Comments on the draft of *Commissioned Paper: Healthcare Disparities Measurement*, with Author Response

Below are the content of my email sent to the authors of Commissioned Paper: Healthcare Disparities Measurement on September 8, 2011, commenting on the draft of the Commissioned Paper that had been made available online. Bracketed numbers indicated the number assigned to each portion of material in the excel file maintained online as Comments on Disparities Measurement Commissioned Paper. Italicized material reflect the authors' responses to each item. The material is reproduced here because the online version does not contain the electronic links to referenced materials that were in the original email.[1]

A few days ago I found on the internet the draft of your guidance paper for the National Quality Forum (NQF) Steering Committee. It is too late for a member of the public to submit comments on the NQF site, though I may formally write the NQF after I have reviewed the draft a little more.

Meanwhile, I thought I would contact the authors directly to urge you not to finalize the draft paper without a substantial revision.

[112] The guidance paper is a potentially quite important document. Yet the current draft fails to address crucial measurement issues, and publication in this form will eventually detract from the prestige of institutions associated with it and, more important, contribute to considerable waste of resources by health disparities researchers and healthcare administrators.

Commissioned paper authors considered comment; no additional changes were made to the paper.

[113] My central concerns can be gleaned easily enough from my October 22, 2009 letter to Janet Corrigan, NQF President and CEO, which is available online along with similar correspondence on the Institutional Correspondence sub-page of the Measuring Health Disparities page (MHD) of jpscanlan.com. (Dr. Corrigan responded that the points raised were being considered by the NQF's Performance Measurement team.) The letter describes near universal problems with health disparities measurement arising, not simply from the fact different measures of disparities may yield different interpretations as to the size of disparities or the direction of change over time, but from the fact that each standard measure of differences between outcome rates (proportions) tends to be systematically affected by the overall prevalence of an outcome.

Changes were made to the commissioned paper to address the comment. The following revisions are now included in section 4c, "especially when the prevalence of the outcome changes."

[114] Most notably, the rarer an outcome, the greater tends to be the relative difference in experiencing it and the smaller tends to be the relative difference in avoiding it. Thus, as healthcare improves, relative differences in rates of appropriate healthcare tend to decrease while relative differences in rate of failing to receive appropriate healthcare tend to increase. Absolute differences and odds ratios also tend to be affected by the overall prevalence of an outcome,

---

[1] Bracketed material within an item reflect minor corrections to the original material.

though in a more complicated way. Roughly, as uncommon outcomes become more common absolute differences tend to increase; as common outcomes become even more common, absolute differences tend to decrease. Differences measured by odds ratios tend to change in the opposite direction of absolute differences.

Comment appreciated; no additional changes made to paper.

[115] Subject to a few qualifications, as appropriate healthcare becomes more common, the absolute difference tends to change in the same direction of the smaller of the two relative differences, creating a situation where one commonly can find a relative difference that will contradict perceptions based on absolute differences, and which, as I'll show, the draft does by reference both to relative differences in favorable outcomes and relative differences in adverse outcomes. I emphasize that my point is not simply that the different measures lead systematically to different conclusions about the comparative size of disparities – and it is certainly not that one standard measure is superior to another. Rather, the point is that none of the measures is a useful indicator of the size of a disparity, at least not without consideration of how it is affected by overall prevalence in a particular setting.

(While the draft, which I have not read fully, does not much discuss odds ratios, inasmuch as your note 6 discusses the 1999 Schulman article and the odds ratios, I should mention that any full discussion of the patterns of changes should be mindful of the way differences measured by odds ratios changes as prevalence changes. As noted, differences measured by odds ratios tend to change in the opposite direction of absolute difference, which would mean they tend to change in the same direction as the larger of the two relative difference[s].)

Commissioned paper authors considered the comment and agree prevalence is an important indicator of disparities. However, they believe the detail specified is not needed within the paper.

[116] About 140 references explaining these patterns in particular settings may be found on the Measuring Health Disparities page (MHD) of jpscanlan.com and the nuances of the patterns are explained on the Scanlan's Rule page (SR) of the same site. The Mortality and Survival page discusses the way that, particularly in cancer journals, researchers discuss disparities in mortality and disparities in survival interchangeably without recognizing, for example, that general increases in cancer survival tend to reduce relative differences in survival but increase relative differences in mortality. The Relative Versus Absolute sub-page of MHD, which is a sort of response to you reference 62 (Harper 2010) (and which could as well be a response to several point in the draft) explains why choice of measure does not involve a value judgment; rather, there can be only one underlying reality as to whether one disparity is larger than another (though, to be sure, it often will be difficult to divine that reality). The Solutions page of MHD describes a method (probit) for measuring health inequalities that is not affected by the overall prevalence of an outcome. Section E.7 of MHD describes the extent of scholarly agreement with my thinking on these issues. A few key references are found after the signature.

Comment appreciated; no further response required.

[116] I suggest that the essential points are hardly debatable. For example, no one confronted with the data would fail to recognize that lowering poverty will tend to increase relative differences in poverty rates while reducing relative differences in rates of avoiding poverty (see Table 1 of Chance 2006); that lowering test scores will tend to increase relative differences in failure rates while reducing relative differences in pass rates (see Table 1 of BSPS 2006); that lowering blood pressure will tend to increase relative differences in hypertension rates while reducing relative differences in rates of avoiding hypertension (see Figure 10 of ICHPS 2008); that improving folate levels will tend to increase relative differences in low folate while reducing relative differences in rates of adequate folate (Table 1 of NHANES Illustrations). See also Comment on Morita, which discusses an award-winning study where the authors (relying on relative differences in the favorable outcome) found that dramatic increases in immunization rates led to dramatic decreases in immunization disparities, while NCHS (which would rely on relative differences in adverse outcome) would have found dramatic increases in disparities.

Comment appreciated; no further response required.

[117] Once one recognizes these patterns, there may well be points to be made about whether underlying distributions that cannot be directly observed will in fact be normal and other issues things that I have covered in various places. But the complications suggested by such points only go to whether in fact it will be possible to appraise disparities while taking the distributionally-driven patterns into account. They do not provide a basis for reliance on standard measures without consideration of the way such measures may be affected by the prevalence of an outcome.

Comment appreciated; no further response required.

[118] Yet while the draft discusses that one may appraise the size of a disparity differently depending on whether one examines the relative difference in the favorable outcome or the relative differences in the adverse outcome, at no point does it suggest that one may reach different conclusions as to directions of changes over time (the main issue of concern in health disparities research) depending on which relative difference one chooses – much less that this will be the case commonly if not most of the time. At the point when the report (at 37) seems about to address the possibility of different interpretations as to change over time depending on whether one examines relative differences in favorable outcomes or relative difference in adverse outcomes, the report instead addresses the fact that in 4 of the 7 cases where Trivedi et al found decreasing absolute differences one would find increasing relative differences in the adverse outcome. That is a point about relative differences versus absolute differences, not a point about the differences between relative differences in favorable or adverse outcomes, which goes completely unaddressed. In all 7 cases, incidentally, the relative difference in the favorable outcome decreased (suggesting that in 3 of 7 cases (i.e., where both relative measures decreased) there occurred a meaningful reduction in disparities while in 4 cases one cannot tell whether there occurred a change that was other than the distributionally-driven consequence of the increasing prevalence of the outcome). Possibly, the italicizing of "adverse events" at the top of 38 is a vestige of an earlier version that more directly explained the possibility for the two relative differences to yield opposite results. But I think few readers would regard the draft as explaining that the two relative differences can yield different interpretations as to changes over time.

Comment reviewed by authors of the commissioned paper; no additional changes made to paper.

[118] NCHS has treated this issue in your reference 60 and elsewhere. As I discuss in many places (e.g., . Section E.7 or MHD, Comment on Keppel, Sec. A.6 of SR, refs 1 and 5 below), NCHS has done so in a misguided manner. But it has at least acknowledged that one may interpret changes over time differently depending on which relative difference is examined. I do not think that such possibility would occur to most readers of the draft.

Comment appreciated; no further changes made to paper.

[120] Turning to the some further examples in the draft, Figure 6 shows difference in interpretations of changes over time depending on whether one examines absolute differences or relative differences in the adverse outcome. But the data presented show that the relative difference in the favorable outcome – which, as reflected by Morita, is still the predominant method for analysis of healthcare disparities – decreased. More pertinent to my key concern, each of the changes is in accordance with the way such measures typically changes when a common outcome (receipt of test) becomes even more common.

Comment reviewed by authors of the commissioned paper; no additional changes made to paper.

[121] The draft's Figure 7 does not appear properly in the online version, so I skip that (though I trust I could derive the numbers from Harper's figure 1). But after the discussion of the difference between changes measured by absolute difference and relative differences in the adverse outcome reflected in the figure, the report then challenges the Werner finding of increased CABG disparities based on absolute differences. It does so by reference to larger black than white percentage increases in CABG rates. But in contrast to the discussion of Figure 7, the draft is relying on relative changes in the favorable outcome rather than the adverse outcome. As discussed in the 2/8/08 Comment on Werner, like the absolute [difference,] the relative difference in the adverse outcome increased. Again, the main point is that each of the observed measure changes [in the manner that] typically occurs when a very uncommon outcome generally increases (as also reflected, for example, in the Comment on McGuire/Escarce). The comment on Werner also discusses what I maintain is the only useful measure of the disparity based on the Werner figures (though see the 2/10/08 Comment on Werner (same link) regarding the potential shortcomings of that approach in the particular setting).

As already mentioned,

[122] the report does discuss that relative and absolute differences may yield different interpretations as to changes over time. But, in addition to failing to recognize that one relative difference will commonly change in the same direction as the absolute difference while the other will commonly change in the opposite direction of the absolute difference, the report fails to show any recognition of the role of overall prevalence with regard to either the patterns of changes in absolute difference or patterns of changes in relative difference in the adverse outcome. Further, the report then merely recommends that both absolute differences and relative differences should be calculated (though not indicating which relative difference) and should be presented when they lead to different conclusions about changes over time, but leaves it up to the reader to determine which figure is more meaningful. Given that I do not regard any of the measures as useful (unless interpreted with regard to overall prevalence), I hesitate to say much

about the guidance offered. But the draft offers no guidance on how the reader should make the determination. As suggested, I do not think there exists useful guidance for choosing whether a change in [the] relative or [the] absolute difference is more meaningful. But for the draft to fail to offer any guidance to the reader still seems inappropriate.

Comment reviewed by authors of the commissioned paper; no additional changes made to paper.

[123] The failure to give guidance is particularly problematic in the context of the later suggestion that performance in pay-for-performance programs should take into account effects on disparities. As I discuss on the Pay for Performance sub-page of MHD, pay-for-performance is the area where efforts to measure health disparities without understanding the way that changes in outcome rates affect various measures are likely to have the most serious detrimental consequences. That is, most discussions of changes in health and healthcare disparities are largely academic, with misperceptions' rarely having important policy consequences. But tying P4P to disparities reductions means entities will actually be paid according to perceptions about changes in disparities that do not have a sound statistical basis.

Comment was reviewed by the authors of the commissioned paper, no further changes made to paper. NQF will consider the suggestions within the context of the Healthcare Disparities consensus standards project

Thus, [124] I suggest that you review some of the comments collected in Section D of MHD – including those relating to various of references in the draft, e.g., Comment on Trivedi NEJM 2005 Ref. 64 (and Comment on Trivedi *JAMA 2006*, which further discusses Trivedi NEJM 2005), Comment on Casalino Ref. 83, Comment on Baicker Ref. 93, Comment on Chien Ref. 101 – as well as the discussion in Section E.7 of varied researchers here and in Europe who have been recently recognized the measurement issues that so far are unmentioned in the draft. Section D should suggest to you that almost all health disparities research that has gone beyond determining that a disparity exists (that is, that in some manner endeavors to appraise the size of such disparity) has been fundamentally flawed for failure to recognize the problematic nature of standard measurement tools. And the latter should persuade you that it is rather late in the day to be creating substantial works on the measuring of health or healthcare disparities without at all considering the extent to which various measures change solely because of changes in the overall prevalence of an outcome. I note that Harper and colleagues have yet to discuss patterns by which measure are affected by overall prevalence (save in a brief AJE reference to the Houweling article discussed in Section E.7). But Sam Harper was on the same panel in which I presented reference 6 and I don't think he will continue provide guidance on measuring disparities while ignoring these issues. I also suggest that you carefully review the Solutions page and consider ways in which the approach described there can be improved upon. The development of refinements on that approach or the development of other approaches that are able to identify changes that are not simply function of changes in overall prevalence would be a signal contribution to health and healthcare disparities research.

Comment appreciated; no further response required.

[125] References:

1. Can we actually measure health disparities? *Chance* 2006:19(2):47-51:
http://www.jpscanlan.com/images/Can_We_Actually_Measure_Health_Disparities.pdf
2. Race and mortality. *Society* 2000;37(2):19-35 (reprinted in *Current* 2000 (Feb)):
http://www.jpscanlan.com/images/Race_and_Mortality.pdf
3. Divining difference. *Chance* 1994;7(4):38-9,48:
http://jpscanlan.com/images/Divining_Difference.pdf
4. The perils of provocative statistics. *The Public Interest* 1991;102:3 14:
http://jpscanlan.com/images/The_Perils_of_Provocative_Stat.pdf
5. Measurement Problems in the National Healthcare Disparities Report, presented at American Public Health Association 135th Annual Meeting & Exposition, Washington, DC, Nov. 3-7, 2007: PowerPoint Presentation:
http://www.jpscanlan.com/images/APHA_2007_Presentation.ppt; Oral Presentation:
http://www.jpscanlan.com/images/ORAL_ANNOTATED.pdf; Addendum (March 11, 2008):
http://www.jpscanlan.com/images/Addendum.pdf
6. Measuring Healthcare Disparities," presented at 3rd North American Congress of Epidemiology, Montreal, Canada, June 21-24, 2011:
http://jpscanlan.com/images/NACE_Presentation.ppt