

James P. Scanlan
Attorney at Law
1529 Wisconsin Avenue, NW
Washington, D.C. 20007
(202) 338-9224
jps@jpscanlan.com

July 1, 2015

Richard Kronick, Ph.D.
Director
Agency for Healthcare Research and Quality
540 Gaither Road
Rockville, MD 20850

Re: (a) Failure of AHRQ to Recognize Patterns by Which Measures of Differences Between Health and Healthcare Outcome Rates of Advantaged and Disadvantaged Groups Tend to be Affected by the Frequency of the Outcome and Implications of That Failure Respecting AHRQ Activities Involving Health and Healthcare Disparities Research;

(b) Anomaly in Identifying Increases or Decreases in Healthcare Disparities Beginning with the 2010 NHDR Due to Altering Approach From (i) Quantifying Changes in Terms of Absolute Changes in Relative Differences to (ii) Quantifying Changes in Terms of Differences in Absolute Changes in Outcome Rates.

Dear Dr. Kronick:

On occasion I write to institutions or organizations whose activities involve the interpretation of data on demographic differences in the law or the social or medical sciences alerting them to ways in which their interpretations are undermined by the failure to understand patterns by which standard measures of differences between favorable or adverse outcome rates of advantaged and disadvantaged groups – or differences between the proportion a group comprises of persons potentially experiencing an outcome and the proportion it comprises of persons actually experiencing the outcome – tend to be systematically affected by the overall frequency of an outcome. Recipients of other letters involving interpretive issues of the type discussed in this letter include [Robert Wood Johnson Foundation](#) (Apr. 8, 2009),¹ [National Quality Forum](#) (Oct. 22, 2009), [Institute of Medicine](#) (June 1, 2010), [The Commonwealth Fund](#) (June 1, 2010), [United States Department of Education](#) (Apr. 18, 2012), [United States Department of Justice](#) (Apr. 23, 2012), [Board of Governors of the Federal Reserve System](#) (Mar.

¹ To facilitate consideration of issues raised in letters such as this I include links to referenced materials in electronic copies of the letters. All such letters may be found by means of the Institutional Correspondence subpage of the Measuring Health Disparities page of jpscanlan.com.

4, 2013), [Harvard University](#) (Oct. 9, 2012), [Harvard Medical School, Massachusetts General Hospital, et al.](#) (Oct. 26, 2012), [Mailman School of Public Health of Columbia University](#) (May 24, 2013), [Investigations and Oversight Subcommittee of House Finance Committee](#) (Dec. 4, 2013), [Education Trust](#) (April 30, 2014), [Annie E. Casey Foundation](#) (May 13, 2014), [Institute of Medicine II](#) (May 28, 2014), [IDEA Data Center](#) (Aug. 11, 2014), [Education Law Center](#) (Aug. 14, 2014), [Financial Markets and Community Investment Program, Government Accountability Office](#) (Sept. 9, 2014), [Wisconsin Council on Families and Children’s Race to Equity Project](#) (Dec. 23, 2014), [Portland, Oregon Board of Education](#) (Feb. 25, 2015), [Vermont Senate Committee on Education](#) (Feb. 26, 2015), [United States Department of Justice and City of Ferguson, Missouri](#) (Mar. 9, 2015), [Senate Committee on Health, Education, Labor and Pensions II](#) (Mar. 20, 2015), [Texas Appleseed](#) (Apr. 7, 2015), and [City of Minneapolis, Minnesota](#) (June 8, 2015). An *amicus curiae* [brief](#) I filed on November 17, 2014, in *Texas Department of Housing and Community Development, et al. v. The Inclusive Communities Project, Inc.*, Sup. Ct. No. 13-1371 (TDHCD brief), might be deemed a similar communication to the United States Supreme Court.²

Introduction

This letter addresses two related subjects. The first concerns the way activities of the Agency for Healthcare Research and Quality (AHRQ) involving health and healthcare disparities research are generally undermined by failure to recognize patterns whereby, for reasons inherent in the shapes of underlying risk distributions, standard measures of differences between outcomes rates tend to be systematically affected by the frequency (prevalence) of an outcome. This subject is pertinent to essentially all efforts by researchers and institutions in the United States and abroad to quantify the differences in the circumstances of advantaged and disadvantaged groups reflected by the differing rates at which the groups experience an outcome (or its opposite). The second subject concerns anomalies in the identification of directions of changes in the size of health and healthcare disparities (and quantification of those changes) in the National Healthcare Disparities Report (NHDR) arising from modifications of methodology commencing with the 2010 report. Among other things, such modifications created situations

² In recent letters such as this, I usually reference only formal letters since 2009. I note, however, that for some decades I have communicated by letter or email with many individuals and institutions regarding issues of the type addressed in this letter. Some of these communications (and responses thereto) are of particular pertinence to the federal government’s efforts to measure health and healthcare disparities. Among these are (a) a November 28, 1988 [letter](#) to Surgeon General David Satcher and Department of Health and Human Services Assistant Secretary for Planning and Evaluation Margaret A. Hamburg; (b) a January 20, 1999 [letter](#) from National Center for Health Statistics (NCHS) Director Edward J. Sondick; and (c) a January 25, 1999 [letter](#) to Dr. Sondick (which exchange is alluded at in my “[Race and Mortality](#),” *Society* (Jan./Feb. 2000) (at 6) and “[Race and Mortality Revisited](#),” *Society* (July/Aug. 2014) (at 332)). Email exchanges of particular pertinence to the subject of this letter include a number of exchanges with NCHS statistician Dr. Kenneth B. Keppel beginning in 2001 (which led to the recognition by NCHS, discussed *infra*, that determinations of whether health and healthcare disparities were increasing or decreasing would commonly turn on whether one examines relative differences in favorable outcomes or relative differences in the corresponding adverse outcomes. Other email exchanges of particular pertinence include those with staff and leadership of the Agency for Healthcare Research and Quality or the Centers for Disease Control and Prevention, some of which emails will be mentioned *infra*.

where disparities that would previously have been regarded as showing large increases now would show large decreases, as well as situations where, even though the NHDR would find a decrease in disparity between Year A and Year B, AHRQ would still find the disparity to be larger in Year B than in Year A.

By way of more detailed summary, Section A discusses the general problems with the health and healthcare disparities research activities of AHRQ arising from the agency's failure to recognize the patterns by which standard measures of differences between outcome rates of advantaged and disadvantaged groups tend to be systematically affected by the frequency of an outcome. Such patterns include that whereby, as the frequency of an outcome changes, the relative difference in experiencing the outcome and the relative difference in failing to experience the outcome tend to change in opposite directions. That is, to use the example cited by the National Center for Health Statistics (NCHS) when first explaining that determinations of directions of changes in the size of health and healthcare disparities would commonly turn on whether one examines relative differences in the favorable outcome or relative differences in the corresponding adverse outcome, as mammography rates generally increase, relative demographic differences in mammography rates (the favorable outcome) tend to decrease while relative demographic differences in rates of failure to receive mammography (the corresponding adverse outcome) tend to increase. Such patterns also include that whereby, as the frequency of an outcome changes, the absolute (percentage point) difference tends to change in the same direction as the smaller relative difference. I have treated this subject in a great many places with respect to activities of AHRQ and otherwise, and I will rely heavily on certain such treatments in Section A.³ Those treatments show that it is not possible for researchers or institutions to provide useful information on the strength of the forces causing health and healthcare outcome rates of advantaged and disadvantaged groups to differ (or, as it might otherwise be put, the differences in the health and healthcare-related circumstances of advantaged and disadvantaged groups reflected by their outcome rates) without consideration of the implications of the frequency of the outcomes examined. Given that essentially all health and healthcare disparities research to date, whether conducted or funded by AHRQ or otherwise, has failed to consider such implications, such research has provided little of value, and much that is misleading or incorrect, respecting such things as whether the strength of the forces

³ These places includes scores of articles, web pages, and online journal comments; dozens of presentations at conferences of public health, statistical, demography, and epidemiology organizations in North America and Europe; and seven methods workshops at the statistics, sociology, demography, epidemiology or law arms of universities in the United States ([American University](#) (2012), [Harvard University](#) (2012), [University of Kansas](#) (2013), [University of Minnesota](#) (2014), [University of Maryland](#) (2014), [George Mason University](#) (2014), and [University of California, Irvine](#) (2015)). One of the online comments that provides a fair summary of the key issues is my December 4, 2006 [comment](#) on an article on effects of marital status on longevity that you coauthored with Dr. Robert M. Kaplan (Marital status and longevity in the United States Population. *J Epidemiol Community Health* 2006;60:760-765). The same issues apply with respect to efforts to identify changes in the effects of marital status over time. See my March 5, 2012 [comment](#) on Syse and Kravdal (Changes over time in the effect of marital status on cancer survival. *BMC Public Health* 2011;11:804 (doi:10.1186/1471-2458-804)) and my November 7, 2011 [comment](#) on Bernsten (Trends in total and cause-specific mortality by marital status among elderly Norwegian men and women. *BMC Public Health* 2011; 11:537).

responsible for health and healthcare disparities has increased or decreased over time or otherwise is larger in one setting than another.

Inasmuch as you have served as an officer of the Medicaid Division of the Massachusetts Department of Public Welfare, I note here that many recent references discuss that the failure to understand the measurement issues addressed in those works has caused Massachusetts to include in its Medicaid pay-for-performance program a healthcare disparities element and to do so in a manner whereby the inclusion of that element is more likely to increase healthcare disparities than reduce them. That program will receive some attention in Section A. The subject is highly relevant to the AHRQ 2016 budget request for funding for research into the effect of incentive programs on disparities.

Further, while the attention to AHRQ in Section A is largely devoted to health and healthcare disparities research issues, the statistical issues the section addresses may apply to many data interpretation activities of AHRQ. These include activities involving interpretation of data on subgroup effects and any situation where an inference is drawn on the basis of the comparative size of two relative differences, subjects addressed in many of the references to be discussed below and also addressed in my comment mentioned in note 3 on the 2006 *Journal of Epidemiology and Community Health* article authored by you and Dr. Robert M. Kaplan, who I understand is now AHRQ's Chief Science Officer.⁴

Section B addresses issues in the NHDR's identification of increases or decreases in disparities arising from a change of methodology beginning with the 2010 report and possibly changes in the 2013 and 2014 reports. Prior to the 2010 report, the NHDR measured changes in the sizes of disparities in terms of absolute (percentage point) changes in relative differences in the adverse outcome, while requiring that an increase or decrease would be found only when that change in the relative difference is greater than one percentage point per year. Possibly commencing with the 2010 report, but certainly by the 2012 report, the NHDR identified changes in disparities in terms of the comparative size of absolute changes in the rates of the advantaged group and the disadvantaged group, an approach that effectively identifies changes in the size of disparities on the basis of changes in absolute differences between rates. Under this approach, in order for the NHDR to identify an increase or decrease in disparity, the difference between absolute changes (which is the same as the change in the absolute difference between rates) must be greater than one percentage point per year.

One implication of this change in method is to cause various situations that would have been deemed to involve increases in disparities on the basis of percentage point changes in relative differences in adverse outcomes to be deemed to involve decreases in disparities on the

⁴ Among other things, the comment discusses that the article explored a hypothesis that the proportionate effect of not being married on mortality would be greater on persons in poorer health than on persons in better health without recognizing the reasons to expect a factor to have the greater proportionate effect on the group with the lower baseline mortality rate (*i.e.*, persons in better health). See also my "[Race and Mortality Revisited](#)," *Society* (July/Aug. 2014) at 339-41.

basis of the comparative size of percentage point changes in rates of experiencing those outcomes. To use an example that I will present among others in Table 6 of Section B, between 2005 and 2010, the rate at which hospital patients 65 and older with pneumonia received a pneumococcal screening or vaccination increased from 44.7% to 88.3% for Hispanics and from 63.9% to 94.5% for non-Hispanic whites. Thus, the ratio of the Hispanic rate of non-receipt of screening or vaccination to the non-Hispanic white rate of nonreceipt of screening or vaccination increased from 1.53 (55.3%/36.1%) to 2.13 (11.7%/5.5%). Under the pre-2010 report methodology (which is also the methodology employed by NCHS in Healthy People 2010), the NHDR would have treated this as an overall increase in disparity of 60 percentage points and a yearly increase of 12 percentage points (though the report would have referred to the percentage point change as a “%” change), something that might well have warranted mention as a particularly large increase. Instead, however, apparently on the basis that the yearly increase over the five-year period was 2.6 percentage points greater for Hispanics than non-Hispanic whites (overall, an increase of 43.7 percentage points for Hispanics compared with 30.6 percentage points for non-Hispanic whites, hence a total difference between the two group’s changes of 13.1 percentage points), the disparity was treated in the 2012 NHDR as one of the fastest decreasing disparities.

Further, AHRQ (in the NHDR and elsewhere) continued to appraise the size of a disparity in terms of relative differences in adverse outcomes even as the NHDR measured changes in disparities in terms of absolute differences between rates. Thus, notwithstanding that the 2012 NHDR found a substantial decrease in the above-referenced Hispanic-non-Hispanic white disparity in pneumococcal screening or vaccination, examining the 2010 and 2005 disparities themselves, AHRQ would find a disparity that is substantially larger in 2010 than in 2005.

Moreover, because, as mentioned above and as will be explained more fully in Section A, given the rate ranges at issue, the change in measurement approach from (a) reliance on absolute changes in relative differences in adverse outcomes to (b) reliance on the comparative size of absolute changes in outcome rates will commonly result in situations where the change in method reverses the direction of disparity changes identified in the report.

A second implication of the change in the 2010 report is that in circumstances where adverse outcome rates are very low (especially where they are well below one percent for all groups), only in the case of an epidemic increase in the outcome would there be a possibility that the percentage point change experienced by one group would be a percentage point higher than that experienced by the other group change even over extended periods. Thus, the method would seem to preclude identifying any changes in disparities for such outcomes in the usual situations examined in the report. Yet, as will be discussed, reports continued to show changes in disparities regarding such outcomes.

Finally, though I am still attempting to clarify this issue with AHRQ staff, it appears that further modifications of the method for identifying changes in disparities in the 2013 and 2014 reports inadvertently implemented a procedure that would cause directions of changes in

disparities again to be consistent with changes in relative differences (as in the pre-2010 reports), but with magnitudes of changes that are different from the magnitudes of changes that would have been identified in the pre-2010 reports.

I emphasize, however, that the attention I give in Section B to changes in methodology in the NHDR should not be read to suggest that any of the methods is sound. The NHDR can only provide useful information if it addresses the issue discussed in Section A. The same holds for all other health and healthcare disparities research activities of AHRQ.

Finally, I note that, as a result of email exchanges or other interactions, the following persons in the government involved with health and healthcare research matters should be familiar with the issues addressed in Section A of this letter: Dr. Makram Talih, Associate Director for Science, Office of Analysis & Epidemiology, NCHS; Dr. Chesley Richards, Deputy Director for Public Health Scientific Services, Centers for Disease Control and Prevention; and Dr. Ernest Moy, Medical Officer, Center for Quality Improvement and Patient Safety, AHRQ.⁵ Clyde J. Behney, Interim Leonard D. Schaeffer Executive Officer of the Institute of Medicine, National Academy of Sciences, should also be familiar with those issues.

A. Issues Concerning Patterns by Which Standard Measures of Differences Between Outcome Rates Tend to be Systematically Affected by the Frequency of an Outcome and Implications of Those Patterns Respecting the Interpretation of Data on Demographic Differences in Outcomes Rates

This section discusses general problems with the health and healthcare disparities research activities of AHRQ arising from the agency's failure to recognize the patterns by which standard measures of differences between outcome rates of advantaged and disadvantaged groups tend to be systematically affected by the frequency of an outcome. As discussed in the introduction, I have addressed the pertinent measurement issues in a great many places. Among the more important papers are my article "[Race and Mortality Revisited](#)," *Society* (July/Aug. 2014),⁶ and my Federal Committee on Statistical Methodology (FCSM) 2013 Research Conference paper "[Measuring Health and Healthcare Disparities](#)" (FCSM paper). Also important is my guest editorial in the Spring 2006 issue of the American Statistical Association magazine *Chance*, titled "[Can We Actually Measure Health Disparities?](#)," which, along with the items referenced in note 6, form the bases for all published recognitions of the pattern whereby relative differences in favorable outcomes and relative differences in the corresponding adverse outcomes tend to change in opposite directions as the frequency of an outcome changes. Each

⁵ A representative of AHRQ also attended the October 2014 methods workshop at the Maryland Population Research Center, mentioned *infra* at page 7. I failed to secure the person's name, however.

⁶ "Race and Mortality Revisited" is an updating of my "[Race and Mortality](#)," *Society* (Jan./Feb. 2000), which along with my "[Divining Difference](#)," *Chance* (Fall 1994), formed the basis for the National Center for Health Statistics first to recognize the pattern by which relative differences in a favorable outcome and relative differences in the corresponding adverse outcome tend to change in opposite directions as the frequency of an outcome changes.

paper has graphical or tabular illustrations of the patterns discussed here, as does an annotated version of the [presentation](#) given at the 2013 FCSM conference. Many graphical and tabular illustrations may also be found in the methods workshops and conferences presentations collected on the [Conference Presentations](#) subpage of the [Measuring Health Disparities](#) page (MHD) of [jpscanlan.com](#). The most comprehensive and useful of these are probably the October 2012 workshop at Harvard's Institute for Quantitative Social Science ("[The Mismeasure of Group Differences in the Law and the Social and Medical Sciences](#)") and the October 2014 workshop at the Maryland Population Research Center of the University of Maryland ("[Rethinking the Measurement of Demographic Differences in Outcome Rates](#)"). Many more illustrations of the patterns may be found on the pages of [jpscanlan.com](#) devoted to measurement issues (along with their subpages)⁷ and in tables created in connection with various of the 140-plus online comments on journal articles collected on the [Journal Comments](#) subpage of MHD.⁸

As discussed in the above-referenced materials and many other places, research into demographic differences in outcome rates has been universally undermined by failure to recognize patterns by which standard measures of differences between outcome rates tend to be affected by the frequency of an outcome. The most notable of these patterns is that whereby the rarer an outcome the greater tends to be the relative difference between rates at which advantaged and disadvantaged groups experience the outcome and the smaller tends to be the relative difference between rates at which such groups avoid the outcome (*i.e.*, experience the opposite outcome). Table 1 below, which is a variation on Table 1 of "Race and Mortality Revisited" and Table 1 of the TDHCD brief, provides a simplified illustration of this pattern. The table is based on a situation where an advantaged group (AG) and a disadvantaged group (DG) have normally distributed scores on a test with mean scores that differ by half a standard deviation (and both distributions have the same standard deviation). The table shows that lowering a test cutoff (and thereby generally reducing the frequency of test failure while generally increasing the frequency of test passage) tends to (a) increase relative differences in failure rates (reflected by the ratio of the failure rate of the DG to the failure rate of AG, which increases from 1.85 to 2.60), but (b) reduce relative differences in pass rates (reflected by the ratio of the pass rate of AG to the pass rate of DG, which decreases from 1.27 to 1.09).⁹

⁷ The principal measurement pages are: [Measuring Health Disparities](#), [Scanlan's Rule](#), [Mortality and Survival](#), [Immunization Disparities](#), [Immune Disparities](#), [Educational Disparities](#), [Disparate Impact](#), [Discipline Disparities](#), [Lending Disparities](#), [Employment Discrimination](#), [Feminization of Poverty](#). The pages have close to 100 subpages addressing or elaborating on particular issues or studies.

⁸ Three particularly pertinent items are the October 9, 2012 [letter](#) to Harvard University mentioned in the first paragraph, which was written in conjunction with Harvard workshop noted in the text above; a 2009 presentation at the 9th International Conferences on Health Policy Statistics ("[Perverse Perceptions of the Impact of Pay for Performance on Healthcare Disparities](#)"); and a 2011 presentation given at the 3rd North American Congress of Epidemiology ("[Measuring Healthcare Disparities](#)"). Each of these items, however, reflects the misperception about the manner in which the NHDR measured health and healthcare disparities discussed at the beginning of Section B.

⁹ Most of my illustrations of these patterns (save in the 2006 *Chance* editorial) present the matter in terms of rate ratios with the higher rate as the numerator of the ratio. In such circumstances, the relative difference is the rate ratio minus 1.

Table 1. Illustration of effects on four measures of differences between pass (or fail) rates of lowering a cutoff from a point where 80% of AG passes to a point where 95% of AG passes (when mean scores differ by approximately half a standard deviation)

Cutoff	AG Pass	DG Pass	AG Fail	DG Fail	AG/DG Pass Ratio	DG/AG Fail Ratio	Abs Diff (perc pts)	DG/AG Fail Odds Ratio
High	80%	63%	20%	37%	1.27	1.85	17	2.35
Low	95%	87%	5%	13%	1.09	2.60	8	2.85

Table 1 also presents the absolute difference between rates and the ratio of DG’s failure odds to AG’s failure odds. I will defer the principal discussion of those measures for some paragraphs. I note at this juncture, however, that the table happens to reflect the following situation of the nature discussed in Section B. Assuming that the pass rates in the first and second rows reflected such things as immunization rates at an earlier and a later point in time, the NHDR pre-2010 methodology would find that the disparity increased by 75 percentage points. But the methodology subsequently implemented would find that the disparity decreased on the basis of the fact that the adverse outcome rate of DG declined by 24 points compared with decline of 15 percentage points for AG (a 9 percentage point difference that corresponds to the decline in the absolute difference between rates).

The pattern whereby reducing the frequency of an outcome tends to increase relative differences between rates of experiencing the outcome while reducing relative differences between rates of avoiding the outcome can be found in virtually any data where one can derive the rates at which advantaged and disadvantaged groups fall above and below various points on a continuum of values associated with experiencing an outcome or its opposite. Income and credit score data show that the lower the income or credit score, the greater tends to be the relative difference in rates of falling below it while the smaller tends to be the relative difference in rates of falling above it. Life tables show that the lower the age, the greater tends to be the relative difference in failing to survive to the age while the smaller tends to be the relative difference in surviving to the age. National Health and Nutrition Survey data show that generally reducing blood pressure will tend to increase relative differences in hypertension while reducing relative differences in rates of avoiding hypertension or that generally improving folate levels will tend to increase relative differences in rates of low folate while reducing relative differences in rates of adequate folate. And, while the described patterns will not be observed in every case,¹⁰ they are commonly evident in situations where an outcome increases or decreases over time or is larger in one setting than another. In addition to the references noted above, see the [Collected Illustrations](#) subpage of the [Scanlan’s Rule](#) page of [jpscanlan.com](#).

¹⁰ As noted in many places, observed patterns are functions of both (a) the frequency-related patterns described here and (b) the magnitudes of the differences between the underlying distributions of the advantaged and disadvantaged groups in the settings being compared (as well as any irregularities in the distributions). This point applies as well to the discussion *infra* about absolute differences and odds ratios.

Implications of the pattern include that as mortality declines, relative differences in mortality tend to increase while relative differences in survival tend to decrease; as rates of appropriate healthcare increase, relative differences in receipt of such care tend to decrease while relative differences in failing to receive such care tend to increase. Relative differences in mortality tend to be greater, while relative differences in survival tend to be smaller, for more survivable than less survivable cancers. Within populations where adverse outcomes are comparatively rare (*e.g.*, persons with high education or high income, British civil servants, Norway and Sweden, Massachusetts and Minnesota, the young, the healthy) relative demographic differences in adverse outcomes tend to be greater, while relative differences in the corresponding favorable outcomes tend to be smaller, than within populations where adverse outcomes are comparatively common.¹¹ And, as is implicit in the last sentence, a factor that affects an outcome rate will tend to cause a larger proportionate change in the outcome rate for groups with lower baseline rates while causing larger proportionate changes in the opposite outcome rates for other groups.)¹² Where inferences are drawn based on the comparative size of a factor's relative effect on a particular outcome for groups with different baseline rates for the outcome, examination of the comparative size of a factor's relative effect on the opposite outcome will commonly support an entirely different inference.¹³

As discussed in "Race and Mortality Revisited" (at 331-35) and the FCSM paper (at 11-12), in 2004 the National Center for Health Statistics (NCHS) first recognized that relative differences in favorable health and healthcare outcomes and relative differences in the corresponding adverse outcomes tend to change in opposite directions as the frequency of an outcome changes (though its belief that it could deal with that issue simply by choosing to measure disparities in terms of relative differences in adverse outcomes reflected a fundamental misunderstanding of society's purpose in examining outcome rates of advantaged and disadvantaged groups). As discussed more generally in those papers, no other federal agency involved with health and healthcare disparities research has yet shown an understanding that it is even possible for the two relative differences to change in opposite directions as the frequency of an outcome changes, much less that they tend to do so systematically.

¹¹ See especially my "[It's easy to misunderstand gaps and mistake good fortune for a crisis](#)," *Minneapolis Star Tribune* (Feb. 8, 2014).

¹² The common pattern whereby relative demographic differences in mortality decrease with age (which observers often note) while relative differences in survival increase with age (which is rarely noticed) could as well be described in terms whereby (a) being in the disadvantaged demographic group causes a larger proportionate increase in mortality among the young than the old but a larger proportionate decrease in survival among the old than the young, or (b) aging causes a larger proportionate increase in mortality for the advantaged group but a larger proportionate decrease in survival for the disadvantaged group. See the [Life Table Illustrations](#) subpage of the Scanlan's Rule page of [jpscanlan.com](#). See also Table 6 of the 2008 Nordic Demographic Symposium presentation "[Measures of Health Inequalities That Are Unaffected by the Prevalence of an Outcome](#)" and pages 11-12 of the October 9, 2012 [letter](#) to Harvard University mentioned in the first paragraph.

¹³ See "Race and Mortality Revisited" at 339-41.

The Institute of Medicine, which provides guidance to AHRQ on the NHDR and the National Healthcare Qualities Report, has yet to show an awareness of this pattern and it is doubtful that any of the data support contractors identified in introductory materials to the NHDRs are aware of the pattern.¹⁴

Federal agencies monitoring other types of demographic differences have no greater an understanding of these issues than agencies monitoring demographic differences in health and healthcare outcomes. Since at least 1994, agencies enforcing fair lending laws (including, among others, the Departments of Justice and Housing and Urban Development, the Comptroller of the Currency, and the Federal Reserve Board) have been encouraging lenders to relax lending standards under the mistaken belief that doing so will tend to reduce relative differences in adverse borrower outcomes. More recently, the Departments of Justice and Education have been encouraging public schools to relax public school discipline standards under the mistaken belief that doing so will tend to reduce relative racial differences in suspensions and expulsions. Both beliefs are the exact opposite of reality. While relaxing standards tends to reduce relative differences in meeting the standard, it tends to increase relative differences in failing to meet the standard. And because federal enforcement agencies continue to appraise the fairness of lending and discipline practices on the basis of relative differences in adverse outcomes, there exists the perverse situation where by relaxing standards in response to federal encouragements, lenders and public schools increase the chances that the government will sue them for discrimination. See the TDHCD brief generally and "Race and Mortality Revisited" at 343. Relatively succinct treatments of the failure of federal civil rights enforcement agencies to understand these issues may be found in my "[Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies](#)," *Amstat News* (Dec. 2012), and "[Things government doesn't know about racial disparities](#)," *The Hill* (Jan. 28, 2014).

An illustration of the depth of the misunderstanding by the Department of Justice may be found in its March 4, 2015 report on the racial impact of police and court practices of the city of Ferguson, Missouri. A premise of the report is that over policing and harsh court procedures played a large role in causing black residents of Ferguson to comprise a much higher proportion of persons experiencing adverse interactions with the police and the courts than they comprised of the city's population. But reducing the frequency of the interactions would tend to increase, not reduce, the proportion blacks comprise of persons experiencing those interactions. See the March 9, 2015 [letter](#) to officials of the Department of Justice and City of Ferguson mentioned in the first paragraph. See also the June 8, 2015 [letter](#) to officials of the City of Minneapolis mentioned in that, which addresses a recent American Civil Liberties Union study of police

¹⁴ This is not to suggest that pattern has been entirely overlooked by researchers outside NCHS. See discussion of treatments by researchers and scholars in the United States and abroad in the [Consensus Subpage](#) of the Scanlan's Rule page of [jpscanlan.com](#).

practices in Minneapolis that reflects the same failure of understanding as the Department of Justice Ferguson report.

Congress does not understand this issue any better than executive branch agencies. As discussed in "Race and Mortality Revisited" (at 342), the Individuals with Disabilities Education Improvement Act of 2004 contains a provision calling for actions in response to "significant discrepancies" in suspensions of special education students that typically would increase the size of the discrepancies as they are commonly measured. See the March 20, 2015 [letter](#) to the Senate Committee on Health, Education, Labor and Pensions mentioned in the first paragraph, which explains that a key premise of the recently-introduced Keep Kids in School Act – that it would tend to reduce relative racial differences in suspensions and expulsions – is the opposite of reality.

Appraisals of the comparative size of differences between outcome rates measured in absolute terms or in terms of odds ratios are unaffected by which outcome one examines.¹⁵ But in order for a measure to effectively quantify the strength of the forces causing the outcome rates of advantaged and disadvantaged groups to differ, the measure must remain constant when there occurs a general change in the frequency of an outcome akin to that resulting from a lowering of a test cutoff. And, like the two relative differences, the absolute difference and the difference measured by odds ratios tend to change systematically as the prevalence of an outcome changes. They do so, however, in more complicated ways than the two relative differences.

Roughly, as uncommon outcomes (less than 50% for both groups being compared) become more common, absolute differences tend to increase; as common outcomes (greater than 50% for both groups being compared) become even more common, absolute differences tend to decrease. The pattern is more difficult to predict when the outcome is neither common nor uncommon or either group's rates crosses 50% during the period at issue. Both of the two relative differences and the absolute difference can change in the same direction, in which case one can infer a true change in the strength of the forces causing the outcome rates to differ. But when a relative difference and the absolute difference do not change in the same direction, the absolute difference will change in the same direction as the smaller relative difference.¹⁶ Given the rate ranges at issue for most of the outcomes examined in the NHDR (with, especially in recent years, favorable outcome rates greater than 50% for both groups being compared), unless both of the relative differences and the absolute difference change in the same direction when the frequency of the outcome changes, the absolute difference will yield a result as to the direction of changes over time different from that yielded by the relative difference in adverse outcomes.

¹⁵ But see the September 20, 2013 University of Kansas School of Law faculty workshop paper "[The Mismeasure of Discrimination](#)" regarding some nuances of characterization respecting differences measured by the odds ratio.

¹⁶ See the introductory section of the [Scanlan's Rule](#) page regarding some minor departures from this pattern when one group's rate is above 50% and the other group's rate is below 50%.

As the frequency of an outcome changes and all measures do not change in the same direction, the difference measured by the odds ratio tends to change in the opposite direction of the absolute difference.

Table 2 is a variation on Table 5 of "Race and Mortality Revisited" (at 335) and Table 1 of the FCSM paper (at 13). It is based on the same specifications as Table 1 above.

Table 2. Standard measures of differences between outcome rates of an advantaged group (AG) and a disadvantaged group (DG) at four levels of prevalence

Row Num	Outcome	Situation	AG Fav OC Rt	DG Fav OC Rt	AG/DG Fav Ratio	DG/AG Adv Ratio	Abs Diff (Per points)	Odds Ratio
1	A	Year A or Low Performer	20.0%	9.0%	2.22	1.14	11.0	2.53
2	A	Year B or High Performer	40.0%	22.6%	1.77	1.29	17.4	2.28
3	B	Year A or Low Performer	70.0%	51.0%	1.37	1.63	19.0	2.24
4	B	Year B or High Performer	80.0%	63.4%	1.26	1.83	16.6	2.31

In the referenced papers the illustration is cast in terms of hiring patterns of different employers where the observer must endeavor to draw conclusions about the comparative size (or likelihood) of bias reflected in the various rows. The table is used in those papers to refute the notion that choice of measure of health and healthcare disparities involves a value judgment, as well as to demonstrate the unsoundness of each of the four measures shown in the table for quantifying the strength of the forces causing the outcome rates to differ.¹⁷ The table also usefully illustrates a sound measure of health and healthcare disparities. That method involves deriving from a pair of outcome rates the difference between the means of the underlying distributions and would indicate that such difference reflected by the outcome rates in each row would be half a standard deviation (just as it would be in Table 1 above).

¹⁷ The refutation of the value judgment argument was principally directed at Harper S, Lynch J, Meersman SC, *et al.* ([Implicit value judgments in the measurement of health disparities.](#) *Milbank Quarterly* 2010 (Mar);88(1):4-29)). The Harper article discussed the choice between the absolute difference and a relative difference without acknowledging that there existed a second relative difference and without apparent recognition that any time the direction of change in disparity indicated by the relative difference the observer happens to be examining differs from that indicated by the absolute difference, the unmentioned relative difference will necessarily have changed in the opposite direction of the mentioned relative difference and the same direction of the absolute difference. A more recent expression of the value judgment argument may be found in Mackenbach JP (Should we aim to reduce relative or absolute inequalities in mortality? *Eur J Pub Health* 2015;25(2):185). The Mackenbach article differs from the Harper article in that Mackenbach article and several of its references recognize that the two relative differences tends to change in opposite direction as the frequency of an outcome changes. See my May 27, 2015 [comment](#) on the Mackenbach article. See also my discussion of a 2012 Mackenbach *Social Science and Medicine* article in my August 18, 2014 [comment](#) on Marmot and Goldblatt (Importance of monitoring health inequalities. *BMJ* 2013;347:f6576). Earlier recognition of the pattern by Mackenbach and others is alluded to in "Race and Mortality Revisited" at 343.

I encourage AHRQ to review of the referenced tables, and other tables in those papers, with respect to those purposes.¹⁸ And I suggest that fully understanding the points in those papers regarding that particular table and related matters, one will have difficulty concluding other than that, without consideration of the points, the NHDR can provide very little of value and much that is misleading or incorrect. For example, were the NHDR to cause observers or Congress to opine about (or devote resources to studying why) general increases in appropriate healthcare rates led to increasing relative differences in failure to receive such care without consideration of the extent to which the latter increases were a function of the former increases, and without consideration of whether relative differences in receipt of care decreased, it would not be serving the purpose Congress intended in requiring the report.

The table, as adjusted, can also usefully demonstrates several things about the measurement of healthcare disparities in terms of absolute differences between rates, as has been commonly done in the pay-for-performance (P4P) context and, as I will show in Section B, as apparently was done in the NHDR at least in the 2012 report.¹⁹ In considering the points that follow, however, it is important to keep in mind that there exists no rational basis for maintaining that the strength of the forces causing the outcome rates to differ varies among the rows in Table 2.

Assume that rows 1 and 2 involve an uncommon healthcare outcome and rows 3 and 4 involve a common healthcare outcome. The odd and even numbered rows – with the higher values for favorable outcomes in the latter – can be alternatively regarded as reflecting outcome rates at different points in time when favorable outcome rates are increasing or at lower-performing and higher-performing hospitals at particular points in time (which alternative formulations are indicated in the second column of the table).

Now consider the matter with respect to efforts to determine the effects on healthcare disparities of incentive programs that generally increase favorable healthcare outcome rates, which would appear to be the subject of the AHRQ 2016 budget proposal item discussed above. By and large such efforts have focused on absolute differences between rates. Those employing that measure while examining the effects of incentive programs on uncommon outcomes will tend to find that incentive programs increase disparities; those relying on such measure while examining the effects of incentive programs on common outcomes will tend to find that such

¹⁸ Tables in those papers provide numerous illustrations of the referenced method for appraising the strength of the forces causing outcome rates of advantaged and disadvantaged groups to differ that is not affected by the frequency of an outcome, as do the Harvard and University of Maryland workshops.

¹⁹ The standard effect of improvements in care on disparities is easier to describe when disparities are measured in terms of relative differences in adverse outcomes. Very simply, improvements in care will tend to increase disparities. See discussion regarding Table 6 (at 19) of the FCSM Paper concerning the fact that the decision of NCHS to measure all health and healthcare disparities in terms of relative differences in adverse outcomes would cause what had been regarded as instances of decreasing disparities in immunization rates during a period of general increases in immunization now to be regarded as instances of increasing disparities in rates of failure to be immunized.

programs decrease disparities. As discussed in "Race and Mortality Revisited," in point of fact in the United States a focus on uncommon outcome has led to a perception that incentive programs will tend to increase disparities, while in the United Kingdom a focus on common outcomes has led to the perception that incentive programs will tend to reduce healthcare disparities. See also the [Pay for Performance](#) subpage of [Measuring Health Disparities](#) page of [jpscanlan.com](#) and the 2011 International Conference on Health Policy Statistics [presentation](#) mentioned in note 8. Neither perception has a sound statistical basis.

But the perception in the United States that incentive programs will tend to increase healthcare disparities led to a call for including a disparities element in P4P programs, with performance on the element to be evaluated either in on the basis of changes over time at particular provider institutions or on the basis of the sizes of disparities across provider institutions.

A P4P program relying on absolute differences to measure changes in disparities over time would find general improvements over time to be associated with increasing disparities for outcome A, but decreasing disparities for outcome B. A program that relied on absolute differences to evaluate cross-institution disparities, would find higher-performing institutions to have larger disparities for outcome A, but smaller disparities for outcome B, than lower-performing institutions. Simplified illustrations of these points may be found in Tables 9 and 10 (at 25-26) of the FCSM [presentation](#) and Tables 26 and 27 (at 111-12) of the Maryland Population Research Center [workshop](#).

The Massachusetts Medicaid P4P program, which may well serve as model for others, makes cross-provider judgments about disparities with respect to outcomes that have generally quite high rates and, in doing so, relies on a measure that is a function of absolute differences between rates²⁰ The approach therefore tends to systematically find smaller disparities at higher-performing institutions. Thus, the disparities element of the program tends to reward higher-performing institutions for reasons that may have nothing to do with cross-institution equity. Given that disadvantaged groups comprise a higher proportion of patients at lower-performing than higher-performing institutions, the diversion of resources to higher-performing institutions at the expense of lower-performing institutions will tend to exacerbate the disparities that result from the fact that disadvantaged groups are disproportionately served by lower-performing institutions.

This tendency, it should be noted, differs from the potential for P4P programs to generally favor higher-performing institutions. For here it is the disparities element of the program itself that tends to cause the diversion of resources away from institutions where disadvantaged groups comprise high proportions of patients.

²⁰ See the [Between Group Variance](#) subpage of the [Measuring Health Disparities](#) page of [jpscanlan.com](#).

AHRQ cannot conduct (or monitor) research into the effect of incentive programs on healthcare disparities without knowing these things. Currently, however, so far as I am able to tell, neither researchers seeking AHRQ funding, nor AHRQ personnel monitoring funding, seem even to be aware that choice of measure can affect conclusions about the association of the size of disparities with other factors. Exemplary of the situation is a multimillion dollar AHRQ contract with the Institute for Medicine and Public Health of the Vanderbilt University Medical Center aimed at evaluating the effectiveness of quality improvement in reducing disparities in healthcare. The contract yielded a 475-page, peer-reviewed report, issued in August 2012, that cites 4258 sources that were examined in fulfillment of the contract. But the report reflects no recognition whatever of the ways the various measures employed in those studies may be affected by the frequency of an outcome or even that it is possible that various measures could yield different conclusions as to directions of changes in disparities in conjunction with quality improvements. In discussing findings of various studies, the report does not identify the measures that were used. See the [AHRQ's Vanderbilt Study](#) subpage of the Measuring Health Disparities page of [jpscanlan.com](#).²¹

A study of such nature cannot provide useful information respecting the effects of improvement in healthcare on healthcare disparities. But the same may be said of any activity involving research into health and healthcare disparities that fails to consider, not merely that different measure may commonly yield different results about such things as whether disparities are increasing or decreasing, but the implications of patterns by which each measure tends to be affected by the frequency of an outcome with respect to the utility of the measure for quantifying a disparity.

B. Issues Concerning Effects of Modifications in Methodology for Analyzing Changes in Healthcare Disparities Over Time in the NHDR Beginning With the 2010 Report

This section addresses some anomalies regarding NHDR determinations of the directions of changes in healthcare disparities, and quantification of those changes, resulting from modification of the methodology for analyzing changes in disparities beginning with the 2010 NHDR. As noted in the Introduction, the attention given to these anomalies should not be read as suggesting that any of the methods are sound.

Consistent with Healthy People 2010, the NHDR has attempted to measure disparities in terms of relative differences in adverse outcomes. While I have been writing since 2007 about measurement problems in the NHDR, I have only recently fully understood this. In fact, in many places, apparently on the basis of the fact that for the NHDR will find a disparity meaningful if either the relative difference in the favorable outcome or the relative difference in the adverse outcomes is greater than 10%, I have stated that the NHDR (or AHRQ) will measure disparities

²¹ The web page indicates that the value of the contract was approximately \$6 million. "Race and Mortality Revisited" (at 333) gives a figure of \$10 million. The latter figure is based on a contract document indicating an anticipated maximum amount of \$10 million. I do not know which figure is more accurate.

in terms of whichever relative difference is larger.²² That confusion on my part, however, is not pertinent to the subject of this section.

Prior to the 2010 report, consistent with Health People 2010, the NHDR measured changes in disparities in terms of percentage point changes in relative differences in adverse outcomes. See, *e.g.*, the 2006 report at 5 note xix. But in contrast to Health People 2010 which uses the terms “percentage point,” the NHDRs have consistently used the percent sign (%) to refer to percentage point changes and differences, while also using the percent sign to refer to percent changes and differences.²³ Whether or not it is technically incorrect to use the percent sign or the term “percent” to refer to percentage point differences or changes, it is extremely bad practice in a scientific document and leads to much confusion. Such confusion can be particularly great (a) among persons who do not know the difference between a percent and a percentage point difference or change; (b) among persons who do not know that percentage point changes or differences (whether regarding (i) outcome rates themselves, (ii) relative difference between outcome rates, or (iii) absolute differences between outcome rates) can mean very different things from percent changes or differences in terms of quantification; (c) among persons who do not know that conclusions about which of two groups experiences a larger change in its rate that are based on percentage point changes often differ from conclusion that are based on percent changes; and (d) among person who do not know that directions of changes in absolute differences between rates are often the opposite of directions of changes in relative differences between rates.²⁴

The figures in Table 1 above provide the basis for a simple illustration of the NHDR pre-2010 methodology for identifying directions of changes in health and healthcare disparities and quantifying such changes. The lowering of the cutoff caused the failure rates to decline from 37% to 13% for DG and from 20% to 5% for AG. That caused the DG/AG failure rate ratio to increase from 1.85 to 2.60. Assuming that the failure rates in the first row reflect adverse healthcare outcome rates at an earlier point in time and the failure rates in the second rate reflect those rates at a later point in time, Health People 2010 would regard this as a 75 percentage point increase in disparity, as would pre-2010 NHDRs (though they would characterize it as a 75% increase). Table 3 illustrates the matter using only the necessary figures.

²² See, *e.g.*, the references in note 8. See also the comments of [March 17, 2014](#), and [February 14, 2013](#), on Epstein K (Persistent health disparities in the US signal for new thinking. *BMJ* 2012;345:e6204 doi: 10.1136/BMJ.e620).

²³ As indicated in the Section A, discussion of the measurement of disparities in Healthy People 2010 should not be read as suggesting that the Healthy People 2010 approach is sound. Further, Healthy People 2010 suffers from a number of presentation issues, including the failure to make clear to readers that there are circumstances where the relative difference in the favorable outcome would show a directions of change that is the opposite of that shown by the relative difference in the adverse outcomes. See FCSM paper at 26-27.

²⁴ The [Percentage Points](#) subpage of the [Vignettes](#) page of [jpscanlan.com](#), which discusses this matter at some length, gives considerable attention to the usage in the NHDR.

Table 3. Illustration of method of appraising changes in healthcare disparities under the NHDR pre-2010 approach (using hypothetical data from Table 1)

Period	DG Adverse Rt	AG Adverse RT	DG/AG Adverse Rate Ratio
Year A	37%	20%	1.85
Year B	13%	5%	2.60

The percentage point change would have to be statistically significant and be greater than one percentage point per year for the NHDR to regard the change to involve an increase or decrease in disparity. But that aspect of the matter can be ignored for instant purposes.

The method of identifying directions of change and quantifying those changes was changed in the 2010 report. In describing measures of changes in quality, access, and disparities, the report states (at 6):

As in past NHQRs [referring to quality and access], regression analysis is used to estimate annual rate of change for each measure. Annual rate of change is calculated only for measures with at least 3 years of data. For most measures, trends include data points from 2001-2002 to 2007-2008. We label measures going in a favorable direction at a rate exceeding 1% per year as improving, going in an unfavorable direction at a rate exceeding 1% per year as worsening, and changing at a rate less than 1% per year as not changing.

This year, for the first time, we introduce a similar method for assessing change in disparities using regression results. When a selected group's rate of change is at least 1% higher than the reference group's rate of change, we label the disparity as improving. When a selected group's rate of change is at least 1% lower than the reference group's rate of change, we label the disparity as worsening. When the difference in rates is less than 1%, we label the disparity as no change.

In the case of quality and access (subjects of the first quoted paragraph), "1%" change meant a one percent (*i.e.*, relative) change in whichever outcome yields the larger such change (as made clear at page 28 of the 2009 report). I note, for clarity, that when measuring changes in quality or access rates (where one is examining overall outcome rates rather than comparing outcome rates of different groups), whether one examines relative changes in the favorable outcome or the adverse outcome, while affecting the sizes of changes, will not affect determinations of whether the situation is improving or worsening.

In the case of disparities (subject of the second quoted paragraph), as with the usage employed in the pre-2010 monitoring of disparities, the "1%" difference presumably meant a one percentage point difference (and the language impliedly meant one percentage point per year). But what is not clear is whether determinations regarding the one percentage point difference in the two group's yearly changes is based on (a) relative (percent) changes in each group's rate or

(b) absolute (percentage point) changes in each group's rate.²⁵ The same holds for similar language in the 2011 and 2012 reports.²⁶

The phrase "rates of change" commonly refer to relative changes. If the report was in fact examining percentage point differences between relative changes in the rates of the two groups being compared, such approach would necessarily identify a direction of change in disparity that would be consistent with the pre-2010 approach (assuming that the report examines relative changes in the adverse outcome). Determinations of the magnitude of what are characterized as changes in disparities, however, could differ substantially from those under the pre-2010 approach. For (a) the percentage point difference between the relative changes in two groups rates between two points in time is different from (b) the percentage point difference between the relative differences at the two points in time, even though they are correlated qualitatively (*i.e.*, with respect to directions). Thus, statements such as those in Table H.4 of the 2010 report (at 10), and like statements in the reports for the following years, that decisions about improvements or worsening of disparities were based on one percentage point changes in disparities would not be correct even if the report was basing determinations of changes on

²⁵ At page 251, in the Priority Populations chapter, the 2010 report describes a method of measuring changes in disparities that is the same as that employed in earlier reports. I assume, however, that this language (which is identical to that found at page 181 of the 2009 report, even as to time periods examined) was inadvertently carried over from the earlier report. The language does not appear in later reports.

²⁶ The 2011 report states (at 6) after describing measures of changes in quality and access:

A similar method for assessing change in disparities using weighted least squares regression results is used. When a selected group's rate of change is at least 1% higher than the reference group's rate of change and this difference in rates of change is statistically significant, we label the disparity as improving. When a selected group's rate of change is at least 1% lower than the reference group's rate of change and this difference in rates of change is statistically significant, we label the disparity as worsening. When the difference is less than 1% or not statistically significant, we label the disparity as static. As with trends, because of the addition of significance testing, this year's results cannot be compared with results in previous reports.

The 2012 report states (at 7) after describing measures of changes in quality and access:

Weighted least squares regression was also used to assess change in disparities. When a selected group's rate of change is at least 1% higher than the reference group's rate of change and this difference in rates of change is statistically significant, we label the disparity as improving. When a selected group's rate of change is at least 1% lower than the reference group's rate of change and this difference in rates of change is statistically significant, we label the disparity as worsening. When the difference is less than 1% or is not statistically significant, we label the disparity as static.

The 2011 report (at 235) and the 2012 report (at 29 and 232), after using language similar to that from the reports' pages 6 (2011) and 7 (2012) quoted above, refer to 1% changes in disparities rather than 1% differences in changes in rates. This seems to suggest that the size of percentage point differences in changes in rates would equal the size of percentage point differences in disparities. But this would only be the case if both the changes in rates for the groups being compared and the disparities were appraised in absolute terms.

percentage point differences between the two groups relative rates of change. These points are illustrated in Table 4.

Table 4. Illustration of relative changes in adverse outcome rates of AG and DG (using hypothetical data from Table 1)

Group	Year A Adverse Rt	Year B Adverse Rt	Relative Change
AG	20%	5%	75.0%
DG	37%	13%	64.9%

According to the figures in Table 4, the relative decrease in the adverse outcome rate for AG (75.0%) was 10.1 percentage points greater than the relative decrease for DG (64.9%). Thus, as with the pre-2010 approach, the disparity had increased. But the increase would be based on a 10.1 percentage point difference in rates of change, which is different from the 75 percentage point increase in the relative difference between the two groups' rates that would have been identified under the pre-2010 approach.

On the other hand, if the report was examining percentage point differences between absolute (percentage point) changes in the rates of the two groups being compared, such approach would often identify a change of disparity that was in the opposite direction of the change that would have been identified under the pre-2010 approach. The size of the percentage point difference between absolute changes, however, would correspond to the size of the percentage point change in the absolute difference. These points are illustrated in Table 5.

Table 5. Illustration of absolute changes in adverse outcome rates of AG and DG (using hypothetical data from Table 1)

Group	Year A Adverse Rt	Year B Adverse Rt	Perc Point Change
AG	20%	5%	15
DG	37%	13%	24

The table shows that DG's absolute decrease (24 percentage points) was 9 percentage points greater than AG's absolute decrease (15 percentage points). The 9 percentage point difference corresponds to the 9 percentage point decrease in the absolute difference, which was initially 17 percentage points (37% - 20%) but declined to 8 percentage points (13% - 5%).

Further, although this approach would find a decrease in disparity, assuming AHRQ continued to measure disparities in terms of relative differences in adverse outcomes, it would continue to find a substantially larger disparity at the end of the period than at the beginning of the period.

The approach of measuring changes based on the percentage point differences between the percentage point changes in the two groups' rates would not always find the direction of change to be different from the direction that would be identified on the basis of changes in

relative differences in the adverse outcome. For reasons explained in Section A, the two approaches could find the same direction of change (a) when both of the relative differences and the absolute difference change in the same direction and (b) when the absolute difference changes in the same direction as the relative difference in the adverse outcome and the opposite direction of the relative difference in the favorable outcome. But even in the cases where the approach of measuring percentage point changes in absolute differences yields the same directions of changes as the pre-2010 method, the approach would not be identifying percentage point changes in the disparities that the report was purportedly relying on (*i.e.*, the relative differences in the adverse outcomes).

I have been unable to figure out how the NHDR measured disparities in 2010 and 2011 (even putting aside the issue of outcome rates that are below 1% for each group, which I discuss below). But, based on advice from AHRQ staff,²⁷ I have come to the view that at least in the 2012 report, AHRQ attempted to measure changes in disparities in terms of percentage point differences between the advantaged and disadvantaged groups' absolute changes in outcome rates and continued to attempt to do so through the 2014 report.

Table 6 is a slightly modified version of Table 5 of the 2013 FCSM presentation. It shows 4 situations highlighted in Table H.2 of the 2012 NHDR (at 14) as among the fastest decreasing disparities, even though the relative difference in the adverse outcome had increased, occasioning the same anomaly as that just noted with respect to hypothetical data. It is important to recognize that the measures in this table are measures of disparities at two points in time rather than measures of each groups' change in rates. But the table is nevertheless satisfactory for instant purposes.

When initially creating this table, I was uncertain whether the NHDR found decreasing disparities notwithstanding increases in relative differences in adverse outcomes because the NHDR was relying on (a) relative differences in favorable outcomes (which would correspond with, as to directions of changes, those identified on the basis of comparisons of the two groups' relative changes in favorable outcome rates) or (b) absolute differences between rates (which would correspond with, as to both directions of changes and size of changes, those identified on the basis of comparisons of the two groups' absolute changes in favorable or adverse outcome rates). I am now persuaded that the 2012 report's findings of decreasing disparities was based on the fact that the percentage point decreases in the adverse outcome rates were greater for the disadvantaged groups than the advantaged groups.²⁸

²⁷ I refer to a emails from Karen Chaves dated May 22, 2013, and from Ernest Moy dated May 14, 2015.

²⁸ I mention adverse outcome rates because I assume it was the adverse outcome rates that were examined. Under an approach based on the comparative size of absolute changes in rates, however, it does not matter which outcome one examines.

Table 6. Situations highlighted in 2012 NHDR as among the fastest decreasing disparities where relative differences in adverse outcome rates increased

Ref	Year	AG Fav Rt	DG Fav Rt	AG/DG Fav Ratio	DG/AG Adv Ratio	AbsDf ²⁹ Perc Points	EES
3	2006	66.50%	49.40%	1.35	1.51	17	0.44
3	2010	83.10%	72.40%	1.15	1.63	.11	0.36
4	2005	63.90%	45.70%	1.40	1.50	18	0.46
4	2010	94.50%	91.70%	1.03	1.51	3	0.21
10	2005	63.90%	44.70%	1.43	1.53	19	0.49
10	2010	94.50%	88.30%	1.07	2.13	6	0.40
11	2005	57.90%	41.50%	1.40	1.39	16	0.41
11	2010	92.90%	87.40%	1.06	1.77	6	0.32

In any event, the table show that, though the 2012 NHDR characterized these situations as reflecting the fastest decreasing disparities, the relative differences in adverse outcomes on which the NHDR purports to rely were increasing. In the case of the outcomes identified as 10 and 11, the contrast in terms of the size of the changes in disparities, as well as the difference in direction, would be striking. And, as with the hypothetical illustration, even though the NHDR finds substantial decreases in disparities, AHRQ would find larger disparities in the later period than the earlier period.³⁰

The final column (EES, for estimated effect size) contains the measure mentioned in Section A (at 12-13) whereby one may derive from the pairs of rates the difference between means of the underlying risk distributions. That measure shows a decrease in disparity in each case.

Assuming that possibly as early as the 2010 report, and certainly in the 2012 report, the NHDR was measuring changes in disparities on the basis of percentage point differences between the advantaged and disadvantaged groups' percentage point changes in rates, I have not been able to figure out how the NHDR could find changes in disparities with respect to rare adverse outcomes. To identify a change, the report requires that the difference between the

²⁹ In many presentations regarding these issues (including the FCSM presentation), mainly for space considerations, I have identified absolute changes with a decimal preceding the number. While this approach is less likely to lead to confusion than the use of the word "percent" or the percent sign to denote a percentage point change or difference, it is also problematic and probably incorrect. More recently, I have eliminated the decimal point and spelled out or abbreviated "percentage point," as in the tables of "Race and Mortality Revisited" and the University of Maryland workshop.

³⁰ The numbers are based on the ordering in Table H.2 of the report, and, as indicated in Table 14 of the FCSM presentation, reflect the following disparities: (3) black-white disparity for short-stay nursing home residents who were assessed and given pneumococcal vaccination; (4) Asian-white disparity for hospital patients age 65+ with pneumonia who received a pneumococcal screening or vaccination; (10) Hispanic-white disparity for hospital patients age 65+ with pneumonia who received a pneumococcal screening or vaccination; and (11) Hispanic-white disparity for hospital patients age 50+ with pneumonia who received an influenza screening or vaccination.

advantaged and disadvantaged groups' percentage point changes must be at least one percentage point per year. With respect to situations where adverse outcome rates are well below 1% (such as new AIDS case, hospital admissions for lower extremity amputations, and maternal mortality) it would be impossible to find situations where the two group's percentage point changes in that outcome (or its opposite) differed by one percentage point (even when the two groups' rates changed in opposite directions), save in circumstances of an epidemic increase in the outcome.

Nevertheless, Table H.3 (at 15) of the 2010 report, Table H.3 (at 11) of the 2011 report, and Table H.2 (at 14) of the 2012 report show a number of changes for these and other rare outcomes. I am awaiting a response to an inquiry of AHRQ staff regarding how changes in disparities of this nature were identified .

A final issue concerns whether changes in methodology in the 2013 and 2014 reports altered the 2012 approach to one of analyzing disparities in terms of the comparative sizes of relative changes in outcomes rates (presumably the adverse outcome) of the groups being compared.

I have been advised by AHRQ staff that the NHDR continues to employ the approach whereby determinations of the directions of changes, and the size of those changes, are based on the difference between the size of the percentage point changes in the rates of the advantaged and disadvantaged groups. But language in the 2013 and 2014 reports raises a question as to whether that is the case.

The 2013 report added this language (at 15-16, emphasis added):

A new approach to assess change in disparities is introduced this year. First, a selected group's rate of change and its reference group's rate of change are calculated using weighted least squares regression. Next, this difference in rates of change is assessed for statistical significance. Then, *the difference in rates of change relative to the reference group's baseline estimate is calculated.*

When the difference in rates of change is significant and when *the difference relative to the reference group's baseline* is greater than 1% per year, we label the disparity as improving if the selected group's rate is higher than the reference group's rate and worsening if the reverse.

When the difference relative to the reference group's baseline is less than or equal to 1% or the difference is not statistically significant, we label the disparity as static.

I do not understand the meaning of the highlighted phrases respecting "relative to the reference group's baseline rate" and am unfamiliar with a procedure whereby each group's change would be examined relative to the reference (advantaged) group's rate.

Richard Kronick, Ph.D., Director
Agency for Healthcare Research and Quality
July 1, 2015
Page 23

In any event, the Introduction and Methods section of the 2014 report contains the following language (at 14):

For each group, estimates were divided by earliest estimate so that earliest indexed estimate equaled one and subsequent indexed estimates were relative to the earliest estimate.

This seems mean that each group's absolute change is divided by the group's earliest estimate. If so, the report would now appear to be again measuring changes in relative terms, but quantified in the manner shown in Table 4 above. This is another matter on which I am awaiting clarification from AHRQ staff.

It is possible that I have misinterpreted some of the matters addressed in this section. There seems little question, however, that, at least with respect to outcomes addressed in Table 6, the 2012 NHDR was reporting as among the fastest decreasing disparities situations where the relative differences in adverse outcomes on which the NHDR purports to be measuring disparities had increased. AHRQ should determine the reasons for that anomaly and the other possible anomalies discussed above.

More important, however, AHRQ should examine whether the NHDR can provide anything of value regarding whether the forces causing health and healthcare outcome rates of advantaged and disadvantaged groups to differ, or regarding how particular policies may affect those forces, while measuring health and healthcare disparities without consideration of the way the measures employed are affected by the frequency of an outcome.

Sincerely,

/s/ James P. Scanlan

James P. Scanlan

cc: Robert M. Kaplan, Ph.D.
Chief Science Officer
Agency for Healthcare Research and Quality