Method to Develop Health Care Peer Groups for Quality and Financial Comparisons Across Hospitals

*Margaret M. Byrne, Christina N. Daw, Harlan A. Nelson,
Tracy H. Urech, Kenneth Pietz, and Laura A. Petersen*

Objective. To develop and explore the characteristics of a novel “nearest neighbor” methodology for creating peer groups for health care facilities.

Data Sources. Data were obtained from the Department of Veterans Affairs (VA) databases.

Statistical Methods and Findings. Peer groups are developed by first calculating the multidimensional Euclidean distance between each of 133 VA medical centers based on 16 facility characteristics. Each medical center then serves as the center for its own peer group, and the nearest neighbor facilities in terms of Euclidean distance comprise the peer facilities. We explore the attributes and characteristics of the nearest neighbor peer groupings. In addition, we construct standard cluster analysis-derived peer groups and compare the characteristics of groupings from the two methodologies.

Conclusions. The novel peer group methodology presented here results in groups where each medical center is at the center of its own peer group. Possible advantages over other peer group methodologies are that facilities are never on the “edge” of a group and group size—and thus group dispersion—is determined by the researcher. Peer groups with these characteristics may be more appealing to some researchers and administrators than standard cluster analysis and may thus strengthen organizational buy-in for financial and quality comparisons.

Key Words. Peer groups, cluster analysis, nearest neighbor, Euclidean distance, quality of care comparisons

Measuring and reporting health care facility performance via clinical measures of quality has become a major strategic initiative in improving the quality of care for Medicare and other health care payers and delivery systems (e.g., Centers for Medicare and Medicaid Services 2008). Establishing appropriate peer groups (i.e., grouping entities by similarity on specific characteristics) for such comparisons can help health care leaders and administrators

make equitable comparisons across hospitals or health systems. Peer groups can be used in health care systems for a variety of purposes including promoting fair allocation of resources, evaluating efficiency or financial performance, as well as assessing quality of care and outcomes. For these evaluations, peer groups help control for systematic risk and constraints presented by various influences on a hospital's finances or clinical outcomes (Ellis and McGuire 1988; Stefos, LaVallee, and Holden 1992). These constraints and risk are generally not easy for administrators to change within a reasonable time. Hence, the rationale for peer groups is to place hospitals or health systems facing similar structural and patient characteristics together and facilitate "like-to-like" fair comparisons.

In the health care field, peer groups have become key elements of hospital industry quality benchmarking analyses (Zodet and Clark 1996; Chen et al. 1999; Austin et al. 2004; Solucient 2006). Washington State has incorporated hospital peer groups in its hospital rate schedule for state-funded care (Holubkov et al. 1998). Canada has developed regional health peer groups within its provinces, where the peer group is a "cluster of health regions that have similar social and economic health determinants" (Statistics Canada 2002; MacNabb 2003). The World Health Organization Regional Office for Europe uses peer groups as a tool to assess quality improvement internationally in European hospitals (World Health Organization 2007).

However, although often used, peer groups as commonly constructed have several characteristics that may not be desirable. Peer groups customarily have mutually exclusive membership, with characteristics or attributes defined in categories (e.g., categories of bed size, urban versus rural, teaching versus nonteaching, categories of patient severity or case mix). Hence, peer groups may have some members on the "edge" of a group with respect to certain characteristics, making comparisons with group members appear imbalanced or unfair. For example, a hospital with 152 beds may be placed in a group where bed size is 150+ beds, and nearly all other hospitals in the group have a much greater number of beds. As hospitals are increasingly

Address correspondence to Laura A. Petersen, M.D., M.P.H., Health Services Research and Development (152), Houston Veterans Affairs Medical Center, 2002 Holcombe Blvd., Houston, TX 77030; e-mail: laurap@bcm.edu. Margaret M. Byrne, Ph.D., is with the Department of Epidemiology and Public Health, University of Miami, Miami, FL. Christina N. Daw, M.P.H., Ph.D., Tracy H. Urech, M.P.H., and Kenneth Pietz, Ph.D., are with the, Health Services Research and Development Center of Excellence, Houston VA Medical Center, and Section for Health Services Research, Baylor College of Medicine, Houston, TX. Harlan A. Nelson, M.S., is with the InVentive Clinical Inc., Indianapolis, IN.

scrutinized in quality and financial evaluations, they might be concerned about being on the edge in terms of characteristics that are used to define the groups. In addition, traditional cluster analysis, a common tool for creating peer groups in hospital systems (Klastorin 1982; Alexander, Evashwick, and Rundall 1984; Stefos, LaVallee, and Holden 1992), may create groups of widely varying sizes. It is not uncommon for clustering analysis to produce groups with as little as one member, or as many as dozens of members.

To counter these potentially undesirable characteristics of peer groups, we sought to develop a new methodology for constructing peer groups using the Department of Veterans Affairs (VA) medical centers as the sample population of facilities. The VA is an ideal place to explore the development of a new methodology for creating peer groups. The VA consists of 21 regional Networks with 137 medical centers and hundreds of outpatient clinics. The VA medical centers vary widely in size, patient illness burden, geographic location, and numerous other characteristics. This variation is essential for exploring peer groupings, as systems with too similar facilities may not be amenable to formation of distinct peer groups. In addition, the development of valid peer groups in the VA is essential to ensure fairness in funding allocation across medical centers (Stefos, LaVallee, and Holden 1992) and for comparison of efficiency and quality of care.

In this research, our main objective was to develop and explore a novel methodology that can be used to create peer groups in any health care system. Our work is innovative in that, unlike peer groups based on traditional cluster analysis, our methodology develops a peer group customized to each medical center by identifying the “nearest neighbor” medical centers, according to the selected characteristics for comparison. We use a Euclidean distance measure as the metric to evaluate “nearness.” Thus, each medical center is the hub of its own peer group, but will also appear in peer groups of other medical centers. In this paper, we explore the characteristics of our “nearest neighbor” peer groups and compare with peer groups constructed using traditional cluster analysis.

METHODS

Sample and Characteristics for Peer Group Development

Our sample of facilities was composed of 133 VA medical centers, excluding sites in Puerto Rico, the Philippines, New Orleans, LA, and Biloxi, MS. (Puerto Rico and the Philippines were excluded because of incomplete data,

while New Orleans and Biloxi were excluded due to substantial disruptions in service due to hurricanes during the period of study.) To create our peer groups, we used characteristics that reflected patient complexity, academic mission, medical research, and size (Stefos, LaVallee, and Holden 1992; Hogan, Franzini, and Boex 2000; Maciejewski et al. 2002; Larson and Fleishman 2003; Wasserman et al. 2003). For example, our patient-level measures included patient illness burden and veteran reliance on the VA (versus Medicare) for health services among the medical center's Medicare-eligible veterans (Wright et al. 1999; Shen et al. 2003; Ash, Shwartz, and Peköz 2003; Petersen et al. 2005; Byrne et al. 2006). Where possible, we used measures that were continuous rather than binary or categorical, as continuous variables provide more refined information for the empirical analysis.

We identified the following domains as important for equitable comparison of medical centers, and selected 16 measures that fall into these domains.

1. Complexity (including tertiary care, referral care)
2. Size (of medical center and patient population)
3. Academic mission
4. Disease burden in the patient population
5. Patient reliance on the VA for health care
6. Care delivery structure, for example, VA-specific medical care programs
7. Infrastructure at the medical center level
8. Community and environment (density and resources in the surrounding community)

Data were from the Federal Fiscal Year 2005. It is important to note that the selection of appropriate variables for use in forming peer groups is a crucial step in the peer group formation process. However, the approach to identifying candidate variables and final selection of characteristics will be highly dependent on the type of facilities that are being grouped and the intended use of the peer groups. Therefore, the appropriate approach and final characteristics will vary with each application.

"Nearest Neighbors" Peer Groups

The main goal of this research was to develop a novel methodology for determining peer group identity. Our approach draws conceptually from traditional cluster analysis, which is a methodology that evaluates similarities between entities, in order to form groups of these entities. The advantage of

using a cluster analysis type approach for developing medical center peer groups is that several characteristics of interest can be examined simultaneously. Thus, the medical centers can be compared for similarity with each other based on multiple attributes. Like cluster analysis, our methodology also allows for simultaneous consideration of multiple characteristics and attributes. However, there are differences in our approach that we believe confer advantages in some situations.

Specifically, traditional cluster analysis results in a set of mutually exclusive peer groups. These peer groups are almost always of varying sizes; in particular, a single medical center can constitute a group (MacNabb 2003). Our peer grouping methodology creates groups in which each medical center is the reference point for a cluster. Peers for that reference medical center are those centers that are “nearest neighbors” based on the characteristics that we chose for comparison. Therefore, the nearest neighbor groupings are not mutually exclusive, and clusters are not necessarily of varying sizes.

The steps in our methodology for developing peer groups are as follows. First, to mitigate the effects of the widely disparate scales of the characteristics of interest, we standardized all variables (including binary ones) to have a mean of 0 and standard deviation of 1. Second, we used the SAS[®] software’s *Proc Distance* procedure (SAS 2002) to calculate the Euclidean distance between all hospitals in multidimensional space. The Euclidean distance between medical centers is calculated by summing the squares of the individual values from the measures, then taking the square root of the result. The metric of Euclidean distance operates in multidimensional space, where each characteristic provides a dimension.

Third, we set criteria for determining the number of peers in a group. Unlike traditional cluster analyses, the size of the peer groups is not determined in the clustering procedure. There are, in fact, a number of possible ways that one could select the number of peers per group. For instance, one could simply select, based on external criteria, the number of peers that are desired per group. Alternatively, one could select a radius in Euclidean distance and identify all medical centers within that distance or closer as peers. As with traditional cluster analysis, such a method may provide too few peers (in some cases) or too many peers (in other cases) for practical comparison. For the exploration of the characteristics of our peer groups, we created peer group sets of varying sizes.

Traditional Cluster Analysis Peer Groups

For comparison with our methodology, we generated peer groups using a standard two-stage cluster analysis technique (Sharma 1995). We used the

same characteristics and data as in the nearest neighbor approach. Specifically, we generated cluster seeds by Ward's method of hierarchical clustering. Based on the resultant R -squared, which is the proportion of variation in distance that is explained by the clustering, we elected to create seven peer groups from the set of medical centers. We then used these seeds as input to the standard k -means iterative algorithm for cluster analysis. This iterative partition process assigns each of the medical centers to one of the seven peer groups.

Exploration of Peer Group Characteristics and Comparison with Traditional Cluster Analysis

We explored the attributes of the peer groups created with our "nearest neighbor" methodology, and compared them with peer groups formed using traditional cluster analysis.

Characteristics of Nearest Neighbor Peer Groups. We present and discuss the characteristics of peer groups (here using 15 medical centers per group) to illustrate: (1) the nondiscrete nature of the groups, and (2) the specific attributes of medical centers that end up being peers.

Effect of Number of Medical Centers in a Peer Group on Distance to Furthest Peer. To explore how the number of medical centers in a peer group affects the Euclidean distance from the reference medical center to furthest peer, we developed five sets of peer groups using 5, 10, 15, 25, and 40 as the number of medical centers per peer group. We then calculated for each peer group in each of the sets, the Euclidean distance from the reference medical center to farthest peer. We present graphically the distribution of these Euclidean distances for each of these sets of peer groups.

Comparison with Traditional Cluster Analysis. As nearest neighbor and traditional cluster analysis have different methodologies, and result in different numbers of groups, there is no obvious metric to use for comparison. Note that we cannot use the Euclidean distance from reference facility to furthest neighbor, as there is no "central" medical center in cluster analysis-derived peer group. Thus, to indicate how the two different methodologies compare, we computed the square root of the sum of squares (RSS) of the Euclidean distances between all pairs of members of the same group, for all peer groups. The RSS of distances is a metric that takes into account the

distances between all members of each peer group, and thus in an informal way can be considered an overall measure that represents the diffusion of the peer group. Although not formally a diffusion measure, for simplicity we will use that term. Note that higher values indicate a more diffuse or less dense group. We calculated the diffusion of all the peer groups both in the traditional cluster analysis and in the nearest neighborhood peer group sets of 5, 10, 15, 20, 25, 35, and 40 members. We present minimum and maximum RSS of distances across peer groups for the nearest neighbor groups, and the RSS of distance measure for all seven of the peer groups created using traditional cluster analysis.

RESULTS

As the nearest neighbor methodology creates a unique peer group for each medical center, our analysis produced 133 peer groups. These peer groups are not mutually exclusive. (In our example in the introduction of a medical center with 152 beds, this hospital is now much more likely to be placed in a peer group [of its own] that includes medical centers with fewer as well as more beds than it has, rather than in a group with only larger facilities.) As described earlier, the number of medical centers that will be included in each peer group is a parameter that can either be chosen by the researcher or determined by some rule such as maximum distance allowed. In contrast, the traditional cluster analysis resulted in seven mutually exclusive groups that ranged in number of medical centers from 6 to 42.

Characteristics of Nearest Neighbor Peer Groups

Nondiscrete Nature of the Groups. Using the nearest neighbor methodology, medical center A may be a peer in medical center B's group, without B appearing in A's group. That is, the groups are not discrete. Table 1 provides an example of this phenomenon. Medical center KKKK is a peer facility for medical center BBB (in italics, ranked 8), but BBB is not a peer for KKKK. This situation would not occur in the nonoverlapping peer groups constructed using traditional cluster analysis methods. BBB's 15th peer (the most distant of its peer group) is medical center SSSS with a Euclidean distance of 4.701 while KKKK's 15th peer (LLLL) is closer with a distance of 3.667. Medical center KKKK's distance from BBB is 3.963, which explains why BBB is outside of KKKK's peer group. The 15 peer facilities for KKKK

Table 1: Example of Two Peer Groupings and the Euclidean Distance from Reference Medical Center to Each Peer

<i>Peer Rank</i>	<i>Peer Group for Medical Center BBB</i>	<i>Euclidean Distance to Peer</i>	<i>Peer Rank</i>	<i>Peer Group for Medical Center KKKK</i>	<i>Euclidean Distance to Peer</i>
1	Medical center PPPP	2.909	1	Medical center NNN	1.589
2	Medical center BBBB	3.323	2	Medical center LL	2.186
3	Medical center NNN	3.546	3	Medical center OO	2.397
4	Medical center UUUUU	3.719	4	Medical center TTTTT	2.696
5	Medical center UU	3.724	5	Medical center EE	2.841
6	Medical center LL	3.770	6	Medical center AAAAA	2.921
7	Medical center CC	3.777	7	Medical center YYY	3.016
8	Medical center KKKK	3.963	8	Medical center SSSS	3.121
9	Medical center WWW	3.993	9	Medical center PPPP	3.145
10	Medical center IIII	4.187	10	Medical center UUUU	3.328
11	Medical center GGGG	4.361	11	Medical center HHHH	3.403
12	Medical center AA	4.398	12	Medical center CC	3.434
13	Medical center AAAAA	4.546	13	Medical center UUU	3.617
14	Medical center II	4.619	14	Medical center HH	3.644
15	Medical center SSSS	4.701	15	Medical center LLLL	3.667

are captured in a shorter range of distance compared with the 15 peer facilities for BBB; therefore, KKKK's peer group is less dispersed than BBB's peer group. Also note that medical centers PPPP, NNN, LL, CC, and SSSS appear in both peer groups.

Specific Attributes of Facilities That End Up Being Peers. To explore in more detail why two medical centers are closer or further apart in Euclidean distance, one can examine medical centers' actual values on the characteristics used to create the peer groupings. Medical centers that are close on characteristics will also naturally be close in Euclidean space. Table 2 shows examples from peer group data, listing the values for four of the 16 measures for medical center Y and two of its peers, medical center MM and medical center GG.

Y's closest peer is MM. These two medical centers share similar scores on hospital beds (295 and 248, respectively) and Diagnosis-Related Groups (DRG) Index (2,832.46 and 2,638.51, respectively), to give two examples. (We defined the DRG Index as the sum of DRG relative weights in FY 2005 for all DRGs with relative weight > 3). The two medical centers are ranked rather closely on the above measures, among the 133 VA medical centers. Y's furthest peer, GG, has somewhat more divergent values for hospital operating

Table 2: Selected Measures for Medical Center Y and Two Peers: Values (Rank among VA Medical Centers)

<i>Medical Center</i>	<i>Euclidean Distance</i>	<i>Hospital Operating Beds</i>	<i>DRG* Index</i>	<i>Resident Slots/ 10k Patients</i>	<i>DCG[†] Relative Risk Score</i>
Medical center Y	–	295 (10)	2,832.46 (12)	41.43 (6)	1.473 (6)
Medical center MM (closest peer to Y)	6.775	248 (20)	2,638.51 (14)	32.30 (11)	1.312 (23)
Medical center GG (15th in Y's peer group)	7.997	161 (40)	1,930.55 (30)	22.19 (41)	1.252 (41)

*Diagnosis-Related Groups Index: Sum of DRG relative weights, for DRGs with relative weight > 3, recorded in Federal Fiscal Year 2005.

[†]Diagnostic Cost Groups, from DxCG[®] risk adjustment software.

beds (Y has 295 while GG has 161) and DRG index (2,832.46 and 1,930.55, respectively). Yet their relative rankings on these measures are still within 35 places of each other, among all 133 medical centers. Illustrating again the nondiscrete nature of the groups, GG appears in Y's peer group but Y does not appear in GG's peer group (data not shown). The 15 peer medical centers for GG are captured in a smaller range of Euclidean distance, or multidimensional space, than medical center Y's 15 peer medical centers (3.458 versus 7.997, not shown). Thus, GG's peer group is less spatially dispersed, than the Y peer group.

Effect of the Number of Medical Centers in a Peer Group on Distance to Furthest Peer

Figure 1 shows the distribution of the Euclidean distances from reference medical center to furthest peer for each set of 133 peer groups, with membership determined at 5, 10, 15, 25, 40 per group, respectively. For example, for the set of peer groups where each peer group contains five medical centers, the largest number of peer groups (34) has a Euclidean distance of 2–2.5 from reference medical center to furthest peer. By construction, this radius distance increases as the number of medical centers in a peer group increases. However, the center of the distributions did not move to the right as quickly as we had expected as the number of medical centers per group increased. In fact, peer groups with 40 members still have the largest number of peer groups (34) in a Euclidean distance range (3.5–4) less than twice that of peer groups with five members. Not surprisingly, the distributions

Figure 1: Distribution of distances from reference medical center to furthest neighbor included in the peer group; calculated for 133 per groups, with membership composed of 5, 10, 15, 25, and 40 medical centers, respectively

are right skewed with a few peer groups in all five sets having much larger Euclidean radius distances.

Comparison with Traditional Cluster Analysis

Table 3 provides the results of the comparison of the traditional cluster analysis with the nearest neighbor method. The RSS of distances between members of a peer group, a measure of diffusion, is an increasing function of the number of members in the group. The median RSS for the clusters generated by cluster analysis (34.7) is similar to the median of our method for $n = 15$ (36.5). Note, however, that the maximum RSS for our method for $n = 15$ (62.7) is somewhat higher than the RSS for the 15-member group from the cluster analysis (47.4). This illustrates the fact that standard cluster analysis is designed to create groups with approximately the same density, while in this instance we base our groups on a constant number of members around a reference medical center. When nearest neighbor groups are constructed in this way, there can be a great variation in diffusion among groups. However, if

Table 3: Summary Measures for the RSS of Euclidean Distances (Diffusion) for Groups Generated by Traditional Cluster Analysis and the Nearest Neighbor Method

<i>Measure</i>	<i>Cluster Analysis</i>	<i>Nearest Neighbor</i>
Total number of groups	7 (mutually exclusive)	133 (nonmutually exclusive)
Range, mean, and median number of peers per group	(6, 42) 19.0, 13	Discretionary
List of number in cluster and RSS of distance across peer groups (Cluster Analysis); Minimum, maximum	6-member: 22.10 8-member: 31.05 9-member: 30.91	5-member: 6.60, 22.07 10-member: 15.63, 42.19 15-member: 26.30, 62.73
RSS of distances across peer groups (nearest neighbor analysis)	13-member: 34.70 15-member: 47.39 40-member: 111.26 42-member: 115.93	20-member: 38.26, 82.78 25-member: 51.75, 100.73 35-member: 78.73, 138.74 40-member: 92.26, 158.47
Median distance across peer groups	34.70	5-member: 11.32 10-member: 23.28 15-member: 36.54 20-member: 48.93 25-member: 62.31 35-member: 93.91 40-member: 108.34

one constructed nearest neighbor groups based on a specified diffusion or Euclidean distance, this variation of course would not occur.

DISCUSSION

Health care policy makers and administrators frequently wish to compare the performance of health care facilities. Indeed, measuring and reporting health care facility performance via clinical measures of quality has become a major strategic initiative in improving the quality of care for Medicare and other health care payers and delivery systems (e.g., Centers for Medicare and Medicaid Services 2008). We have developed a methodology for developing customized, nonmutually exclusive peer groups, with a given health care entity being the reference point or center of a unique comparison group. This methodology can be applied to health care systems, regions, clinics, or pharmacies. Our nearest neighbor methodology differs from traditional cluster analysis in a number of ways. Most conventional cluster analyses result in a set of mutually exclusive peer groups, which are almost invariably of varying sizes. Also, the number of facilities/entities in each peer group is

determined during the creation of the groups. In contrast, our nearest neighbor methodology creates peer groups that are not mutually exclusive and each entity is the reference point and center of its group. In addition, the number of peers in a group is determined at the discretion of the researchers.

Nearest neighbor peer groups can be nondiscrete, in that medical center A can be a member of medical center B's peer group but not the converse, depending on how group membership is defined. If membership is defined based on the number of facilities, then varying Euclidean distances between medical centers and their peers can lead to this nondiscrete quality. Indeed, this lack of discreteness in the resulting peer groups illustrates the flexibility of the nearest neighbor method in individualizing the groups to a given reference medical center. However, if membership is defined by Euclidean distance to furthest peer or total density/diffusion of the group, this nondiscrete quality will not occur. In our exploration of the novel nearest neighbor peer groups, we found that the maximum Euclidean distance from reference point to furthest peer facility did not increase substantially as the number of members in a group increases. Instead, even when the membership of peer groups was increased eight-fold, the Euclidean distance for the plurality of peer groups was not even doubled.

In this research, we also provided a comparison of peer groups formed from using the nearest neighbor methodology to those created using standard cluster analysis. Such a comparison is difficult and somewhat tenuous, as the different methodologies do not yield results that lend themselves to obvious comparisons. In the comparisons we did perform, we found that the median diffusion of nearest neighbor peer groups with 15 members was approximately equivalent to the median diffusion of peer groups created using cluster analysis (median peer group size was 13). Thus, as a rough approximation, it appears that the nearest neighbor methodology produces peer groups that are at least as dense as those created using cluster analysis that is designed to maximize group density. However, for every given peer group from cluster analysis, the (approximate) same-sized nearest neighbor set of peer groups, constructed using a number-based determination of membership, had a maximum group diffusion that was larger. Thus, some peer groups created using the nearest neighbor method and including a fixed number of facilities will comprise facilities that are not very similar.

The likely reason for the maximum diffusion of the nearest neighbor methodology peer group being greater than a comparably sized cluster analysis generated peer group (when a fixed number of facilities is used to form nearest neighbor peer groups) is because of the inclusion in the nearest

neighbor groups of facilities that are not very similar to other facilities in that group. We can, in fact, identify seven facilities that are most responsible for the highest RSS scores. These medical centers are dissimilar to the other centers on most of the characteristics that we used to create the peer groups.

No peer grouping methodology can alleviate the concern that some medical centers will not have close peers. However, with the nearest neighbor methodology, we specifically know the Euclidean distance from a given medical center to its nearest peer(s). Therefore, a researcher or user of this methodology can decide a priori, based on a comparison with other peer distances or on some other calculation, the maximum distance they would consider for neighbors. In this way, medical centers that do not have any appropriate peers can be identified, and users can decide whether or not to include them in comparisons.

There are a couple of important issues that potential users of our methodology must address when they consider constructing their own peer groups. Of first importance is the question as to which characteristics to include. As discussed, the variables that should be used for peer group construction will be dependent on the entities being grouped and the purpose of the groupings. For our analyses, we focused upon structural characteristics, such as academic affiliation and number of beds, that were not readily changed by managers, and thus represent somewhat fixed constraints on performance. Second, although not included in the analysis here for clarity of presentation, it is possible to weight the characteristics that are used in terms of importance. Weights could be applied in a variety of ways; one way would be to multiply the desired weights to the squared characteristic values before they are summed and the square root taken. Thus, higher importance can be placed on specific variables that users believe should be more influential in forming the groups, such as rural location or other characteristics. Third, as discussed throughout, users of this methodology have discretion in choosing how to form final peer groups. Groups can be formed based on a certain number of members, if a specific number of comparators are desired. Alternatively, if a specific similarity or "closeness" of peers is desired, that can be specified to determine peer membership. Finally, users of our methodology should be aware that certain types of statistical analyses using the groups might not be valid. For example, the use of random effects models is questionable because the hospitals in a group cannot be considered a random sample of a larger population. Moreover, the same analysis on different groups that contain one or more of the same hospitals may yield different results. These issues highlight the need for further research on the use of significance testing with peer groups determined by this method.

One of the advantages of the nearest neighbor method is that the peer groups are more refined than often-used groupings, reflecting the multidimensional diversity of health care providers. Also, the multiple characteristics and dimensions in the methods account for the myriad complex factors that contribute to health care provider performance. While some health care managers might find traditional lists of peers to be relatively clear-cut (as, for example, they may show rural versus urban, or teaching versus nonteaching institutions), the nearest neighbor method can incorporate as many characteristics that leaders deem critical. This method also facilitates the use of continuous variables as measures.

Managers of hospital and health care systems struggle at times to achieve mutually acceptable means of evaluating and comparing hospitals under their leadership. Some researchers argue that comparing hospitals only to their peers, rather than all hospitals, perversely lowers the standards for those medical centers whose characteristics may be generally associated with lower quality performance (Romano 2004). However, this perspective is countered by the practical consideration that health care facilities or systems may have structural and patient-based differences that cannot be changed but do affect financial or quality outcomes. An advantage of the nearest neighbor peer group methodology is that it may strengthen organizational buy-in for use of peer groups for use in comparisons of financial performance, efficiency, and quality. As the use of benchmarking and quality reviews increases and evolves, the methods presented here may facilitate the formation of peer groups for comparison.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: The authors thank Mark Kuebeler for invaluable programming assistance. This research was conducted with support from a VA contract (Project XVA 33-097) at the request of Veterans Integrated Service Networks 6, 7, 9, 12, 15, and 23. This work was supported in part by the Houston VA HSR&D Center of Excellence (HFP90-020). Dr. Byrne holds a National Cancer Institute career development award (K07 CA101812). Dr. Petersen was a Robert Wood Johnson Foundation Generalist Physician Faculty Scholar (grant no. 045444) at the time that this work was conducted, and an American Heart Association Established Investigator Awardee (grant no. 0540043N).

Disclosures: None.

Disclaimers: The views expressed are solely of the authors, and do not necessarily represent those of the VA.

REFERENCES

- Alexander, J. A., C. J. Evashwick, and T. Rundall. 1984. "Hospitals and the Provision of Care to the Aged: A Cluster Analysis." *Inquiry* 21 (4): 303–14.
- Ash, A. S., M. Schwartz, and E. A. Peköz. 2003. "Comparing Outcomes across Providers." In *Risk Adjustment for Measuring Healthcare Outcomes*, Third Edition, edited by Iezzoni, L., pp. 297–334. Chicago: Health Administration Press.
- Austin, P. C., D. A. Alter, G. M. Anderson, and J. V. Tu. 2004. "Impact of the Choice of Benchmark on the Conclusions of Hospital Report Cards." *American Heart Journal* 148 (6): 1041–6.
- Byrne, M. M., M. Kuebler, K. Pietz, and L. A. Petersen. 2006. "Effect of Using Information from Only One System for Dually Eligible Health Care Users." *Medical Care* 44 (8): 768–73.
- Centers for Medicare and Medicaid services. "Compare Hospitals." [accessed on July 14, 2008]. Available at http://www.cms.hhs.gov/hospitalqualityinits/25_hospitalcompare.asp
- Chen, J., M. J. Radford, Y. Wang, T. A. Marciniak, and H. M. Krumholz. 1999. "Performance of the '100 top Hospitals': What Does the Report Card Report?" *Health Affairs* 18 (4): 53–68.
- Ellis, R. P., and T. G. McGuire. 1988. "Insurance Principles and the Design of Prospective Payment Systems." *Journal of Health Economics* 7 (3): 215–37.
- Hogan, A. J., L. Franzini, and J. R. Boex. 2000. "Estimating the Cost of Primary Care Training in Ambulatory Settings." *Health Economics* 9 (8): 715–26.
- Holubkov, R., V. L. Holt, F. A. Connell, and J. P. LoGerfo. 1998. "Analysis, Assessment, and Presentation of Risk-Adjusted Statewide Obstetrical Care Data: The StORQS II Study in Washington State—Statewide Obstetrics Review and Quality System." *Health Services Research* 33 (3, pt 1): 531–48.
- Klastorin, T. D. 1982. "An Alternative Method for Hospital Partition Determination Using Hierarchical Cluster Analysis." *Operations Research* 30 (6): 1134–47.
- Larson, S. L., and J. A. Fleishman. 2003. "Rural–Urban Differences in Usual Source of Care and Ambulatory Service Use: Analyses of National Data Using Urban Influence Codes." *Medical Care* 41 (7 suppl): III65–74.
- Maciejewski, M. L., M. K. Chapko, A. N. Hedeem, and J. C. Fortney. 2002. "VA Community-Based Outpatient Clinics: Cost Performance Measures." *Medical Care* 40 (7): 587–95.
- MacNabb, L. 2003. "Application of Cluster Analysis towards the Development of Health Region Peer Groups." Statistical Society of Canada. Annual Meeting, *Proceedings of the Survey Methods Section*: 85–90.
- Petersen, L. A., K. Pietz, L. D. Woodard, and M. Byrne. 2005. "Comparison of the Predictive Validity of Diagnosis-Based Risk Adjusters for Clinical Outcomes." *Medical Care* 43 (1): 61–7.

- Romano, P. S. 2004. "Peer Group Benchmarks Are Not Appropriate for Health Care Quality Report Cards." *American Heart Journal* 148 (6): 921–3.
- SAS. 2002. SAS Institute, Inc., Cary, NC.
- Sharma, S. 1996. *Applied Multivariate Techniques*. New York: Wiley & Sons.
- Shen, Y., A. Hendricks, S. Zhang, and L. E. Kazis. 2003. "VHA Enrollees' Health Care Coverage and Use of Care." *Medical Care Research and Review* 60 (2): 253–67.
- Solucient. 2006. "Top Hospitals Reports." [accessed on July 14, 2008]. Available at <http://www.100TopHospitals.com>
- Statistics Canada. 2002. "Health Indicators." Available at <http://www.statcan.ca/eng/lidh/freepub/82-221-XIE/01201/about.htm>
- Stefos, T., N. LaVallee, and F. Holden. 1992. "Fairness in Prospective Payment: A Clustering Approach." *Health Services Research* 27 (2): 215–37.
- Wasserman, J., J. Ringel, K. Ricci, J. Malkin, M. Schoenbaum, B. Wynn, J. Zwaninger, S. Newberry, M. Suttorp, and A. Rastegard. 2003. *An Analysis of Potential Adjustments to the Veterans Equitable Resource Allocation (VERA) System*. Santa Monica, CA: RAND.
- Wasserman, J., J. Ringel, K. Ricci, J. Malkin, B. Wynn, J. Zwaninger, S. Newberry, M. Suttorp, and A. Rastegard. 2004. *Understanding Potential Changes to the Veterans Equitable Resource Allocation (VERA) System: A Regression-Based Approach*. Santa Monica, CA: RAND.
- World Health Organization. 2007. *Performance Assessment Tool for Quality Improvement in Hospitals*. World Health Organization Office for Europe.
- Wright, S. M., L. A. Petersen, R. P. Lamkin, and J. Daley. 1999. "Increasing Use of Medicare Services by Veterans with Acute Myocardial Infarction." *Medical Care* 37 (6): 529–37.
- Zodet, M. W., and J. D. Clark. 1996. "Creation of Hospital Peer Groups." *Clinical Performance and Quality Health Care* 4 (1): 51–7.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.