E-mail Regarding 2007 NHDR Sent to AHRQ Staff March 12, 2009

This is a follow-up to my email of November 12, 2007, regarding certain technical issues in the National Healthcare Disparities Report (NHDR).

Since the time of that note, I have continued to take issue with the health disparities measurement approach employed by the Agency for Healthcare Research and Quality (AHRQ) in the NHDR (as well as with the somewhat different approach of the National Center for Health Statistics (NCHS)), as summarized in the Section E.4 of the Measuring Health Disparities page (MHD) of jpscanlan.com and Section A.6 of the Scanlan's Rule page of the same site. The main area of disagreement involves the agencies' use of relative differences between outcome rates of advantaged and disadvantaged groups to measure health and healthcare disparities – in the case of AHRQ, use of the larger of (a) the relative difference in adverse outcomes or (b) the relative difference in favorable outcomes; in the case of NCHS, use of the relative difference in adverse outcomes – without recognizing the way that the two relative differences tend to change systematically in opposite directions as the overall prevalence of an outcome changes. Items of particular note include the 2007 APHA presentation that I previously brought to your attention (and to which I added an important Addendum in March 2008, which addresses the implications of my coming to recognize that AHRQ relied on the larger of the two relative differences), and a comment on a 2008 study in *Pediatrics* concerning changing disparities in vaccination rates (Comment on Morita). The comment provides a particularly useful illustration of the central issues. It involves a study where the authors (relying on relative differences in favorable outcomes) found dramatic decreases in disparities in circumstances where NCHS (relying on relative differences in adverse outcomes) would find dramatic increases in disparities, and where AHRQ (relying on the larger relative difference) would concur with NCHS for one period and concur with the authors for another period.

As with my November 2007 note, however, this note is intended mainly to address certain matters apart from the central disagreement over measurement issues. Like the earlier note's treatment of matters where the 2006 report appeared to be inaccurate or misleading, this note addresses matters where the 2007 report may be inaccurate or misleading. As before, I hope these points can be addressed in an errata sheet regarding the published report and can be addressed in the next report. But the earlier points were not addressed in an errata sheet for the 2006 report or in the content of the 2007 report (save that the misdescription of odds ratios discussed in Section A.2 of the November 2007 note has been corrected in the 2007 report). Therefore, I am posting in slightly edited form the November 12, 2007 note, as well as an October 29, 2007 note that raised related questions, on a sub-page of MHD. And I am posting this note as well. For there should always be a publicly available record of instances where a government document may be inaccurate or misleading. This holds especially for scientific document on which researchers continually rely and where errors may find there way into subsequent publications. Ideally, such record will be created and maintained by the governmental entity creating the record. So I hope AHRQ will eventually adopt my recommendation in that regard.

My comments on the 2007 report are grouped under four headings below. Section A addresses the failure of the report to clarify that it is measuring disparities in terms of the larger relative difference, even when it presents rates in terms of the opposite outcome of the one underlying the disparities measurement. This was addressed in Section A.4 of my earlier email, though then I was still uncertain as to precisely how AHRQ was measuring disparities. Section A of this note illustrates the issue with reference to the 2007 report. Section B gives further attention to the matter addressed in Section A.1 of the earlier note, specifically, the treatment of percentage point changes as percent changes. As illustrated in the earlier note, in the case of the only two instances where the 2006 report discussed the size of a change in particular disparities, the confusion of these terms caused an increase in one disparity to be understated by 92% and a decrease in the other disparity to be overstated by 778%. Section C addresses the way certain highlighted points in the 2007 report may be misleading or incorrect. Section D addresses a usage problem that was common in reports prior to the 2006 report, which seemed to have been corrected in the 2006 report, but which is pervasive in the 2007 report.

With regard to the first three of the items, I again stress the utility of providing summaries of the data underlying the reports conclusions about changes in measures, with such summaries' including the actual rates being compared for each core measure (as well as the source from which the rate is drawn). The inclusion of such summaries will make the document substantially more informative and easier to use, will clarify for the readers the way the disparities and changes in those disparities over time are being measured, will likely lead to the avoidance of many of the errors of the kind noted in Section C, and will enable readers to more easily identify situations where such errors may have occurred.

A. Failure to Clarify that Disparities are Presented in Terms of Larger of the Two Relative Differences (i.e., in the Favorable or the Adverse Outcome)

As reflected in the March 2008 Addendum to the 2007 APHA presentation, it took me some time to recognize that AHRQ was relying on the larger relative difference in the measurement of disparities. But, while still uncertain on that point, in the November 2007 note, I urged AHRQ to provide clarification on this issue. And I noted that many of the bar graph illustrations that reflected favorable outcomes were misleading and would better illustrate the relative differences on which the report is actually relying if turned upside down.

The 2007 report is similarly misleading in this regard. Only in note vii at page 18 and in the text at page 147 is there a suggestion that the disparity might be measured in terms of the larger relative difference. In all cases where the core measure is phrased in terms of a favorable outcome, the discussion of the disparity is couched in terms of that favorable outcome even when the disparity is measured in terms of the opposite outcome. The first two highlighted changes in disparities (also discussed in Section C below) involve such situations. And, as discussed in Section C, few people who read that the relative rate of

adequate hemodialysis changed from 1.29 to 1.08 will realize that the relative differences in adequate hemodialysis rates were not 29% and 8% but were in fact approximately 5% and 1%.

The above example involves a situation where only the size of the disparity varies by whether the relative rate is based on favorable outcome rates or adverse outcome rates. But in other cases, as illustrated in the comment on the Morita study, the direction of change may depend on whether the relative rates for the two periods compared are based on favorable or adverse outcome rates. And even when AHRQ gave the greatest attention to issues about measuring disparities in terms of relative differences in favorable or adverse outcomes (which I believe occurred at page 34 of the 2003 report), it did not discuss that the determination of whether a disparity is increasing or decreasing may turn on whether one examines the relative difference in the favorable or the adverse outcome (as NCHS had done, for example, in Keppel K., Pamuk E., Lynch J., et al. 2005. Methodological issues in measuring health disparities. Vital Health Stat 2 (141) (Keppel 2005)). In any event, the fact there may occur situations where the determination as to the change in direction of a disparity turns on whether one examines the favorable or an adverse outcome seems a compelling reasons for the report to be as clear as possible on this issue

Finally, in Section A.4 the November 2007 note, I pointed out that researchers, sometimes funded by AHRQ, were measuring healthcare disparities in ways that would tend to lead to opposite results from those AHRQ would reach, and were doing so while unaware that AHRQ would measure disparities any differently. The Morita study, while not funded by AHRQ, is another example of the ways researchers are measuring healthcare disparities without recognizing that NCHS and AHRQ would employ approaches that may often lead to opposite conclusions. AHRQ cannot fulfill its role its role as the leader in healthcare disparities research without making its methodology abundantly clear to researchers.

B. Difference between Percent Changes in Disparities and Percentage Point Changes

In my November 2007 note, I explained that the confusion between a percent change in a disparity and a percentage point change caused – in the only two instances where the size of a change was specifically discussed – a 122% yearly increase in a disparity to be described as a 9.8% yearly increase and a 0.9% yearly decrease in a disparity to be described as a 7.9% yearly decrease. That confusion apparently also affects all determinations of whether a change in a disparity meets the 1% per year threshold to be regarded as an increase or decrease and whether it meets the 5% change for a larger increase.

The 2007 report contains no instances where the size of a particular change in disparity is described in percent change (or percentage point change) terms. But the description of the method of calculating changes over time in note x at page 3 and the text at page 147 seems to indicate that the 2007 report is relying on percentage point changes for such

calculation. Thus, though the 2007 report does not discuss the size of a particular change, all the analyses of counts of changes described as being greater than 1% and greater than 5% are affected.

The NCHS measures changes in disparities in terms of percentage points for purposes of measuring progress in Healthy People 2010 (though it is very careful to make that clear in various places including the Healthy People 2010 Midcourse Review). Thus, the measurement of changes in disparities in the NHDR is consistent with the approach of NCHS. The problem is that the references to percentage point changes as if they are percent changes may well mislead the public.

In the case of the 2006 report, because for two changes the report provided the specific figures for both groups being compared, a careful reader could divine that the so-called percent changes were in fact percentage point changes. But my reading of the 2007 report reveals no situations where the body of the report provides information that would allow the reader to reach that conclusion.

Finally, I note that the significance of the difference between a percentage point change and a percent change can vary greatly. When the relative rate is, say, between 1.8 and 2.2, a 10 percentage point change is not much different from a 10 percent change. But when the relative rate is 1.2, a 10 percentage point increase or decrease would reflect a 50% change in the disparity. And, when the relative rate is between 10 and 11 (as in the case of the racial difference in new AIDS cases), a 10 percentage point changes is approximately a 1% change. Persons told that the racial differences in new AIDS cases had been declining at the rate of 10% per year, as was highlighted in the 2006 report, would be led to believe that the disparity might be eliminated in ten years. At the actual rate of decline, it would take 100 years to eliminate the disparity.

Finally, I note that AMA Style Manual also specifically addresses the distinction between percentage point differences and percent differences Section 19.7.2 (at 831), noting that that "[t]he two terms are *not* interchangeable." (original emphasis). NCHS is obviously quite cognizant of the distinction as reflected by its invariably careful usage. AHRQ ought to be similarly careful in its usage.

C. Points Highlighted in the Report that are Misleading or that Contain Factual Errors

Because the 2006 and earlier reports were the subject of my 2007 APHA presentation, my review of those reports, especially the 2006 report, was extensive. By contrast, I have reviewed the 2007 report mainly to see whether issues I raised about the earlier reports had been addressed. Nevertheless, I found a number of ways in which the discussion of the most highlighted reductions in disparities was misleading or incorrect. Those issues are addressed below. To the extent that the points I make regarding incorrect numbers are deemed to be well taken, some review of the accuracy of the remainder of the report may be warranted.

Table H.1 at 5 of the 2007 NHDR presents a list of core measures that showed reductions in disparities for selected groups from 2000-2001 to 2004-2005, describing the earlier figures as based on the earliest year in the NHDR. The same table appears as slides 4 and 5 of the PowerPoint summary of the report. In the report itself, an asterisk is placed on four disparities that were eliminated or reduced to nonsignificance from the earliest year to the most recent year. Page 6 then sets out four bullet points discussing the changes in these four disparities, stating: "Overall, four of the core measures showed disparities getting smaller or [being] eliminated." The four bullet points, with numbers added, are set out in italics below. A comment follows immediately after each point. As shown in the comments, items 1 and 2 are misleading with respect to the outcome being measured. Item 2, which purports to show a substantial reduction in disparity from the earliest year in the reports to the most recent year, actually presents a change from other than the earliest year, when there in fact was no change from the earliest year to the most recent year. Item 3 presents the wrong earliest year figure. Item 4, which purports to show a dramatic reduction in disparity, from a 3.7 relative rate to a 1.10 relative rate, and which is the largest reduction shown in the report, actually involves a reduction from a relative rate of 1.14 to a relative rate of 1.10.

1. From 2001 to 2005, the proportion of Black hemodialysis patients with adequate dialysis improved (from 82% to 87%; in 2005 this rate was not statistically different from Whites), and the gap between Blacks and Whites decreased (from a relative rate of 1.29 to 1.08).

Comment on item 1:

The 2005 87% figure for blacks in item 1 is from Table 34a of the 2007 report, which also shows a rate of 88% for whites. The corresponding 12% and 13% figures for failing to receive adequate hemodialysis would yield the 1.08 relative rate set out in the report. The 2001 82% figure for blacks is apparently from page 45 of 2003 report, which also shows an 86% figure for whites. The corresponding 18% and 14% figures for failing to receive adequate hemodialysis yield the 1.29 relative rate for 2001.

There is no problem with the figures here. But, although it is possible to divine from 2001 black rate that the 1.29 relative risk could not be based on rates of receiving adequate dialysis, it is likely that few readers would in fact recognize that the relative rates shown are based, not on rates of receiving adequate hemodialysis, but on rates of failing to receive adequate dialysis. And many readers would be surprised to know that in the two years identified the differences in adequate hemodialysis rates were not 29% and 8%, but 5% and 1%.

Thus, in addition to clarifying its approach to measurement early in the document, I suggest that AHRQ ought to much more clearly indicate the instances where, though it describes a favorable outcome, it calculates the disparity in terms of the adverse outcome. That is especially so where the rates set, as in the case of the 82% and 87% figures in parentheses, are rates for the opposite outcome from that on which the disparity is calculated. Ideally, if the report is relying on the relative difference in the opposite

outcome from that in which the core measure is cast, the report should simply state the matter in terms of the adverse outcome.

2. From 2002 to 2004, the proportion of Asians with a usual primary care provider improved (from 69.3% to 75.2%), and the gap between Asians and Whites in having a usual primary care provider decreased (from a relative rate of 1.40 to 1.13).

Comment on item 2:

It first warrants note that, as indicated above, Table H1 had listed the same 1.40 figure as being from 2000-2001, and as being from the earliest year in the NHDR, while this bullet point indicates that the earlier figures are from 2002. As discussed below, this difference has significant implications.

The 2004 figures in item 2 are from Table 220a of the 2007 report, which shows rates of having a primary care provider of 75.2% for Asians and 78.1% for whites. The corresponding 24.8% and 21.9% figures for failing to have a primary care provider yield a relative rate of 1.1, as in the report. The 2002 figures are from Table 187 of the 2005 report, which shows rates for having a primary care provider of 69.2% for Asians and 78.1% for whites. The corresponding 30.8% and 21.9% figures for failing to have a primary care provider yield a relative rate of 1.406, essentially the same as that set out in the report.

But the 2002 figure are not the earliest figures in the NHDR. Table 132a of the 2004 report shows, for 2001, rates of having a primary care provider of 74.9% for Asians and 77.9% for whites. The corresponding 25.1% and 22.1% figures for not having a primary care provider yield a relative rate of 1.136. That is essentially the same relative rate as in 2004 (actually 1.136 in 2001 compared with 1.132 in 2004).

Thus, Table H.1 is incorrect both in describing the relative rate for the earlier period as being from 2000-2001 and in describing it as being from the earliest year in NHDR. And, while the bullet point is correct as to the year for the earlier figure, it stills gives the incorrect impression that the change is from the earliest time to the most recent. As noted, the recent relative rate is the same as the earliest relative rate. Hence, while the bullet point suggests progress regarding this disparity from the earliest point examined until the most recent, in fact the relative rate first got larger and then returned to its original size.

And, of course, this bullet point suffers in the same way as the first bullet point in that, while describing favorable outcomes, it presents relative rates based on the adverse outcomes.

3. From 2001 to 2004, the proportion of Hispanics who had a hospital admission for perforated appendix decreased (from 322.4 per 1,000 admissions to 291.8 per 1,000 admissions), and the gap between Hispanics and non-Hispanic Whites decreased (from a relative rate of 1.06 to 1.01).

Comment on items 3:

The 2004 figures in item 3 are from Table 242 of the 2007 report, which shows rates of admission with a perforated appendix per 1000 appendicitis admission of 291.8 for Hispanics and 287.8 for whites, which yields a relative rate of 1.01, just as stated in the report. The 2001 figures presumably are from Table 191a of the 2004 report, which shows rate of 329.274 (not 322.4 as in the report) for Hispanics and a rate of 300.416 for non-Hispanic whites. These figures yield a relative rate of 1.10 rather than the 1.06 figure in the report.

This appears to be a simple data transcription error of very minor consequence. But it warrants note that it is among several data errors in these four bullet points (including that discussed below with regard to item 4). Whether or not the existence of such errors among these highlighted four bullet points raises issues of the general accuracy of the report, it provides further reason for setting out all the figures in a summary sheet.

4. From 2001 to 2004, the proportion of people living in poor communities who had a hospital admission for perforated appendix decreased (from 344.2 per 1,000 admissions to 307.7 per 1,000 admissions), and the gap between poor and high income decreased (from 3.47 to 1.10).

The 3.47 earliest year relative rate is also shown in Table H.1. But the range of figures shown in items 3 and 4 by themselves suggests that something is likely amiss with this figure. If the figures shown in item 4 were correct, the rate of the advantaged group (in this instance, the highest income group) would have been 99.2 per 1000 in 2001 and 279.7 in 2004, certainly an improbable increase.

In any case, reference to the underlying tables reveals the following: The 2004 figures are from Table 242 of the 2007 report, which shows a figure of 310.7 for poor people and a figure of 283.4 for the highest income group. The 307.7 figure appears to have been erroneously drawn from the figure for all women, which is in the row above the figure for the poor. But the relative rate of 1.10 does appear to be based on the correct figure for the poor for the latter year.

Figures for 2001 should be drawn from Table 191a of the 2004 report. That table shows a rate of 348.781 (not 344.2) for the poor and a rate of 304.720 for the highest income group, which figures yield a relative rate of 1.14. I do not know the basis for the 344.2 figure or how the 3.47 relative risk might have been derived. But it is does seem clear that what appears to be a substantial change (one which could be characterized either as a 237 percentage point reduction or a 96% reduction) was a more modest reduction (one which could be characterized either as a 4 percentage point reduction or a 29% reduction).

Finally, I note again that setting out the rates for both groups being compared in a summary sheet would make errors like this likely. But I also note that Figure 3.11 at 130,

which sets out the actual rates, also makes clear that that the 3.7 relative rate shown in Table H1 and discussed on page 6 is incorrect.

Like item 3, this may involve a simple data transcription error and reports that contain many numbers often will have such errors. But if such errors cannot be entirely avoided, they should be avoided in the case of points that are highlighted. And given that in this instance the highlighted figure involves what a reader would interpret as, by far, the largest percentage (or percentage point) reduction of any disparity in the report, it should be corrected in an errata sheet.

D. Use of "Times Higher" Rather than "Times as High" to Quantify Disparities

When one group's rate of experiencing an outcome is 9% while the other group's rate is 3%, it will properly be said that the former figure is 3 times as high as the latter figure or that the former group is 3 times as likely to experience the outcome as the latter group. I believe that you will find that NCHS is very careful in this usage. Increasingly, however, less careful observers tend to describe the situation by stating that the first rate is 3 *times higher* than the second or that the first group is 3 *times more likely* to experience the outcome (which technically is the same as being 4 times as likely).

Such usage is incorrect and tends to give reader an exaggerated impression of the size of the disparity. This issue is addressed in the *New York Times Manual of Style and Usage* (1999 edition at 335). It is occasionally addressed in online discussion (e.g., link). Since 1991, the issue has been treated in *Precision Journalism* (which treatment may be accessed online at Precision Journalism). Thus, I doubt that anyone would seriously contend that 9% is three times higher than 3%, and certainly the NCHS would state that such usage is incorrect.

This error was common in the health disparities reports up until 2006. But in the 2006 NHDR the usage was correct in 12 of 14 cases. Because I interpreted that fact to indicate that AHRQ had considered this matter and deliberately adopted the correct usage save for a couple of oversights, I did not mention the matter in my comments on the 2006 report. In the 2007 report, however, while the correct "times as likely" and the incorrect "times more likely" are both used some number of times, in all 30 instances where the report quantitatively described one rate as being higher than another it uses the phrase "times higher."

There is no reason for a scientific document issued by the United States Government to be other than precise on a matter like this.