ACO Metrics for Recommended Changes

Metric	Discussion	Recommendations	Methodologic Flaw
Amb Sensitive Conditions for Admission (9,10)	Claims based measures are problematic if diagnosis codes are based on hospital claims – our review found 30% of hospital discharge diagnoses were incorrect, so we cannot rely on them.	Claims based measures should go into GRPO so we have an opportunity to review the denominator.	Non-level playing field: Denominator with variable accuracy
Medication Reconciliation (12)	All ACO's involved felt strongly that this was an extremely important area to measure. However, there are tremendously variable systems issues out of the direct control of each ACO. For example, the measure assumes that ACO's receive notification of discharges—some ACO are situated in areas dense with hospitals and only know when discharges happen from the home hospital. Other ACO's do not have access to accurate and timely discharge medication information.	There is no fair methodology for benchmarking this metric. We suggest reporting only or that each Pioneer beat its own standard by a negotiated rate for year-over-year improvement. We also strongly urge CMMI to require that hospitals notify ACOs of discharges, and to provide a clean discharge medication list to the patient.	Non-Level Playing Field: ACO's in areas with multiple hospital systems or no direct hospital linkages are disadvantaged; inconsistent discharge information from different hospitals
Falls Assessment (13)	The many geriatricians among are highly alert to falls risk. However, the utility of uniform application of falls risk assessment in most community dwelling elders is questionable. While the best predictor of a fall is a fall, there is no validated screening tool in the ambulatory setting that is a good predictor of falls risk. Furthermore, there is no evidence that generalized screening moves outcomes.	Given these very direct guidelines from the USPSTF, we believe that more information is needed before workflows and resources	No evidence- based benefit for in community dwelling elderly populations

	evidence that generalized screening moves outcomes. Moyer, et. al. **Prevention of Falls in Community-Dwelling Older Adults: U.S. Preventive Services Task Force Recommendation Statement**. Annals of Internal Medicine. May 29, 2012. "The USPSTF does not recommend automatically performing an in-depth multifactorial risk assessment in conjunction with comprehensive management of identified risks to prevent falls in community-dwelling adults aged 65 years or older because the likelihood of benefit is small."	are devoted to increasing a score on this measure. Hold in observation status until a validated falls assessment screening tool exists and a more specific population for the denominator (eg aged 75+) can be defined.	
Flu/Pneumovax Immunization (14, 15)	Depending on geographic location, immunizations are often given at work or in pharmacies. Claims data is not always be submitted, or submitted only with a significant claims lag—creating large structural measurement inaccuracies depending on the pattern of the location administration.	Please exempt from benchmarking. Consider using a negotiated minimum standard. Isn't there an existing standard? Med Adv? Could we propose using that?	Non Level Playing Field: claims rates differentially reflect actual implementation rates

Metric	Discussion	Recommendations	Methodologic Flaw
BMI Screening and Followup (16)	The rate reported for this measure does not help compare quality between all ACOs. This numerator has conditional criteria for compliance. A patient may be counted as non-compliant for this measure if their BMI was	We recommend creating disaggregated metrics to clarify what we are measuring.	Flawed Aggregation Logic Lack of evidence based benefit for screening and follow-up for elders with BMI <23

not measured OR may be non-compliant if the patient had an abnormal BMI for which they did not receive follow-up care OR patients may be non-compliant if they have an abnormal BMI but received follow-up care outside the 6 month window. An ACO that does not measure BMI on a consistent basis could have the same rate as an ACO that consistently does BMI measurements but does not consistently provide or document the proscribed follow-up care for patients with abnormal BMI measurements and either of these ACOs could have the same rate as an ACO that has a standard to provide follow-up care in 8 months (for example) rather than 6 months.

These are clearly different quality of care issues, however this measure treats them as if they are the same.

From an ACO perspective, it is difficult to know what actions to take to improve this rate solely based on the rate itself because there are multiple reasons a patient can be non-compliant for this measure.

Additionally, regarding the portion of this metric devoted to those elderly with a BMI <23, CMS cites a single observational study which suggests that patients below this threshold may have nutritional deficiencies. This is not sufficient evidence to justify nutrition interventions for almost half of all seniors, and is not consistent with the community standard of care.

The initial sub-measure should be one of **assessment**: a numerator that is the number of patients with a measured BMI in the measurement period and the denominator is patients 18-74. If this were to be made equal to HEDIS, we could begin benchmarking this sub-metric.

An **obesity** sub-measure should look at Numerator = patients with BMI >=30

There is no clear mortality risk for those in the 25-29 range.

An **underweight** sub-measure should look at those with BMI <23, and compare to their *weight in the following measurement period (2013 reporting)*. While 23 is a perfectly acceptable BMI for stable healthy older adults…it is weight loss, e.g. loss of homeostasis, that denotes

with BMI <23

		risk. This measure should be place in observation only for at least 2 years, as preliminary chart review suggests that the volume sampled would not be statistically significant. We recommend a **follow up measure** for obesity only: Numerator = patient with documented follow up plan; Denominator = patients with BMI >=30. Until we receive data on weight loss, the size of the numerator involved is unknown, and we recommend data collection only for the underweight category.	
Metric	**Discussion**	**Recommendations**	**Methodologic Flaw**
Tobacco screening and Cessation (17)	The validity of comparing ACOs with different patient mixes is questionable – ACOs with more tobacco users have a much higher hurdle to pass than ACOs with fewer tobacco users because this measure requires some type of cessation intervention if the tobacco screening is positive for compliance but if the tobacco screening is negative then the patient is compliant without additional intervention. An ACO whose population rates of tobacco use rates tend to be lower could have a better score than ACO with excellent systems for follow-up but located in an area of	We recommend disaggregation of the numerators into 2 sub-measures. The initial sub-measure should **assess** smoking, and we suggest using the same	Flawed Aggregation logic

	the country with high tobacco use rates. This measure also has the same issue around determining appropriate actions to improve the measure rate because non-compliance can be due to no tobacco screening in time frame or a positive screening without cessation intervention. ACO's need to collect all the data anyway—clear disaggregation of the metrics allows transparency for all parties regarding what part of operations need targeting.	specs as the Meaningful Use measure to prevent reduplication of efforts and minimize message confusion amongst providers. A secondary sub-measure should assess **follow up**/counseling/treatment. Because of the newness of the metric, we recommend observation status only for this measure through at least 2013.	
BP screening (21)	All of the undersigned ACO's are committed to BP measurement and control. We typically have workflows that screen most patients in primary care. However, this particular metric drives resource use in ways that are not evidence based. The specification contains a non-validated interpretation of how to follow up BP measurements, including recommended lifestyle changes, which is complex and not directly supportive of JNC 7 guidelines. JNC7 clearly states "The classification of blood pressure is based on the average of two or more properly measured, seated BP readings on each of two or more office visits." It does not specify the length of time between visits.	We recommend reporting straight BP screening percentages only and keep this metric in observation status. We recommend the complex follow up portion of the metrics be dropped, pending JNC 8 which is due out this year.	No evidence- based benefit.

Metric

Discussion

Recommendations

Methodologic

Flaw

Hgba1c < 8

(22)

The utility of this hgba1c target is questionable, and may be detrimental, in the frail elderly and patients with limited life expectancy. ACO's with systems in place to support more frail elderly would actually be at a disadvantage.

Yau, et. al. **Glycosylated hemoglobin and functional decline in community-dwelling nursing home-eligible elderly adults with diabetes mellitus**. J Am Geriatr Soc. 2012 Jul;60(7):1215-21.

In this relatively large study (367 participants) with rigorous end points (death or functional decline in 2 years) out of a PACE program in SF-- HbA1c of 8.0% to 8.9% was associated with *better* functional outcomes at 2 years than HbA1c of 7.0% to 7.9%, suggesting that the current American Geriatrics Society guideline recommending a HbA1c target of 8.0% or less for older adults with limited life expectancy may be lower than helpful to maintain function.

Associated studies indicated that following guidelines for a Hgba1c of 7.0-7.9 in the frail elderly actually resulted in greater ER and Hospitalization usage due to hypoglycemia, with attendant higher costs.

While Pioneers have already accepted patient assignments, it is also important to acknowledge the unintended consequence of lending authority to this type of non-stratified metric would be to discourage systems from enrolling frail elderly in the future—the opposite of what we would like to achieve.

Recommend holding this metric in observation for the duration of the ACO contract while developing hgba1c refinements appropriate for a Medicare FFS population.

Analyze the effectiveness of using the hgba1c >9 in the “all or none metric” in 2013. Would that metric alone suffice?

The evidence basis for achieving a goal Hgba1c <8 in the frail elderly is questionable.

Metric	Discussion	Recommendations	Methodologic Flaw
Aspirin use in Diabetics with IVD (26)	This measure lies under Diabetic All or Nothing scoring. The denominator is different from the rest of the metrics that are aggregated in the Composite, making it mathematically illogical.	We recommend that this metric be reported separately from the diabetic composite.	Flawed Aggregation Logic
IVSD and lipids (29)	Most studies on the benefits of lipids used an age cut off of 75. Two significant studies showed some benefit up to age 80.	We recommend a rate based on an age cut- off at 75, and data for all patients greater than 75 held for future analysis.	Lack of evidence based benefit in the very elderly
HF and BB (31)	CMS proposes to randomly sample patients via submitted CPT and ICD-9 codes. Many patients with EF values in the range of 40-55 are diagnosed with systolic dysfunction, and others may be mislabeled based on subsequent improvement in EF values after an initial diagnosis. The reporting of values for these measures would not only be a labor intensive process at our end, but there is a high likelihood that many patients will be rejected from the sample.	We recommend improving CMS sampling methodology and holding this metric for analytical observation only. We also recommend broadening the list of Beta Blockers to capture all relevant data.	Denominator size too small

In fact, initial chart review estimated only 18% of the given sample would fall into the denominator definition.

Given the tiny denominators, the statistical power to establish an accurate estimate of any individual rate is highly doubtful. Tying reporting and performance benchmarks to a metric this underpowered is not appropriate.

However, we agree that observation of the aggregated metric across all ACO's may be useful to CMS for anchoring future quality benchmarks.