Dear Ms. Tavenner

Thank you for the opportunity to comment on the Quality Rating System (QRS) Scoring Specifications for qualified health plans (QHPs). At Connect for Health Colorado, the Colorado insurance marketplace for individuals, families, and small employers, our primary goals are to increase access, affordability, and choice. We believe a strong quality rating system will support both access and choice for our customers, as well as affordability through driving higher quality healthcare.

Colorado is one of the most competitive insurance marketplaces in the country and Connect for Health Colorado is a non-purchasing exchange. For the 2014 enrollment year we had 10 insurance companies offering 150 different health plans on the individual exchange. Given this, we have an interest in encouraging our consumers to make informed choices with as much information available as possible, both on new start-up plans and longer established organizations.

We would offer the following comments on the proposed QRS scoring methodology, in order of priority:

1. We would be concerned that the large number of indicators and measures would disadvantage plans with smaller enrollment. We would like to see modeling performed that aims to predict how start up plans with smaller enrollment might behave with respect to adequate sample size and whether they are able to meet the "half scale" rule for the composite and domain ratings and whether they will be able to satisfy the "full-scale" rule for the summary and global ratings.
    a. We would be concerned with the sample size selection and the risk that small samples may incorrectly represent the actual performance of the plan or provider, either due to adverse selection or random variation in performance.
    b. Since the global rating will likely be the only measure we can display in the primary "shopping page" any plan that is not able to report a global measure because of small sample sizes would likely be disadvantaged as compared to plans that do have a rating.
2. Many of these measures were developed by NCQA and are proprietary HEDIS measures. Many of our health insurance partners use URAC as an accreditor. We would hope CMS would make the auditor of any results accreditor-agnostic to avoid disadvantaging either acceptable accreditor.
3. The measure sets are comprehensive, and include measures that will be meaningful to most if not all health care consumers. The nomenclature however is very confusing and we would support finding less technical names for each level in the hierarchy.
4. The roll-ups into composites, domains, summary measures and global measures allows consumers to stay high level or get to the level of granularity that is meaningful to them, assuming they understand this very complex schema and that the results are displayed on the website in such a way to make toggling back and forth between the different levels intuitive and valuable.

a. This complexity concerns us. We would urge CMS to use plan language and develop guides with examples to help consumers understand how to use this information to inform their health plan decisions.
b. Our website does not currently have the functionality to display this type of information and allow drill downs, so it would have to be developed.
c. To allow consumers to define the summary measures, domains or composites that are relevant to them and their families, there may need to be a search tool developed that uses layman's language to direct consumers to applicable results.

5. Although the comment period on the hierarchies has closed, we would like to add that the 5-level hierarchy may be needlessly confusing and would be in favor of a 3- or at most 4-level hierarchy.
6. Although not mentioned in the draft scoring specifications, we would like to inquire as to the proposed timing of the release of results. It will likely take some time to do the analytics at CMS of the submitted data, which means that the information will not be available to consumers or the States for some length of time. If at all possible, we would urge CMS to attempt to release data that is one year old, rather than two or more years. Especially in these early years of the exchanges, the plans are changing rapidly and a two-year lag may disadvantage plans that are rapidly growing and changing.
7. Standardizing scores allows direct comparisons between QHPs and for assessing QHPs, but it does remove the view of how much improvement is needed overall, or not. For example, if we understand the methodology correctly, if top performing plans on the Diabetes retinal exam measure are performing around 60%, meaning almost half of their members did not get recommended retinal exams, they would be standardized to a 90 or higher score, yet their performance by objective measures is still mediocre. Likewise for higher performing measures, where all plans are clustered within a few points, like breast cancer screening, a plan at 70% performance may be standardized to 50 points while a plan at 75% may get 80 points, while in reality there is little clinical difference in performance. We are not sure how to fix this but is seems like there needs to be some overlay of meaningfulness or reasonableness to make the standardization fair.

Thank you for considering our concerns.

Regards,

Connect For Health Colorado