Inclusion of healthcare disparities issues in pay-for-performance programs should await development of reliable means of measuring changes in disparities over time

Casalino et al.[1] offer certain reasons to expect pay-for-performance programs to increase health disparities. Such reasons are not implausible. But the source Casalino et al. cite as evidence that such programs may have increased disparities (Werner et al.[2]) suffers from the failure to recognize that the increase in the absolute difference between black and white CABG rates observed following implementation of a CABG report card program was to be expected in light of the fact that overall CABG rates were increasing, and did not necessarily indicate an increase in disparity in any meaningful sense. I recently explained the reasons for this expectation in a comment on Werner et al.[3] A further recent illustration of these reasons may be found in a comment on Escarce and McGuire,[4] which comment shows that simply because they rely on relative rather than absolute differences, Escarce and McGuire would have found a decline in disparity in the same circumstances where Werner et al. found an increase in disparity. I have also explained in the references to the comment on Werner et al. that all standard measures of differences between rates are problematic for measuring changes in disparities over time, since, solely for statistical reasons, each such measure tends to change whenever the overall prevalence of an outcome increases or decreases.

The same measurement issues raise questions about the wisdom of the recommendation of Casalino et al. that pay-for-performance programs should provide rewards for reduction in disparities as well as for increases in overall quality. The Massachusetts Medicaid pay-for-performance program discussed by Casalino et al. as intending to include reduction of racial disparities as one of its goals illustrates the problem. As reflected in Recommendation on Method number 7.b at page 19 of reference 5 to this comment, it has been recommended that the Massachusetts program measure progress in reducing disparities based on changes in absolute differences between white and minority rates of experiencing an outcome. Thus, for reasons explained in references 5 to 15 of the comment on Werner et al., improvements in healthcare, measured by increases in rates of favorable outcomes, would tend to cause healthcare disparities to be perceived as changing in the following manner (which holds notwithstanding that, according to the recommendation, other factors would go into the development of the denominator of a fraction for which the change in absolute difference will be the numerator):

For relatively uncommon process outcomes, like those examined by Werner et al. and Escarce and McGuire, as well as by Jha et al.[6] (where, solely for statistical reasons, increases in overall rates tend to be accompanied by increases in absolute differences between rates), improvements in care will tend to be accompanied by perceptions of increasing disparities. But for more common process outcomes, like those examined by Trivedi et al.[7] (where, solely for statistical reasons, increases in overall rates tend to be accompanied by increasing absolute differences between rates), improvements in care will tend to be accompanied by perceptions of decreasing disparities.
As discussed in a second comment on Werner et al.[8], however, pay-for-performance programs may well be principally focused on clinical outcomes, like control of hypertension among persons deemed hypertensive. For reasons discussed in references 6-10 and 13-14 of the first comment on Werner et al., improvements in favorable clinical outcome rates would tend, at least initially, to be accompanied by perceptions of increasing disparities. Further, as discussed in the second comment on Werner, in the case of clinical outcomes, even such approaches to measurement of disparities as might effectively evaluate disparities in process outcomes without being affected by changes in overall prevalence (the reliability of which remains unclear) appear to be problematic with regard to evaluating disparities in clinical outcomes.

Reliance on absolute differences creates the curious situation where improvements in healthcare may tend at various times (or with regard to particular types of procedures) to lead to perceptions of increasing disparities and at other times (or with other types of procedures) lead to perceptions of decreasing disparities (something that would also apply to differences measured in odds ratios which tend, as least where distributions are normal, to exhibit patterns of change that are the opposite of those exhibited by absolute differences). Thus, depending on a number of factors including the type of services a health care entity provides, various entities that are achieving the same degree of improvements in care might be perceived as achieving opposite results with respect to changes in disparities. Just to be completely clear, however, that is not an arguments for reliance on measures like relative differences in rates of favorable outcome (as in Escarce and McGuire and as would tend usually to correlate improved care with decreasing disparities) or relative differences in adverse outcomes (as recommended by the National Center for Health Statistics and as would tend usually to correlate improved care with increasing disparities). Like reliance on absolute differences, each approach not only would find disparities to have changed when they have not, but would fail to identify genuine changes in disparities.

Thus, tying pay-for-performance to changes in healthcare disparities, or even attempting to evaluate the impact of pay-for-performance on healthcare disparities, ought to await the development of measures that can reliably distinguish between changes that are solely the consequence of changes in the overall prevalence of an outcome and those that are not.

References:


