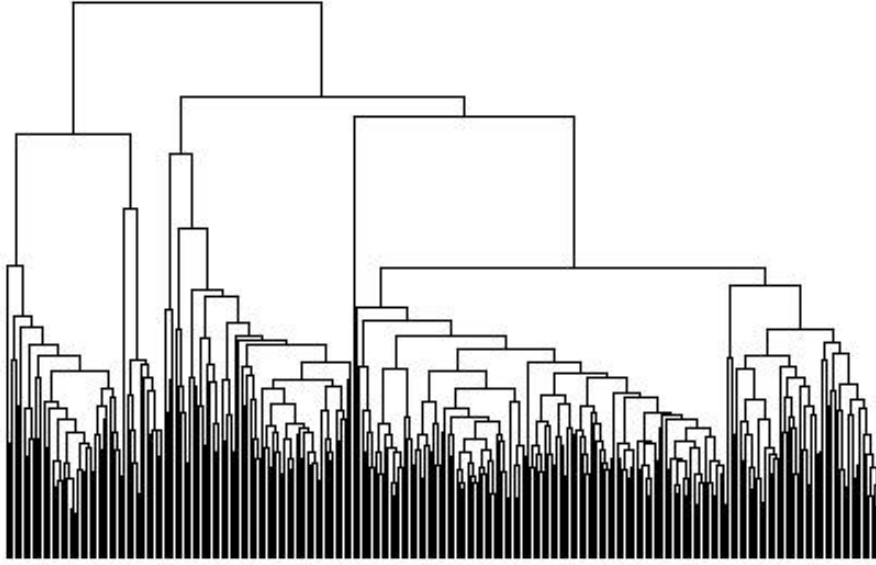


Fig. 5. Clustering Dendrogram. An agglomerative hierarchical clustering algorithm begins with all objects in separate clusters and links them together based on overall similarity until all objects have been grouped into a single cluster. Each time a new cluster is formed, its distance from other clusters is recalculated. The height of the dendrogram approximates the variance within clusters. When each object is in its own cluster, the internal variance is zero. As objects are clustered together, the internal variance increases. In our analysis, HRRs are considered to be individual objects, and their distance from one another is decided by the proportion of their population that uses each service.

2.5 LINKAGE METHODS

The function `hclust` that we used to cluster the data contains several linkage methods. For our analysis, we considered three different methods: single, complete, and average. Under the single linkage method, the distance between two clusters is taken to be the distance between the two closest objects in those clusters. Complete linkage can be thought of as the opposite of this method. In complete linkage, the distance between clusters is taken to be the maximum distance between the two objects. Finally, we considered average linkage, in which as its name suggests takes the average of the location of objects of clusters as the location from which distance will be calculated.

Single linkage tended to result in a single large cluster with a “tail” of small clusters that often consisted of single objects. For our purposes, this was not a useful result, and so we decided against using the single linkage function for our final analysis. Between complete and average linkage, average linkage produced more helpful results and is more intuitive. Thus we selected average linkage as the linkage function for all further analysis.

3. CLUSTER VALIDATION

3.1 INTERNAL METHODS

Applying a clustering algorithm to a set of data, as you can see in Fig. 5, gives no indication of how many clusters we should actually use. In fact, selecting how many clusters to use is one of the real challenges of clustering analysis. As previously noted, the y-axis of a dendrogram can be thought of as the variance within clusters. If we want to consider clusters to be unique and to indicate trends over all their members, we don't want too few clusters as this will result in over-generalizing. On the other hand,

we don't want too many clusters, either, since reducing the number of objects to consider was part of our motivation for performing cluster analysis in the first place.

To find a balance between reducing the number of clusters and keeping variance within clusters low, we use a method known as “elbow analysis” to look at how the distance within clusters changes given the total number of clusters used. As seen in Fig. 6, we simply plot dendrogram height against the number of clusters. We then look for an “elbow” or a point at which increasing the number clusters ceases to drastically decrease the distance within clusters. Another aspect to consider when selecting how many clusters to use is that we have data for four years. It makes our analysis easier if the same number of clusters were selected for each year. In the case illustrated in Fig. 6, each year’s data displays an elbow at around eight clusters, which is the number of clusters we ultimately elected to use. Figure 7 is a map of the eight clusters drawn from the 2007 data.

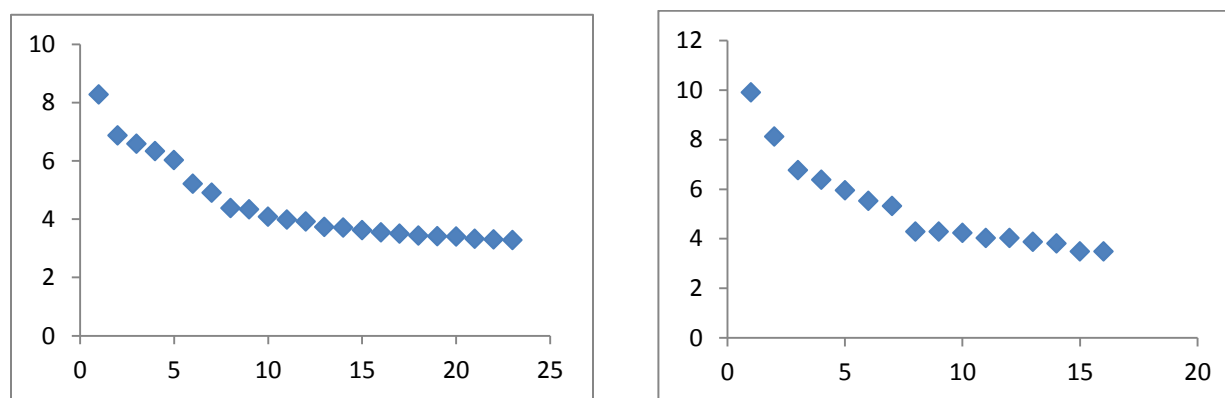


Fig. 6. Elbow plots of dendrogram height against number of clusters. Dendrogram height represents variance within clusters. From left to right are the elbow plots for 2007 and 2009. Ideally, the same number of clusters will be selected for each of the four years covered by our data set. Comparing the elbow plots for different years can be a useful tool when the “elbow” is not immediately clear. In this case, the rate of change of variance levels off around seven or eight clusters. After comparing all four years, we elected to use eight clusters.

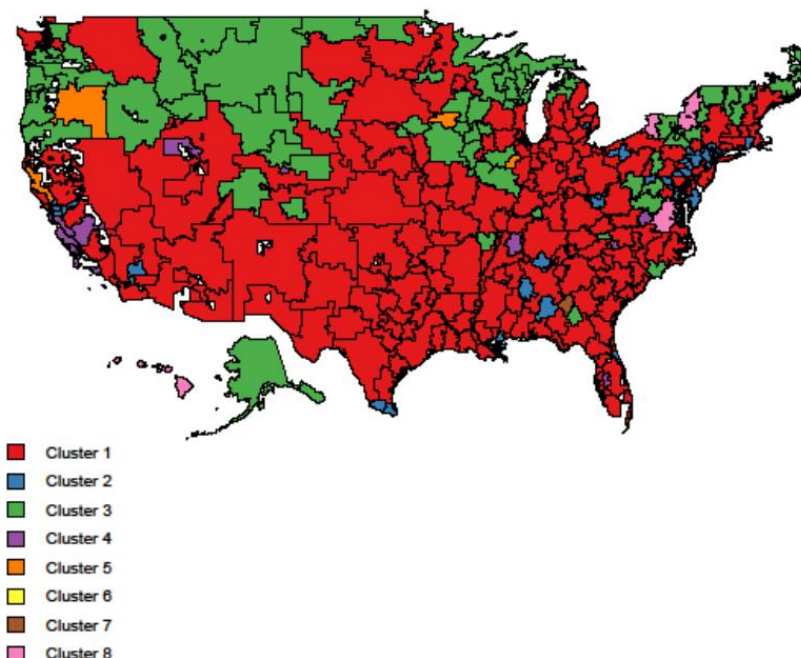


Fig. 7. Map of clusters for 2007. Note that the majority of HRRs belong to either Cluster 1 or Cluster 2, with the rest distributed over six smaller clusters. Clusters 6 and 7 consist of individual HRRs, Santa Cruz, CA and Columbus, GA, respectively.

3.2 EXTERNAL METHODS

3.2.1 Comparison to Cost

In order to determine if the service-usage clusters explain variation in cost, we constructed boxplots for each cluster (see Fig. 8). They do show different overall median costs, although especially in the cases of Cluster 1 and 2, it should be noted that they frequently have high variances and contain high-valued outliers. Ultimately, the connection between cluster identity and cost is not a particularly strong one.

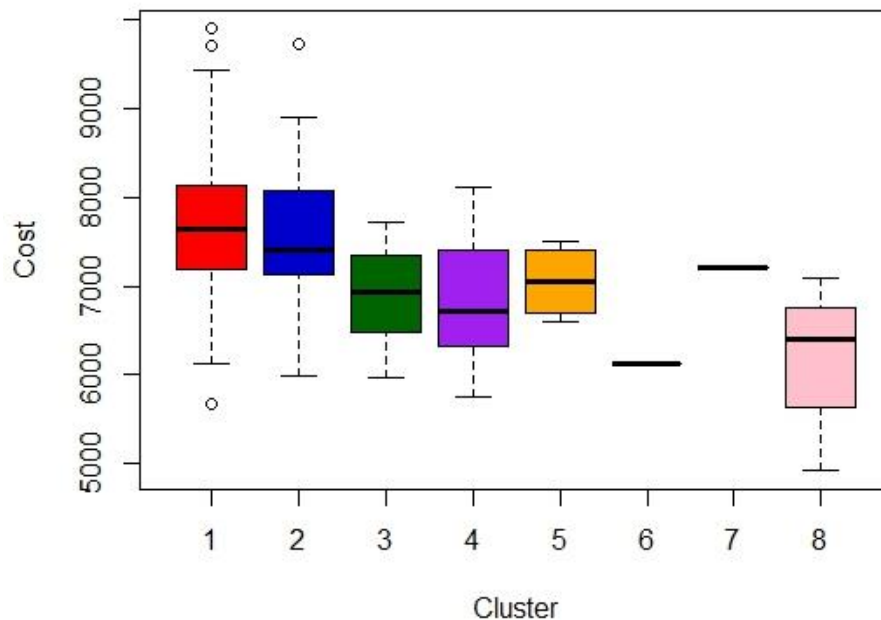


Fig. 8. Boxplot of 2007 clusters versus standardized risk-adjusted cost. The clusters display some difference in median cost, but particularly for Clusters 1 and 2, the spread is so great that comparing them is not particularly informative. Cluster 3, which is quite large, displays surprisingly low variance given its size.

When we used non-transformed data for clustering (see Fig. 9), the connection to cost was clearer, although most HRRs were placed in a single cluster. This difference likely partially due to the fact that clustering of the non-transformed data placed less weight on services used by a low proportion of the Medicare population. While these services may not have a large effect on overall cost, the degree to which they are used often varies considerably between regions.

3.2.2 Geospatial Autocorrelation

We had hypothesized that some of the geographic variations in Medicare service use might be the result of cultural differences or of broadly varied standards of physician practice. In either case, this would suggest that patterns of service usage might be similar in contiguous HRRs. Furthermore, there should be some patient movement between HRRs, which would also contribute to the development of broader geographic trends. To test this idea, we used spatial autocorrelation statistics in R to determine whether cluster identity of HRRs was geographically random, or whether adjacent HRRs were more likely to belong to the same cluster.

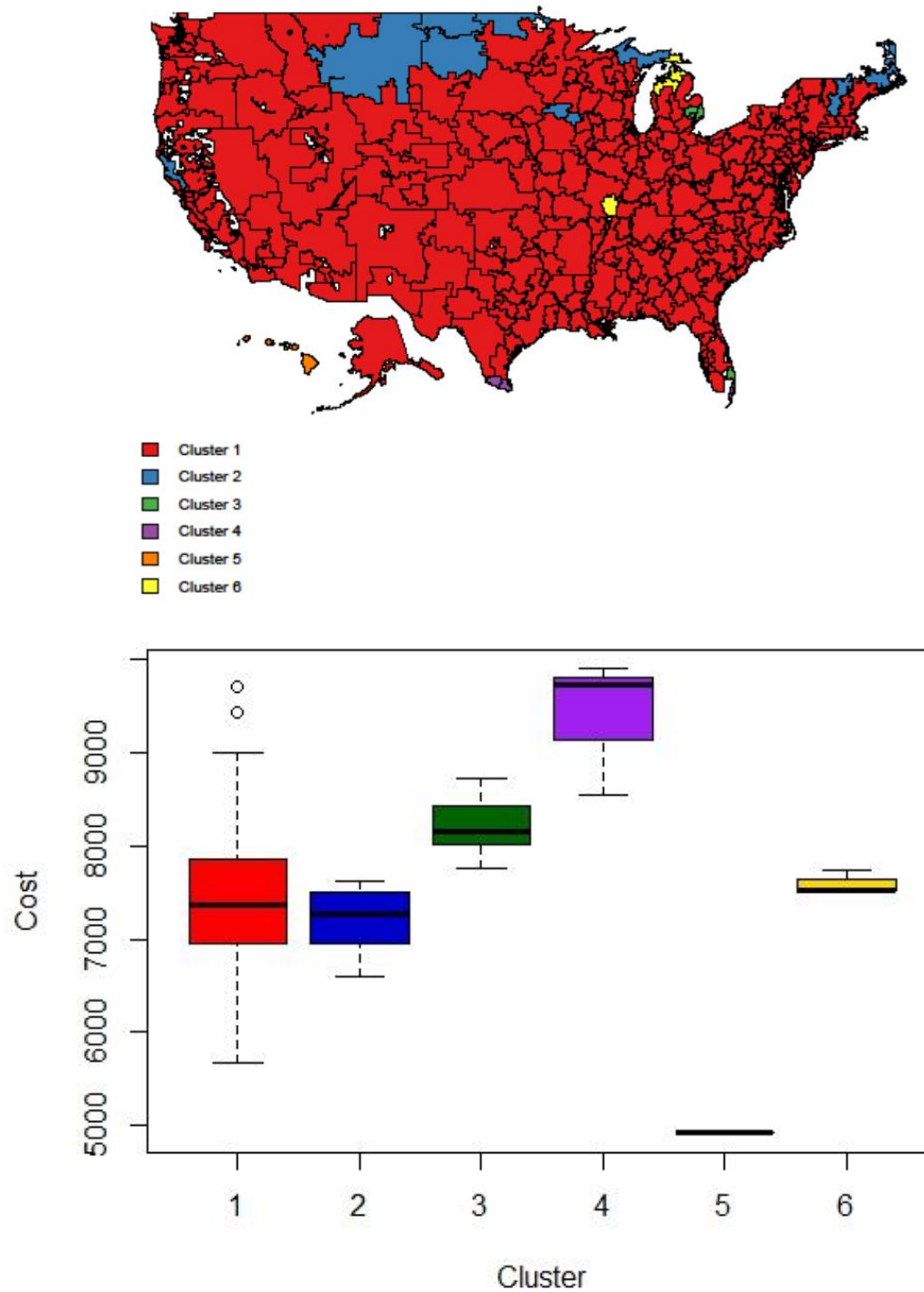


Fig. 9. Clustering and comparison to cost of non-transformed 2007 data. Although most HRR's were placed in a single large cluster, the remaining clusters show a much greater difference in standardized risk-adjusted cost than they did when the data was transformed.

We found that for 2007, HRRs belonging to the large clusters 1, 2, 3, and 4 displayed significant (by z-tests) tendencies towards contiguity, a trend which is visually apparent in Fig. 7. Clusters 5 and 8 were both small (cluster 5 contains four HRRs, while cluster 8 contains only three HRRs) and in no case were adjacent to another HRR with the same cluster identity. Clusters 6 and 7 contained only one HRR each. Clusters 5 and 3 as well as clusters 6 and 4 also showed a tendency towards contiguity. Some pairs of

clusters tended *not* to be contiguous. Clusters 2 and 1, 3 and 1, 3 and 2, and 4 and 1 were all in contact less frequently than would be expected under a random assignment of cluster identity.

The fact that adjacent clusters tend to be of the same cluster identity suggests that cluster assignment is not random, as does the apparent tendencies of some clusters to “avoid” each other. That HRRs are not fully grouped by geographic location merely suggests what we already expect—that there are many factors at play in defining service usage patterns and the geographic location is only one of them. Just why the geographic trends exist is a question ripe for future exploration.

4. CLUSTER PROFILE CONSTRUCTION AND COMPARISON

To compare service-usage patterns between clusters we calculated the mean value of the proportion of beneficiaries who use each service for each cluster. We then constructed a histogram of service usage for all clusters (see Fig. 10). Although some services, such as Evaluation and Management, are used in a fairly consistent manner between clusters, other services such as Long Term Care Hospitals and Home Health vary considerably. The variation in the proportion of the population using the various services across clusters was assessed and roughly divided into three groups: low, moderate and high (see Fig. 11). As Fig. 10 shows, in general, the services that showed the most variation were ones that were used overall by a relatively low proportion of the population. Each cluster displays a unique pattern of use of services with moderate to high variation, as illustrated in Fig. 12.

Some of the services, such as Long Term Care Hospitals,⁷ Critical Access Hospitals,⁸ Ambulatory Service Centers,⁹ and Rural Health Centers,¹⁰ are not consistently distributed throughout the United States, suggesting that the variance in their use may be a case of supply-driven demand; that is, if a service is accessible, it will be used, while obviously if it does not exist within a region, it will not be utilized. It is interesting to note that Long Term Care Hospitals and Critical Access Hospitals have been identified as being particularly cost-ineffective^{11,12} to the extent that there is currently a moratorium on building any more Medicare-supported Long Term Care Hospitals.¹² Ambulatory Service Centers, on the other hand, are a relatively new concept that is considered to be a more cost-effective option than a traditional hospital.⁹

Home health and hospice services both showed moderate variation and are more difficult to link to supply-induced demand. Research by Dartmouth Atlas has suggested that end of life care may be a driver for cost variation,¹³ although end of life care practices are currently a sensitive political issue. Hospice services are a component of end of life care, so the variation in how they are used is interesting to note. Home Health, is an area with a history of fraud.^{9,14}

5. CONCLUSION AND FUTURE WORK

5.1 SUMMARY OF FINDINGS

Our work revealed the usefulness of clustering as a method for working with Medicare data. Clustering let us compare the use of 22 services in 306 regions with relative ease and efficacy of time and let the data speak for itself in a way that uncoupled service use from cost data. The cluster analysis allowed us to identify approximately eight distinct service-usage patterns which are amenable to more fine-grained analysis in order to determine how these patterns are related to Medicare spending. Because the clusters were based on general patterns we could compare clusters directly to one another, a task that would have been nearly impossible if we were looking on the individual HRR level.

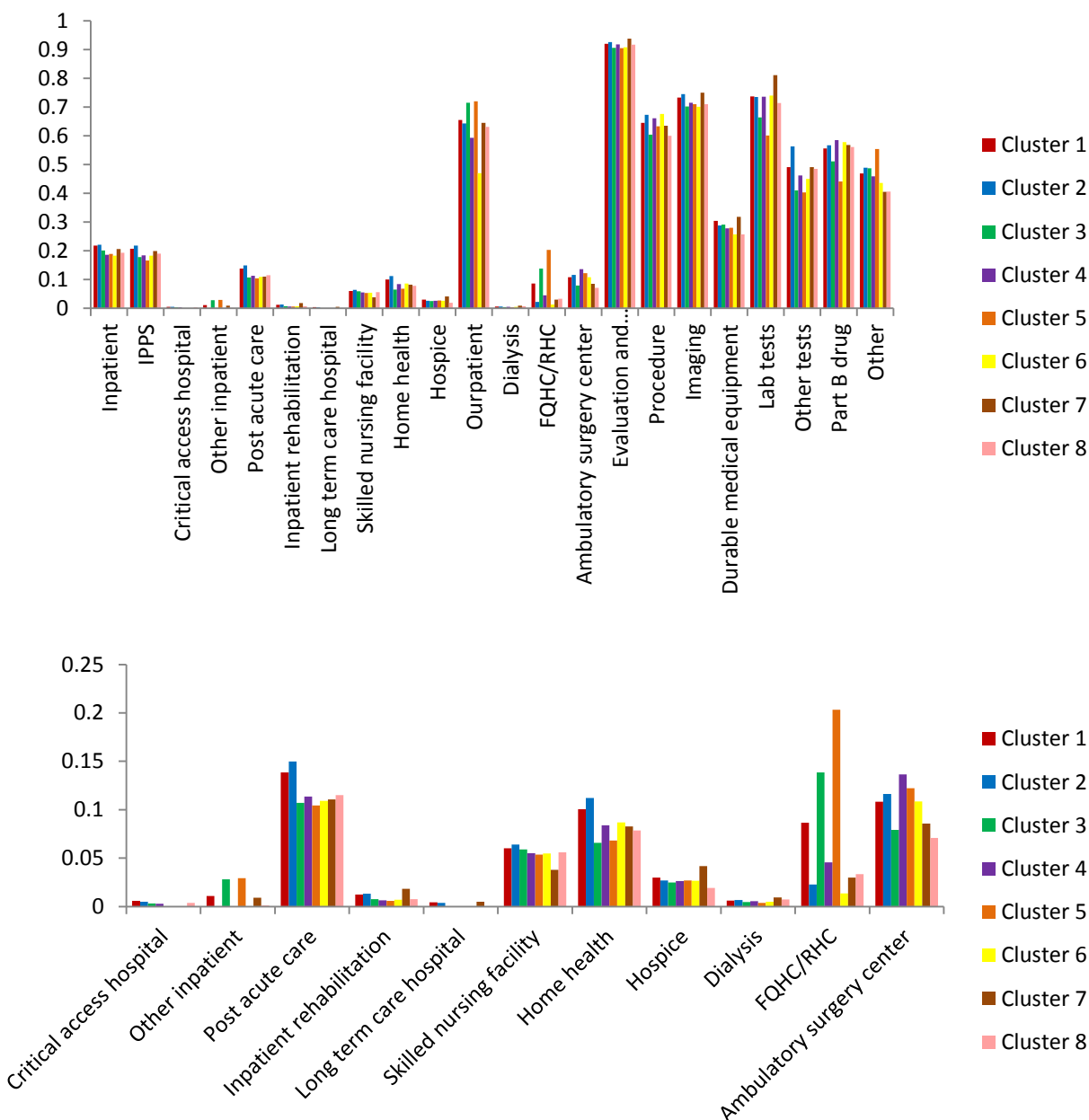


Fig. 10. Histogram of service use by cluster. The proportion of each cluster using a service is plotted for each service. The services that showed the most variation between clusters were services that are not used by a large proportion of the Medicare population.

Contiguous HRRs showed a tendency to appear in the same cluster, suggesting that HRRs were not just being assigned randomly to clusters. The tendency of contiguous HRRs to be assigned the same cluster identity supports our hypothesis that geographic variation in service usage may be influenced by broader regional norms, both in terms of physician standards of practice and lifestyles and medical conditions within the Medicare population. The fact that clusters are not completely geographically uniform merely suggests that, as expected, there are many other factors at play that determine how services in an HRR will be used by the Medicare population.

<u>Low variation</u> Inpatient IPPS Outpatient Evaluation and management Procedure Imaging Durable medical equipment Lab tests Other tests Part B drug Post-acute care Other	<u>Moderate variation</u> Skilled nursing facilities Home health Hospice Dialysis Ambulatory service centers	<u>High Variation</u> Critical access hospitals Other inpatient Inpatient rehabilitation Long term care hospitals Federally qualified/Rural health centers
---	--	--

Fig. 11. Degree of variation of service use. The variance in the proportion of the population using services between clusters ranged from low to high. Twelve services showed low variation between clusters, five showed moderate variation, and five showed high variation. Services that showed moderate to high variation tended to be used by a lower proportion of the population than services that showed low variation.

We identified twelve services that exhibited low variation across clusters, five that exhibited moderate variation, and five that exhibited high variation. In general, services that were used by a high proportion of the Medicare population showed low variation, while service used by a very low proportion of the population varied considerably between clusters. Several of the services that showed high variation in use are facility-specific and are not consistently available throughout the United States, suggesting that their variation between clusters is the result of supply-induced demand.

The connection between cluster identity and standardized risk-adjusted per capita costs was relatively weak, although the median range among all clusters was substantial (\$6500–\$8100). The absence of a stronger link may have been affected by our decision to transform the data in such a way that services used by a low proportion of the population are given similar weight as those services used with greater volume. However, this same choice illuminated variation in low-volume services that might not otherwise have been identified.

5.2 POSSIBLE EXPLANATIONS

Several reasons exist for why low volume services demonstrated such high variance. In some cases, the facilities simply were not equally available in all HRRs. The proportion of a population using other services, such as dialysis, is likely driven by the medical conditions present in the local population, which in turn is affected by underlying demographic factors. In general, the low-use high-variation services were more specific than the high-volume, low-variation services. It is possible that if high volume services such as Evaluation and Management or Procedure were further subset, those subcategories would demonstrate greater individual variation between clusters. Finally, for very-low volume services, we must acknowledge the possibility of some random variation between HRRs.

The question of what drives contiguous HRRs to display similar service-usage patterns is an interesting one. We have hypothesized that it may be a combined effect of people moving between HRRs for treatment, broader geographic trends in medical standards of practice, and broad-scale demographic trends in the Medicare population. Movement of population between HRRs is especially likely in areas of high population density and small geographic HRR size. Studying areas in which cluster identity is particularly consistent and areas where a single HRR of a different cluster interrupts a geographic trend could shed some light on this question.

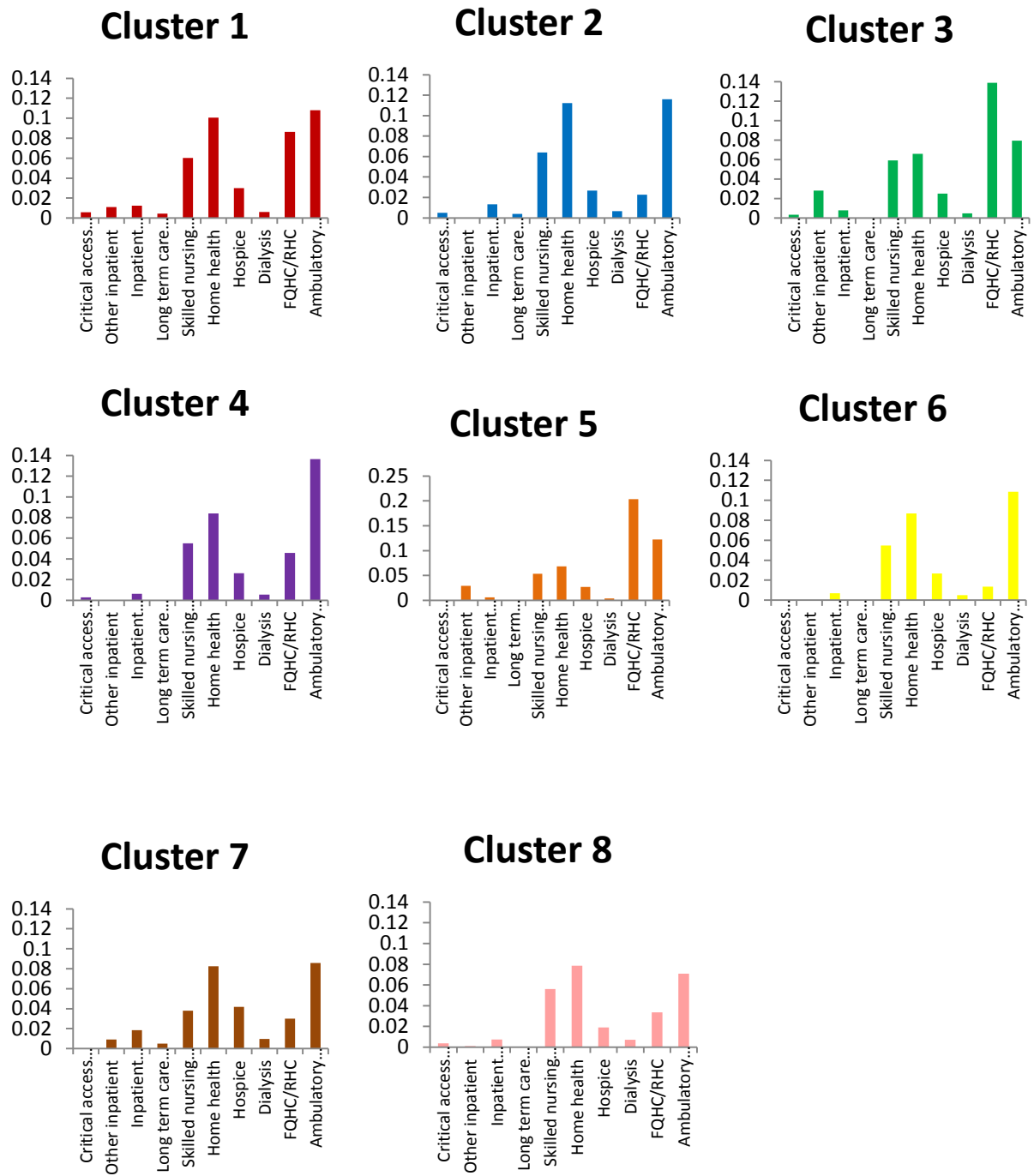


Fig. 12. Histograms of high-variation services by clusters. Each cluster displays its own pattern of service usage. Note that the y-axis on for all clusters ranges form 0 to 14, except in the case of Cluster 5, where the high proportion of FQHC/RHC use necessitated increasing the range to 25. Some services are used by all clusters, while other that are unevenly distributed geographically are not used at all by some clusters.

5.3 OUTSTANDING QUESTIONS

Our initial round of clustering analysis of geographic variation in Medicare service-usage and its relationship to standardized risk-adjusted per capita cost raised at least as many questions as it answered and suggested several avenues of future research. The most direct options are of re-running our analysis with some modifications to the data used. First, because some of the services we identified as having high variation between clusters are not consistently available throughout the US, we suggest removing them from the data set, since we already know that they vary. In addition to the proportion the population that uses a service each year, the CMS data set contains similar information on how much money was spent per capita in an HRR on a service, and how many times that service was used, regardless of the number of individuals utilizing it. It would be interesting to see how clustering HRRs based on this information would compare to our initial analysis. Clearly, it would be redundant to compare groups of HRRs clustered on service-specific spending, but it might reveal variation in service-specific costs in different regions. If sufficient demographic information could be obtained on the HRR level, we would like to compare clusters to their demographic profiles, and if possible even try clustering HRRs based on demographics.

As discussed above, the services listed in the CMS data set vary in their specificity. Subdividing large categories such as Evaluation and Management or Procedure would result in an even larger set of information, but our experience with clustering analysis over 22 service-usage categories suggests that the technique would remain effective for many more factors. Thus if the information is available, it could help further identify variations in patterns of service-usage.

Our largest outstanding question is one of temporal variation. We have not yet developed a methodology for comparing clusters over time, leaving us to focus on individual years even though we have access to four years of data. Even if we could track clusters over time, we would need to factor in the effects of policy changes, which can be difficult to identify clearly. We would also need to consider shifting demographic trends within Medicare population as it changes each year as beneficiaries pass on and new people age in to Medicare eligibility. Nevertheless, determining whether clusters show consistency over time is a vital component to making any generalizations about geographic variation in service-usage, or the connection between patterns of service-usage and cost.

6. REFERENCES

1. T. Rosenthal, "Geographic Variation in Health Care," *Annual Review of Medicine*, 63:593-509, 2012.
2. J. E. Wennberg, et al., "Improving Quality and Curbing Health Care Spending: Opportunities for the Congress and the Obama Administration," *The Dartmouth Atlas of Health Care* (2008), accessed on December 12, 2012.
3. P. Orszag, *Health Care and Behavioral Economics*, presentation to the National Academy of Social Insurance, Congressional Budget Office (2008), accessed October 2012.
4. E. S. Fisher, D. C. Goodman, and A. Chandra, "Disparities in Health and Health Care among Medicare Beneficiaries," *The Dartmouth Atlas of Health Care*, 2012.
5. Institute of Medicine, *Table 1: HRR Level Demographic, Cost, Utilization, and Quality Data*, accessed August 2012. <http://iom.edu/Activities/HealthServices/GeographicVariation/Data-Resources.aspx>
6. Wikimedia Commons, *Logit Function*, accessed November 2012.
7. Centers for Medicare and Medicaid Services, *Long Term Care Hospitals*, 2011, accessed October 2012 through the Rural Assistance Center. <http://www.raonline.org/racmaps/mapfiles/ltch.png>
8. Centers for Medicare and Medicaid Services, *Critical Access Hospitals*, Quarter 2, 2011, accessed October 2012 through the Rural Assistance Center. <http://www.raonline.org/racmaps/mapfiles/cah.png>
9. Karen Milgate, personal communication to Linnet Vacha, Oak Ridge National Laboratory, Oak Ridge, TN, November 29, 2012.
10. Centers for Medicare and Medicaid Services, *Rural Health Clinics*, Quarter 2, 2011, accessed October 2012 through the Rural Assistance Center.
11. A. Berenson, *Long-Term Care Hospitals Face Little Scrutiny*, The New York Times 2010, accessed November 9, 2012.
12. J. Gold, "When 'Critical Access' Hospitals Are Not So Critical," *Kaiser Health News* 2011, accessed October 2012.
13. The Dartmouth Atlas of Health Care, *End-of-Life Care*, accessed December 12, 2012.
14. J. Cloherty and P. Thomas, *Biggest Medicare Fraud in History Busted, Say Feds*, ABC News 2012, accessed December 12, 2012.