Appendix D. The High Performer Study Algorithm

Appendix D Contents:

Overview D-3

Summary of the High Performers Special Study Algorithm D-14

CMS Peer Review Recommendations for the Final HPSS Algorithm D-17

Ranking Performance in U.S. Nursing Homes Using a Modification of the CMS High-Performer Algorithm D-20

Overview

Overview

Background

The purpose of the algorithm in the High Performer Special Study (HPSS) was to distinguish high performers from all other providers nationally. The algorithm needed to meet three criteria: provide an objective measure of clinical performance, withstand CMS and QIO peer review, and receive stakeholder consensus for validity.

Algorithm development began with the findings from the extensive "State of the Art" literature review. Key components of the algorithm needed to (a) focus only on clinical performance; (b) consider both absolute and relative performance; (c) accommodate a broad scope of quality indicators and multiple clinical conditions; (d) reward stability in performance over time; (e) delineate providers by services provided, regardless of size or geographic location; and (f) have face validity.

Initially, eight principles of high performer/performance were identified, with four eventually being adopted after discussion with the CMS team members. The following four guiding principles formed the basis of the algorithm:

1. Hospitals that provide the same service should be held to the same standard, regardless of size or geographic location.
2. Performing well on an indicator on which others perform poorly (a rarely achieved indicator), while performing as well as any on an indicator on which others do well (a frequently achieved indicator), is basic to the definition of high performance.
3. High performance must incorporate sustained excellence over time.
4. Calculations used to identify high-performing institutions should be easily comprehensible.

Criteria of High Performance

In addition to the guiding principles, the literature review also led to the initial criteria incorporated into the high performer algorithm. After piloting the criteria with test datasets and seeking input from a CMS peer review panel, there were follow-up discussions with members of the panel to refine the criteria and improve the algorithm.

CMS approved the final algorithm criteria in October 2004. Applying the following criteria, the algorithm identifies hospitals that:

- Achieve a performance score of at least 90 percent on at least two indicators—where there are three or more indicators.
- Achieve a 95th percentile score on the rarely achieved indicator(s).
- Achieve a 95th percentile score on the frequently achieved indicator(s) or achieve ≥ 95 percent when the median is ≥ 90 percent on a frequently achieved indicator.

- Achieve a 90th percentile score on all "medium" indicators, those between the rarely achieved and frequently achieved indicators.
- Sustain the above four standards for two consecutive six-month periods.

Algorithm Formula

The algorithm was designed to accommodate a variety of health care settings and quality indicators. It is a flexible model that can be expanded or contracted, depending on the number of clinical conditions and quality indicators of interest. The final approved algorithm formula used for the HPSS in the hospital setting is based on three national CMS and JCAHO inpatient clinical conditions—acute myocardial infarction (AMI), heart failure (HF), and pneumonia (PN)—and 20 quality indicators across the three clinical sets:

$$\begin{aligned} HP = \; & (t_1 AMI_{a,b,c,d} + t_1 HF_{a,b,c,d} + t_1 PN_{a,b,c,d}) \; + \\ & (t_2 AMI_{a,b,c,d} + t_2 HF_{a,b,c,d} + t_2 PN_{a,b,c,d}) \; + \\ & (t_{1*2} AMI_{a,b,c,d} + t_{1*2} HF_{a,b,c,d} + t_{1*2} PN_{a,b,c,d}) \; = \; 36 \text{ maximum points} \end{aligned}$$

Where:

HP = High Performer Score, from 0 to a maximum of 36 points
t_1 = Time period 1, first six-month study period
t_2 = Time period 2, second six-month study period
$AMI_{a,b,c,d}$ = 4 HP criteria applied to set of six AMI indicators
$HF_{a,b,c,d}$ = 4 HP criteria applied to set of four HF indicators
$PN_{a,b,c,d}$ = 4 HP criteria applied to set of PN indicators

HP criteria are:

a = Hospital receives one point for achieving 90 percent on at least two of the topic-specific indicators
b = Hospital receives one point for achieving 95th percentile on the rarely achieved topic-specific indicators
c = Hospital receives one point for achieving 95th percentile (or ≥ 95 percent when the median is ≥ 90 percent) on the frequently achieved topic-specific indicators
d = Hospital receives one point for achieving 90th percentile on the medium topic-specific indicators
t_{1*2} = Hospital receives one point if the same HP criteria ($_{a,b,c,d}$) were achieved in each time period (t_1 and t_2) for the same topic (AMI, HF, PN)

A one-page summary of the High Performer Algorithm is presented in the next section of this appendix. It includes the above information and a list of caveats and limitations identified in the algorithm. Copies of the one-page summary, as well as earlier versions, were distributed throughout CMS and the QIO community to solicit input and recommendations.

Development of the Algorithm

Initial tests of the algorithm were conducted using data submitted to the QIO Clinical Warehouse (QCW) by 57 Arizona hospitals during calendar year 2003. Hospitals volunteered to report this data

publicly, in coordination with the national Hospital Quality Alliance (HQA), beginning with July 2002 data. Hospital performance in three clinical topics was assessed—acute myocardial infarction (AMI), heart failure (HF), and pneumonia (PN).

In order to better develop and test the algorithm, the Arizona dataset was expanded to include three additional states. Arizona's two partners in the CMS three-state pilot, Maryland and New York, along with Florida, each ran common SAS code on their own state QCW data. This allowed HSAG to work with a combined dataset that included nearly 400 hospitals before the national dataset was requested.

In early October 2004, HSAG contacted the Task Hospital Reporting QIOSC, custodians of the QCW, to create a national dataset for the project. Once the national dataset was received and verified for data completeness, an analytic file was created. The SAS code that HSAG developed was then in place to run the U.S. data for future QIO and special study requests. Since the submission deadline for 2nd Quarter 2004 discharges was November 15, 2004, it was decided to wait and acquire the latest available data to better reflect hospitals' "real-time" performance.

The study period spanned 12 months of inpatient discharges from July 2003 through June 2004. A total of 4,620 hospitals submitted data during the period, representing all 50 states, the District of Columbia, and Puerto Rico. Hospitals that submitted fewer than 20 eligible cases for the 12-month period in all three clinical topics were excluded. The final study sample, after excluding 753 hospitals with less than 20 eligible cases in each clinical topic, was 3,867 hospitals. A mean of 129 hospitals per state submitted data to the QCW, ranging from 5 in Delaware to 317 in California.

Initially 20 inpatient quality indicators across the three clinical topics were considered for use in the model: AMI (8), HF (4), and PN (8) (Table D-1). Four indicators were removed from analyses due to the low number of hospitals (< 25 percent) that submitted data for those measures; these included appropriate antibiotic selection for PN, influenza vaccination for PN patients, timely thrombolytic for AMI, and timely PTCA for AMI.

Table D-1. List of 20 Inpatient Quality Measures Across Three Clinical Conditions

AMI	Heart Failure	Pneumonia
Aspirin at arrival*	Discharge instructions	Antibiotic < 4 hours*
Aspirin at discharge*	LVF assessment*	Antibiotic selection†
ACEI for LVSD*	ACEI for LVSD*	Blood culture < 24 hours
Smoking cessation	Smoking cessation	Blood culture pre-antibiotic
Beta-blocker at arrival*		Influenza vaccine†
Beta-blocker at discharge*		Pneumococcal screen*
Timely thrombolytic†		Smoking cessation
Timely PTCA†		Oxygenation*

** Indicates "10 Measure Starter Set" for the Hospital Quality Alliance (voluntary public reporting).*

† Quality Indicators excluded from the study due to limited hospital cases/submissions.

Analytic and clinical staff from HSAG and the University of Alabama—Birmingham (UAB) conducted and reviewed the preliminary analyses associated with individual indicators, condition-specific aggregate statistics, and an overall grand summary statistic. In the majority of indicators, wide variation in hospital performance was observed between individual indicators and across indicators. Both of those analytical considerations are key components of the algorithm design. Correlations in performance between clinically similar indicators were also observed (e.g., aspirin at arrival, and beta-blockers at arrival). The initial findings from the Arizona data on quality indicators that were used to develop the algorithm were later confirmed with the four state and national datasets.

In order to apply the algorithm criteria, it was necessary to group the indicators into the frequently achieved, rarely achieved, and remaining indicators within each clinical condition. Categorization of the indicators was driven solely by the data, based on indicator performance rates, variation across hospitals, and cluster analyses (Figure D-1). Depending on the particular dataset, it may not be possible to define three distinct groups in one or more clinical topics.

Within a clinical topic, indicators were compared by performance and a variation metric that was the difference between an indicator's rates at the 95th and 50th percentile. Cluster analysis was used to group indicators and determine the statistical difference between indicators. The HPSS indicator groups based on the full 12 months of national hospital data are listed in Table D-2.

Indicator groups can be expected to vary over time as performance improves and variation is reduced across providers. The algorithm design allows for this, and the potential score and the overall cumulative score within a clinical topic is simply adjusted accordingly.

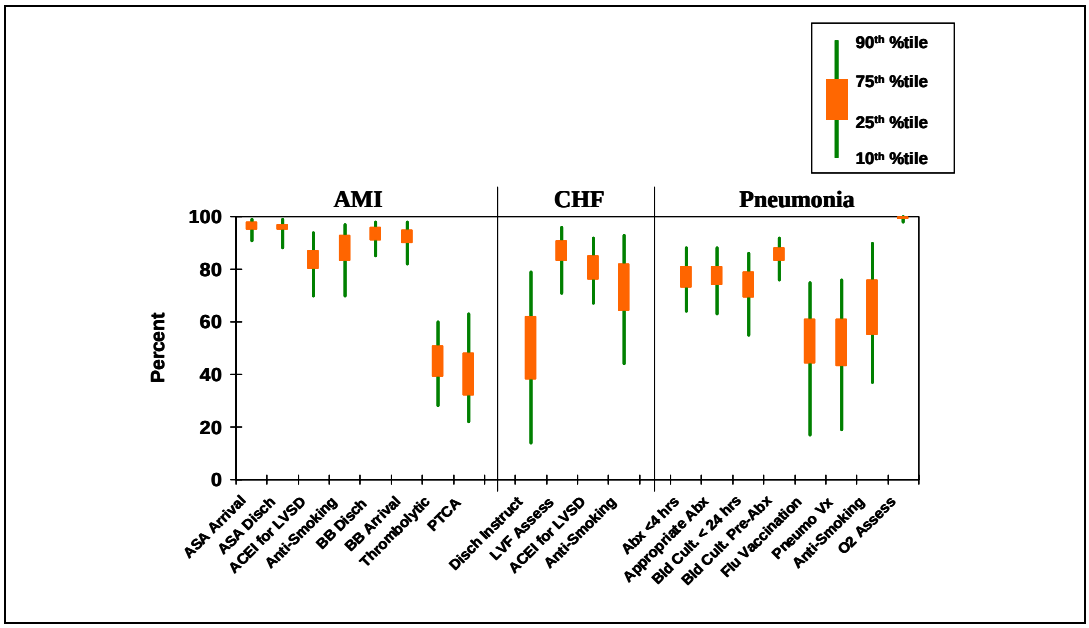


Figure D-1. **Variation in Hospital Rates by Individual Quality Indicator, QIO Clinical Warehouse, July 2003–June 2004**

Table D-2. Quality Indicator Assignment for High Performer Algorithm, QIO Clinical Warehouse Data, July 2003–June 2004	
Quality Indicator	Indicator Group
AMI	
Aspirin at arrival	Frequent
Aspirin at discharge	Frequent
Beta-blocker at arrival	Remaining
Beta-blocker at discharge	Remaining
Smoking-cessation counseling	Rare
ACE-I for LVSD	Rare
Thrombolytic agent	Excluded—Too few reporting
PTCA < 90 minutes	Excluded—Too few reporting
HF	
LVF assessment	Frequent
ACE-I for LVSD	Frequent
Smoking-cessation counseling	Remaining
Discharge instructions	Rare
PN	
Oxygenation assessment	Frequent
Blood cultures before antibiotic	Remaining
Antibiotic < 4 hours	Remaining
Blood cultures < 24 hours	Remaining
Smoking-cessation counseling	Rare
Pneumococcal screening/vaccination	Rare
Appropriate antibiotic	Excluded—Too few reporting
Influenza vaccination	Excluded—Too few reporting

National Algorithm Results

The final algorithm applied to the national HQA data clearly identified hospitals that performed significantly above the national norm. The distribution of hospital scores was highly skewed, with a range between no points scored and a maximum of 24 points (Figure D-2). One-half of the nation's hospitals scored less than 3 points, resulting in an extremely small number of high performing hospitals. Indeed, the score at the 95th percentile was just 11 points. Although no hospital approached the algorithm maximum of 36 points, 45 of the nation's hospitals (1 percent) scored between 16 and 24 points (Figure D-2). The top six hospitals in the nation, those with scores between 20 and 24 points, were identified for the qualitative phase of the study.

The distribution of mean algorithm scores by state is presented in Figure D-3. Each group of similarly shaded states represents a quartile range in algorithm points. States with no shading were in the top quartile. No evident pattern was observed in the geographical distribution of algorithm scores by state.

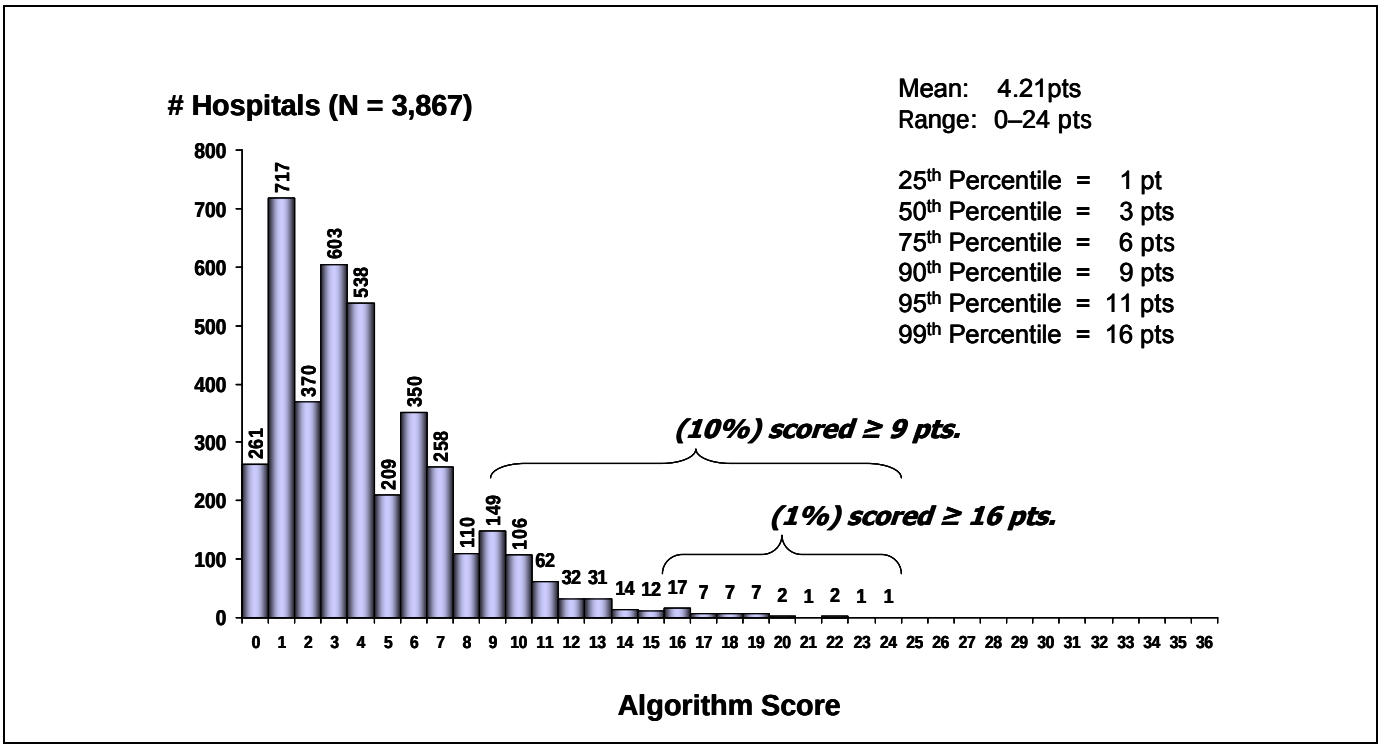


Figure D-2. **Frequency Distribution of Hospital High Performer Algorithm Scores, QIO Clinical Warehouse, July 2003–June 2004**

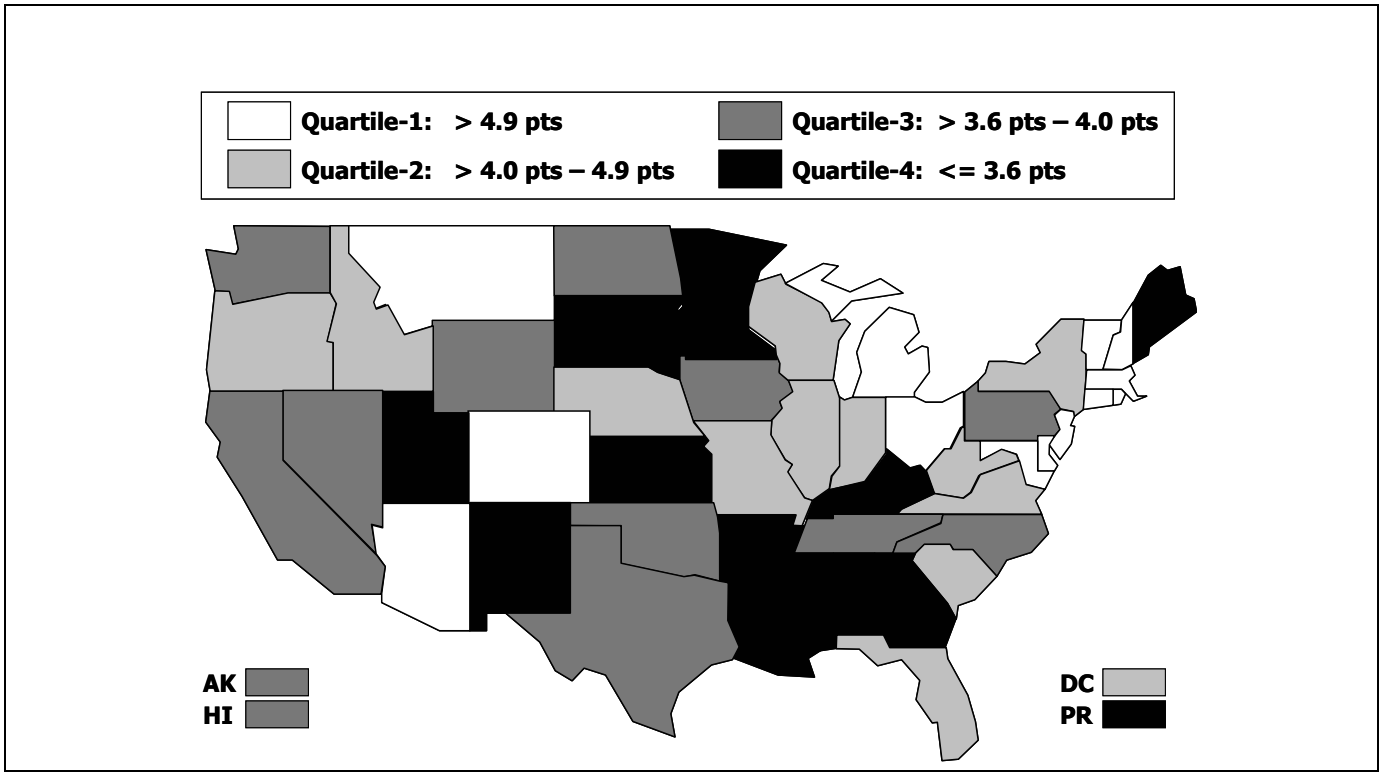


Figure D-3. **Geographical Distribution of Algorithm Scores by State**

Validation of the Algorithm

One of the guiding principles of the algorithm was that it should be transparent and easy to comprehend. Consequently, it was desirable to test the algorithm against established statistical models with more scientific rigor. HPSS team members from the UAB conducted independent analyses to validate the algorithm with a variety of statistical models.

There are basically three types of scaling methods in common use. The first is factor analysis with or without rotation of the resulting factor loadings. This exists in a number of forms, one of which was used here, that of principal axis factor analysis with varimax rotation. The second is called multiple correspondence analysis. It has three commonly used forms, two of which were appropriate in the HPSS application. They include general multiple correspondence analysis and a "fuzzed" or "fuzzy" multiple correspondence analysis. The last is a method called multidimensional scaling, which has at least three main forms, two of which were used to test the HPSS algorithm. They are the classical form of Multidimensional Scaling (MDS) and a so-called non-metric version.

In all, three types of scaling methods were used: (1) Principal Axis Factor Analysis (PFA) with Varimax Rotation; (2) Multiple Correspondence Analyses (MCA); and (3) Multidimensional Scaling (MDS). Two variations of both the MCA and MDS were also employed, accounting for five separate statistical comparisons. The results indicated a statistically significant intraclass correlation ($p < .0001$) between the HPSS algorithm and both variations of the MDS and the PFA.

In addition to statistical modeling to validate the algorithm, HSAG received feedback on algorithm transparency and face validity from audiences at two HPSS presentations. Ninety percent of the respondents at the National Symposium for Healthcare Executives (NSHE) and 100 percent of 2005 QualityNet Conference respondents agreed that the algorithm was easily understood. In fact, 88 percent of the hospital administrators at the NSHE conference desired to know their own hospital's performance score. A third presentation and source of audience feedback, the National Association for Healthcare Quality conference, was scheduled for New Orleans but was cancelled due to Hurricane Katrina.

Structural Measures to Categorize High Performers

In order to identify a matching set of non-high performers (NHPs) to compare to the high performers (HPs) in the qualitative portion of the study, structural measure data was necessary. As part of the Three-State Pilot Project, a Structural Measures Work Group for Public Reporting, led by CMS, reviewed potential data sources for posting hospital characteristics on the Hospital Compare section of the CMS beneficiary Web site (www.medicare.gov). The workgroup concluded that the CMS Healthcare Cost Reporting Information System (HCRIS) was the most reliable, up-to-date, and complete national source of structural measures and hospital characteristics data for U.S. hospitals.

HSAG worked with CMS to acquire the most recent HCRIS data available (September 2004) to supplement the HPSS hospital performance data. Viable HCRIS variables were used to describe each of the six HPs and create six corresponding pools of similarly characterized hospitals that were among the NHPs. The HCRIS variables used to categorize the hospitals are identified in Table D-3.

Table D-3.	Selected HCRIS Variables Used to Match Non-High Performers with High Performers
	State Urban/Rural Ownership/Type of Control Number of Beds Bed Days Available Total Inpatient Days Medicare Inpatient Days Proportion of Medicare (Medicare to Total Inpatient Days) Intern-Resident FTEs Teaching (CMS Approved) Occupancy Rate (Inpatient Days / Bed Days Available) Referral Center Heart Transplant Center

Four additional HCRIS variables were considered but discounted for various reasons: Critical Access Hospital (CAH) designation (none were HPs); if rural, beds < 100 (only 3 entries); and sole community hospital (only 36 entries).

HCRIS data were also supplemented with additional statistics and, when available, hospital ratings/rankings from the US News Web site and JCAHO. Key variables from these data sources included whether hospitals were part of a health care system and whether they were JCAHO accredited.

Not only were the pools of potential NHPs made up of hospitals with similar characteristics as the six HPs, but hospitals eligible for a "pool" had the same scoring potential as the six HPs. Therefore, the base criteria for the NHPs were that they obviously had a low score (< 6 points), but they also had the potential to score as high as the six HPs. A third, required criterion was that an NHP needed to be in the same state as the corresponding HP. This key criterion greatly limited the potential NHPs matched to the HPs. The hospitals represented five states: Florida, Illinois, Indiana, New Jersey, and North Carolina.

Peer Review and Consensus

Throughout the development of the algorithm, HSAG sought advice, consultation, and recommendations from CMS and the QIO community. On June 30 and July 1, 2004, CMS hosted a peer review meeting in Baltimore, Maryland. This strategic meeting was held to review the model based on the four-state dataset, discuss changes, resolve issues, and provide direction. Participants included key staff from the CMS Central Office and three CMS Regional Offices (Seattle, Dallas, and Boston), two Arizona hospital representatives, and team members from HSAG, CMS, and the University of Alabama at Birmingham (UAB).

As follow-up to the meeting, three conference calls with CMS, HSAG, and UAB were held to summarize the input from the peer review meeting and finalize the recommendations and refinements to the model. Section D2 of this appendix, CMS Peer Review Recommendations for the Draft HPSS Algorithm, summarizes the final seven recommendations. In addition to CMS Peer

Review, HSAG sought the advice of individual peer panel members and received input from individuals in the QIO community. The algorithm was presented with results from both the four-state test dataset and the final national dataset at the September 2004 QualityNet Conference and the 2005 American Health Quality Association (AHQA) Technical Conference, respectively.

Single-Clinical-Topic Algorithm Model for Pneumonia

Originally, HSAG proposed an "A" model and a "Z" model algorithm. The latter is the more comprehensive model—with three inpatient clinical topics—that was used to score the nation's hospitals for the HPSS. However, the Z model algorithm does not recognize hospitals that admit few or no patients in some clinical areas. For the purpose of the HPSS, it was desirable that the algorithm identify HPs that were able to perform well across multiple areas, increasing the likelihood that performance was attributable to a system-wide infrastructure. However, CMS wanted to ensure that the algorithm could be adapted to fit a variety of provider types. Clearly, there are best practices to learn from hospitals that do not provide all services but excel in the services that are provided.

In order to test the A model on one clinical topic, PN was selected. More patients are admitted with PN than either AMI or HF. Moreover, many smaller hospitals treat PN patients but transfer AMI and HF patients. The Z model algorithm was easily adapted to analyze the results of just the four PN criteria in each of the two time periods, producing a model with a maximum score of 12 points (4 points for each criterion in each of the two time periods, plus a possible 4 points for sustainability on all criteria over both time periods).

There were high-performing hospitals that scored well with both algorithms. As expected, there were also many rural hospitals that were HPs for their PN care that were not recognized as HPs with the Z model. Among the top 54 hospitals scored using the full three-topic model, 20 new hospitals were listed when only PN care was examined. Conversely, 34 of the 54 high-performing hospitals scored in the top tier, whether they were scored on all three topics or only the single topic. One model does not fit all provider types or datasets. The algorithm is flexible and can be modified to accommodate a wide variety of provider and indicator parameters.

Use of the Algorithm for Nursing Homes

CMS was interested in testing the algorithm on nursing home data. Although the HPSS was focused on the hospital setting and qualitative data collection from the hospital HPs, it was timely for CMS to score nursing homes (NHs) with the algorithm and substantiate other independent NH findings. It was also an opportunity to further test the validity and versatility of the model in another health care setting with its own unique quality measures. Consequently, in December 2004, HSAG worked closely with the NH Data QIOSC (the Colorado QIO) to run modified SAS code on Minimum Data Set (MDS) summary files for U.S. nursing homes.

Unlike the hospital setting with distinct sets of clinical indicators, the nursing home quality measures are not as easily categorized. Nonetheless, because it was another test of the algorithm design, CMS was able to identify two logical measure sets among seven of the NH quality measures for testing purposes. A brief summary of the NH results completed in January 2005 is presented

later in this appendix. The conclusion and caveat of the NH test was that, while the algorithm worked well in the nursing home setting, the final methodical version used for the hospital setting and three clinical topics underwent extensive peer review and validation. Use of the algorithm with MDS data appears to be feasible, but only with similarly refined methods and review.

Summary of the High Performers Special Study Algorithm

Summary of the High Performers Special Study Algorithm

PRINCIPLES OF HIGH PERFORMANCE

1. Hospitals that provide the same service should be held to the same standard, regardless of size or geographic location.
2. Performing well on an indicator on which others perform poorly (a rarely achieved indicator), while performing as well as any on an indicator on which others do well (a frequently achieved indicator), is basic to the definition of high performance.
3. High performance must incorporate sustained excellence over time.
4. Calculations used to identify high-performing institutions should be easily comprehensible.

CRITERIA OF HIGH PERFORMANCE

a. Achieves a performance score of at least 90 percent on at least two indicators, where there are three or more indicators.
b. Achieves a 95th percentile score on the rarely achieved indicator(s).
c. Achieves a 95th percentile score on the frequently achieved indicator(s) or achieves ≥ 95 percent when the median is ≥ 90 percent on a frequently achieved indicator.
d. Achieves a 90th percentile score on all remaining indicators, those between the rarely achieved and frequently achieved indicators.
e. Sustains these standards for two consecutive six-month periods.

ALGORITHM FORMULA

$$\mathbf{HP} = (t_1AMI_{a,b,c,d} + t_1HF_{a,b,c,d} + t_1PN_{a,b,c,d}) + (t_2AMI_{a,b,c,d} + t_2HF_{a,b,c,d} + t_2PN_{a,b,c,d}) + (t_{1*2}AMI_{a,b,c,d} + t_{1*2}HF_{a,b,c,d} + t_{1*2}PN_{a,b,c,d}) = \mathbf{36}$$

Where:

- HP = Score of 36 for the maximum a high-performing hospital can achieve
- t_1 = time period 1, first 6-month study period
- t_2 = time period 2, second 6-month study period
- $AMI_{a,b,c,d}$ = four HP criteria applied to set of AMI indicators
- $HF_{a,b,c,d}$ = four HP criteria applied to set of HF indicators
- $PN_{a,b,c,d}$ = four HP criteria applied to set of PN indicators

HP criteria are:

- a = Hospital receives one point for achieving 90 percent on at least two of the topic-specific indicators
- b = Hospital receives one point for achieving 95th percentile on the rarely achieved topic-specific indicators
- c = Hospital receives one point for achieving 95th percentile (or ≥ 95 percent when the median is ≥ 90 percent) on the frequently achieved topic-specific indicators
- d = Hospital receives one point for achieving 90th percentile on the remaining specific indicators
- t_{1*2} = Hospital receives one point if the same HP criteria ($_{a,b,c,d}$) were achieved in each time period (t_1 and t_2) for the same topic (AMI, HF, PN)

IDENTIFIED ALGORITHM LIMITATIONS

1. Impact of sampling error on high performance determination: the potential exists for mistakenly excluding high performers due to unaccounted for sampling/measurement error.
2. Providers that do not care for patients with one or more of the identified clinical conditions (e.g., heart specialty hospitals, or smaller hospitals that only stabilize and transport AMI patients) will be unable to earn the total 36 points and, therefore, may not qualify as high performers even if they achieve a perfect score on the services they provide.
3. Small numbers for an indicator (i.e., < 20 eligible denominator cases per time period): the formula does not adjust for small numbers or missing indicators and/or topics.
4. The formula presents a fixed absolute standard for a high performer, thus minimizing the possibility of any hospital achieving high performer status.
5. The current approach does not provide for identifying low performers: there is no relative value given to the scores of the non-high performers to evaluate their performance in relation to one another.

These limitations exist as the result of employing a simplified approach (greater face validity, less

scientific rigor) in order to make the algorithm easily comprehensible. Future work can address these limitations and present known statistical methods (which are not as easily comprehensible) that have been used successfully to overcome these limitations in other studies and settings.

CONTACT

Health Services Advisory Group: Andrea B. Silvey, PhD, MSN, (602) 665-6135, asilvey@hsag.com.

CMS Peer Review Recommendations for the Final HPSS Algorithm

CMS Peer Review Recommendations for the Final HPSS Algorithm

Further Expand "Criterion E," Sustained Excellence

It was recognized at the CMS peer review meeting that the model did not reward high-performing hospitals appropriately for sustaining their results over consecutive time periods. The model scored points independently across the two time periods, regardless of the consistency in criteria or clinical topic.

An equitable solution was proposed that still maintained the transparency of the model. In addition to the possible 24 points in the model, another potential 12 points for sustainability were added, resulting in a possible 36-point total. The 12 additional points are based on four criteria multiplied by three topics. If any criterion within any topic scores a point in both time periods 1 and 2, a point for sustainability will be earned.

Create a Single-Topic Model

The model was not designed to learn from high-performing hospitals that do not treat all clinical topics. Peer review recognized that it is possible to learn from a hospital that does not provide all services but excels in the services that it does provide. It was agreed to test a PN-only model, since more patients are admitted with PN than either AMI or HF. Moreover, many smaller hospitals treat PN patients but transfer AMI and HF patients.

Keep All Indicators Weighted Equally

It was agreed that it was beyond the scope of the project to develop a weighting scheme for the quality indicators that would be acceptable to all clinicians and administrators. More importantly, the information to be gleaned from high-performing hospitals and their capacity for system change should be independent of the clinical importance of the indicators.

Do Not Add an "Improvement" Component

During peer review, the point was raised that the algorithm does not measure a facility's improvement over time. The suggestion was made to include a measure of improvement (e.g., reduction in failure) between time periods 1 and 2. However, after considerable follow-up discussion, it was decided that the study should learn from those that have already achieved high performance and are maintaining it, not those currently moving toward high performance.

Do Not Retire High Compliance, Low Variation Indicators

The utility of the oxygenation assessment indicator for pneumonia patients in the model was questioned due to its extremely high rate of performance among the vast majority of hospitals. Meeting participants agreed that an indicator has little value to discern hospital performance when a mean indicator rate nears 100 percent and there is little difference in performance between the

highest and lowest rates. However, it was decided that, for the purpose of the study, all viable CMS indicators should be included in the model. Unless CMS were to remove the indicator from public reporting, it would continue to be used in the HPSS.

Include Absolute Threshold Values for Frequently Achieved Indicators

It is possible that hospitals performing at acceptable high standards on frequently achieved indicators are not recognized in the model. This is most likely to occur when their performance rate on frequently achieved indicators at the 95th percentile nears 100 percent. Until this type of indicator is retired, it was agreed that criterion "c" for frequently achieved indicators (see Section D1 in this appendix) would be modified to adjust for acceptable performance rates. When the median performance rate of a frequently achieved indicator reaches 90 percent, a hospital will now meet the criteria if it is in the 95th percentile (original criterion) or its rate is below the 95th percentile but has a rate > 95 percent.

Refine the Method of Assigning Indicator Groups

The model, as presented for peer review, used arbitrary percentile cut-offs (i.e., quartiles) to categorize and separate the indicators into the three indicator groups. It was evident in the results that were presented for peer review that the performance rates were not normally distributed within all three topics. It was agreed that the categorization of indicator groups needed to be data driven and derived from the distribution of the indicator rates. Indicators should not be assigned to separate indicator groups when there are no significant differences between them.

The flexibility of the model will accommodate varying types of distribution within a topic. However, this may result in less than three indicator groups per topic. Consequently, the maximum possible score overall for the algorithm may vary, depending on the distribution of the rates and indicator assignment. This will not be determined until the national dataset is used to define the final indicator groups within each clinical topic.

Ranking Performance in U.S. Nursing Homes Using a Modification of the CMS High-Performer Algorithm

Ranking Performance in U.S. Nursing Homes Using a Modification of the CMS High-Performer Algorithm

Health Services Advisory Group (HSAG) has been developing a multidimensional algorithm to "score" high performers in the hospital setting. However, the algorithm was designed to be adaptable to other settings and across a variety of measures. HSAG worked closely with the Nursing Home Data QIOSC (the Colorado QIO) to use MDS data to apply the algorithm to the nation's nursing homes. The following is a summary of the algorithm results tested with U.S. nursing home data. The data is preliminary, and the algorithm modified for the nursing home measures has not been refined or subjected to peer review. It does, however, show that the use of such a model is feasible and can be utilized across health care settings.

Parameters

Time Period:

The data spans two 12-month time periods

t1 = July 2002 through June 2003

t2 = July 2003 through June 2004

Measures:

In order to test the application quickly, CMS defined two tiers of measures among seven nursing home measures.

Tier-1 – Chronic Pain, PAC Pain, Pressure Ulcer, and Restraints

Tier-2 – ADL Loss, PAC Delirium, and PAC Walking Improvement

Nursing Homes:

The NH Data QIOSC supplied information on a total of 17,376 nursing homes. There were 450 homes excluded that had insufficient cases across all seven measures in both time periods.

Final Total: 16,926 Nursing Homes

Algorithm Scores:

Nursing homes earned points for meeting specific performance criteria in each time period and sustaining that performance over both time periods. Therefore, the higher the performance, the higher the score. The number of points within a tier or clinical topic is data-driven and dependent on the distribution of the measures and how they cluster together.

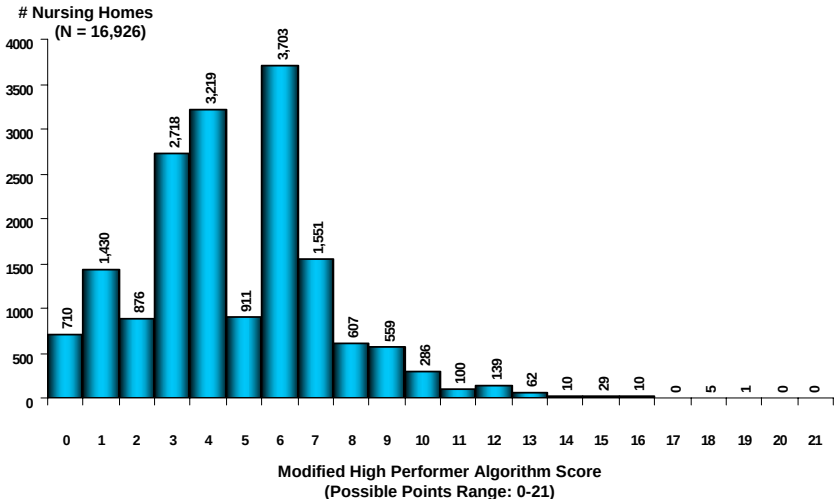
Tier-1: Possible algorithm scores range from 0–9 points

Tier-2: Possible algorithm scores range from 0–12 points

Total: Possible algorithm scores range from 0–21 points

Results

Overall Scores across both tiers and all seven measures ranged from 1 to 19 out of a possible 21 points. The high score of 19 points was achieved by just one nursing home in the nation. Moreover, there were just 5 nursing homes among 16,926 with a score of 18 points (Figure D-4). The mean score overall was 4.7, with a median score of 4.



Note: The model examines performance in seven nursing home measures.

Figure D-4. **Distribution of U.S. Nursing Home Performance Scores**

The maximum possible points for **Tier-1 Measures** was nine, and was achieved by 18 nursing homes. There were no nursing homes with eight points for Tier-1 Measures, and 78 nursing homes scored seven points. One-half of the nursing homes (8,140 homes) earned three points (Table D-4).

Table D-4. **Tier-1: Chronic Pain, PAC Pain, Pressure Ulcers, and Restraints**

Tier 1 Patients	Frequency	Percent	Cumulative
0	2,680	16.74	16.74
1	2,904	18.14	34.89
2	106	0.66	35.55
3	8,140	50.86	86.41
4	1,463	9.14	95.55
5	88	0.55	96.10
6	529	3.31	99.40
7	78	0.49	99.89
9	18	0.11	100.00
TOTAL	16,006	100.00	

Only 2 nursing homes among 16,926 achieved the 12-point maximum for the **Tier-2 Measures**. There were no nursing homes with 11 points and just 23 nursing homes with a score of 10 points. However, 222 nursing homes achieved 9 points, accounting for 1.4 percent of the nation's nursing homes. A total of 3 points was achieved by 6,746 nursing homes (42 percent) (Table D-5).

Table D-5. **Tier-2: ADL Loss, PAC Delirium, and PAC Walking Improvement**

Tier 2 Patients	Frequency	Percent	Cumulative
0	2,541	15.99	15.99
1	2,998	18.87	34.86
2	564	3.55	38.41
3	6,746	42.45	80.86
4	1,448	9.11	89.97
5	676	4.25	94.23
6	343	2.16	96.39
7	292	1.84	98.23
8	35	0.22	98.45
9	222	1.40	99.84
10	23	0.14	99.99
12	2	0.01	100.00
TOTAL	15,890	100.00	

The algorithm results, particularly within individual tiers, indicate a very skewed distribution in nursing home performance. In both tiers, more than one-third of the nursing homes scored less than two points. Likewise, a large cluster of the nation's nursing home scores (approximately 50 percent) was observed in both tiers.

This work was merely an attempt to apply a variation of the high performer algorithm to nursing home data. The final methodical version used for the hospital setting and three clinical topics has undergone peer review and validation. However, with refined methods, this approach with nursing home data appears to be feasible.