

## **5-Star Quality Measures—Are they risk adjusted to “level the playing field” and allow fair/accurate comparisons across facilities?**

Christie L. Teigland, Ph.D., Director of Health Informatics Research, New York Association of Homes and Services for the Aging (NYAHS) and EQUIP for Quality®

In general, using these measures is highly problematic due to the lack of “risk adjustment.” While CMS states all the measures are risk adjusted, in fact the risk adjustments are very limited and do not begin address the most important patient issues related to these outcomes. The facts about each QM/QI used in the rating are discussed in detail below.

Because this is true, ***facility rates are largely determined by the types of residents they admit*** whereas if they were properly adjusted for resident population characteristics (and this is not rocket science, they have done this in acute care/hospital/surgery for years) you could truly compare one facility to another based on quality of care provided. Importantly, this is a disservice to the residents because many critical quality of care problems are missed or buried.

One of the leading academics and key developer of the MDS/QMs, Dr. Vince Mor, stated in 2006: *“The impetus to apply the emerging set of quality ‘tools’...may have outstripped the evidence for their valid application in selecting top providers or for rewarding their superior performance. Observed differences in quality measures could be a result of different providers serving different types of patients—making it difficult to determine whether differences are due to real differences in quality. Additionally, observed differences in provider quality measures may be due to how providers go about assessing their patients. Researchers and government have a responsibility to make sure that the technical aspects of the QMs being used to compare NHs are up to the challenge of being used both to stimulate the org changes needed to redesign care processes to improve care and to allow for legitimate and valid comparisons across providers.”* (Mor, V, Defining and Measuring Quality Outcomes in Long-Term Care, JAMDA, October 2006, 532-540.)

Dr. Robert Wachter, another leading expert in this area stated in 2006: *“The existing measures accurately reflect the immature state of the science of quality measurement, a result of our underinvestment in that science. Now that there is “real skin in the game” (pay for performance), it is vital to continue to refine the measures, making them more relevant to the care of complex patients with multi-dimensional illness and to improve our understanding of case-mix adjustment to facilitate apples-to-apples comparisons of outcomes.”* (Dr. Robert M. Wachter, Department of Medicine, University of California, San Francisco, JAMA, 2006;296:2439.)

**Indeed, the original QMs were never validated or designed to identify facilities providing superior care—there were intended as indicators of potentially poor care for further investigation, nothing more.**

**Aggregation Issue:** Beyond the problems with the measures, there are equally important and unanswered problems related to aggregating this many quality measures in this ad hoc fashion. The issue of aggregation of disparate measures of care (which have inherent problems standing alone, much less combined in some fashion) is very complex. The methodology for scoring and the results have NOT been validated to determine whether these aggregated

results indeed identify above average vs. poor performing nursing homes. There are many unresolved issues with aggregating QMs into one measure. Most often, facilities tend to do well in some areas and poorly in others depending on their focus and resident population, the good and bad average out and they look just “average.”

## Discussion of the 5 Star QMs: Facts About Risk Adjustment

### Long Stay Measures

One important inconsistency to understand is that the first two measures below—ADL change and mobility change—are treated differently in the Quality Measure score calculation in two ways (1) they are compared to the statewide rate, not the national rate like all the other QMs; and (2) they are each weighted 1.66 times more than the other Quality Measures.

This can have a big impact. The statewide rates on the ADL decline QM range from 10.6% in Oregon to 24.2% in North Dakota. The rates on the Mobility decline QM range from 10.2% in Oklahoma to 25.7% in Pennsylvania. Other QM rates similarly vary widely across states.

So the questions this raises are: (1) if residents’ likelihood of declining in function varies so widely across states, does not their likelihood of having other adverse outcomes included in the 5 Star System also vary as greatly? ***Facilities in states like North Dakota and Pennsylvania with very high rates of decline are held to far lower standards of quality than facilities in other states—does this make sense?*** A facility that ranks in top 20<sup>th</sup> percentile in PA may rank in the worst 20<sup>th</sup> percentile in OK. Does this not also impact negatively on the residents in those facilities since they are held to much lower standards of care? Is quality of care that different, or are resident populations really that different? If so, the likelihood of those residents having pressure ulcers or pain is probably much higher as well, yet those rates are compared to national standards.

**ADL change**—is NOT risk adjusted. There are a few exclusions which could be considered partial risk adjustment, but they do very little in practice.

- (1) The measure excludes residents who cannot decline further in function so you can’t measure decline. The result of this exclusion is that facilities that may provide very poor care and have far more residents who have declined to point of needed total assistance in the four activities of daily living (bed mobility, eating, transfer and toileting) are not considered in their rate, so ***these facilities are in fact “rewarded” for having most of their residents totally functionally dependent.*** They also get more reimbursement this way. What incentive do they have to try to improve residents’ functionality, which may be possible with the right nursing restorative programs and other interventions?
- (2) The measure also excludes residents who are comatose, have end-stage disease (6 months or fewer to live) or are in hospice. In practice, this excludes a very small number of people and is rarely coded. The average percent of comatose residents in NYS, for example is less than half of one percent--0.41%. The average percent of residents with end stage disease is 2.4% and residents in hospice 2.2%. So these exclusions do very little to affect any facility rate on this QM.

**Importantly, this measure does not risk adjust for diseases and conditions that are not in control of the facility, but put residents at far greater risk for functional decline.** For example, our research shows that residents with dementia decline much faster than cognitively intact residents; thus facilities with large dementia/Alzheimer's populations tend to have higher rates of ADL decline. Older residents and males are also much more likely to decline, so facilities with older populations or primarily male populations (like Veterans homes) are likely to fare worse on the QM. High BMI/obesity is a risk factor for functional decline, and resident populations in New York have had a large increase in BMI's over the past 10 years. Residents with diseases/conditions such as multiple sclerosis, Parkinson's, lung disease, antibiotic resistant infections, etc. are 2x more likely to have ADL decline. Thus, there are far more risk factors that should be considered if we want to compare facilities on their ability to prevent ADL decline.

**Mobility change**—is risk adjusted using a small number of risk factors (covariates).

- (1) The proportion of residents with a recent fall puts residents at higher risk for mobility decline (which is ability of resident to move in and around their room measured by the locomotion ADL (G1ea)). This means the more residents with falls, the higher we expect the facility mobility decline rate to be, thus their rate is risk adjusted down to account for this. Again, ***this can reward facilities with high rates of falls, which are often preventable and under the control of the facility.***
- (2) Residents needing extensive or total assistance in eating or toileting are considered higher risk and have same impact on the facility risk adjusted rate. It is not clear why needed assistance in eating or toileting should lead us to expect that residents are more likely not be able to move around their room, given the appropriate care and assistance. This risk adjustment is puzzling.
- (3) Uses the same exclusions as the ADL decline measure above, with the same limitations noted above.

Again, the measure does not include resident risk factors that might impact ability move around the room, such as balance/gait problems, osteoporosis/bone density related fractures, etc.

**Pressure ulcers**—is risk adjusted using “stratification method.” However, the “high risk” group is very narrowly defined. To be considered “high risk” you must need (1) extensive or total assistance in transfer or bed mobility; or (2) be comatose or have malnutrition. The latter two are rare occurrences, so really have no impact. Even CMS' own survey protocols cite many other risk factors for pressure ulcer development, including the biggest of them all—a history of pressure ulcers. Indeed, having a history of pressure sores is included in the short stay pressure ulcer measure risk adjustment, but not here—very inconsistent and puzzling. Other very critical risk factors that greatly increase the likelihood of developing a pressure ulcer include diabetes, peripheral vascular disease, CVA, hip fracture, bowel incontinence, desensitized skin...the list goes on and on. None of these conditions are considered as putting residents at high risk.

The short stay measure also risk adjusts for *limited* assistance in bed mobility as well as extensive/total--but not for any level of assistance in transfer as in the long stay (chronic care) measure above. The short stay measure is also risk adjusted for diabetes, peripheral vascular disease and low BMI. ***If these are important risk factors for short stay patients why not for long stay? The fact is these are all very important risk factors that are not***

***under control of facility staff, and can greatly increase likelihood of developing a pressure ulcer.***

Indeed, using the CMS stratified measure for pressure ulcers, if you rank facilities from worst to best and compare the top 20% to a performance ranking based on rates adjusted using the above risk factors, more than 40% of facilities who ranked in top 20% fall out—some to the bottom quartile, meaning they are doing a FAR WORSE job than expected given their resident population characteristics. ***We are in fact rewarding poor performance and/or facilities that do not admit high risk residents. Is this the message we want to send?***

**Pain**—has one risk factor (covariate) that attempts to risk adjust for cognitive impairment using one item from the MDS related to the residents ability to make independent decisions (B4=0 or 1). Our dementia research shows this item does NOT capture a large number of dementia residents—you can have a diagnosis of dementia or Alzheimer’s disease and not be able to communicate pain at all and not be considered “high risk” under this definition.

Furthermore, there are many, many conditions that put residents at risk for pain that are *not* considered. For example, if you have arthritis you are 80% more likely to have pain, a fracture you are 50-200% more likely to have pain, physical therapy you are 58% more likely to experience pain, conditions like edema or vomiting put you at 30-40% higher risk, end stage disease 70% more likely, a surgical wound more than 2x more likely, etc.

Finally, we find that facilities with a large proportion of dementia residents generally have *far lower pain rates* because pain is significantly under assessed for this population who cannot communicate their pain. ***So, you might look “above average” using CMS definition when in fact you are doing a very poor job of assessing and managing pain in cognitively impaired residents.*** And some facilities do not adequately assess pain on a regular basis, so they look good on this measure as defined. There are also many cultural and age related issues around reporting pain.

Facilities that do a good job at assessing pain, including pain in cognitively impaired residents WILL look far worse on this measure. Here is the language copied directly from Nursing Home Compare:

- ***“Comparing these percentages is different from the other measures because the percentages may mean different things.”***
- ***“Generally, a lower percentage on this measure is better. However, this isn't always true.***
  - *For example, two nursing homes could provide the same quality of care and have the same number of residents with pain. However, if one of the nursing homes does a better job checking the residents for pain, they could have a higher percentage on this measure.*
  - *Or, if for personal or cultural reasons, more residents in one of the nursing homes refuse to take pain medication, that nursing home's percentage would be higher.*
- ***In these examples, although the percentage for one nursing home is higher, it does not mean they are not providing good care.”***

Yet, they still use this QM in the 5 Star system!

**Physical restraints**—no risk adjustments or exclusions

**UTIs**— no risk adjustments or exclusions

**Long-term catheters**—risk adjusted for bowel incontinence and stage 3-4 pressure ulcer on prior assessment (which could be 3 months or more old), but no other factors that put residents at higher risk.

### **Short-Stay Prevalence Measures:**

**Delirium**—has 3 exclusions only for comatose, end stage disease and hospice, which have minimal impact as noted above. There is no risk adjustment for cognitive impairment/dementia which puts residents at far greater risk (4x times more likely to have delirium symptoms!); again, population characteristics and coding practices can greatly influence facility rates on this measure.

**Pain**—no risk adjustment or exclusions, not even for cognitive status as in long stay pain measure above.

**Pressure Ulcers**—the short stay QM is actually an incidence measure and is the best risk adjusted measure of all, but is still fraught with problems. It includes all the risk factors cited in the chronic care pressure ulcer QM discussion above. The measure counts (1) residents who had no pressure ulcer when admitted and have one on next assessment (which means if they missed a stage 1 red spot on 5 day assessment but caught it on 14 day assessment, it will count); and (2) residents admitted with a pressure ulcer at any stage that did not improve from 5 day to 14 day assessment. Since there can be as little as 3 days between these two assessments, if a resident has a Stage 3 or 4 pressure ulcer on admission, it is highly unlikely to be reduced in stage in a very short time period. Even a stage 2 pressure ulcer cannot heal within 3 days in many cases. And short stay assessments beyond the 14-day assessments don't count at all (e.g., 30 days, 90 days, etc.), so wound care centers who specialize in treating deep tissue wounds do not get credit for the wonderful job they do at healing these very complex wounds.

So, again, depending on type of short stay residents admitted, the timing of the initial two assessments, and the accuracy of coding, a facility can look good or bad on this measure regardless of the quality of care provided.