

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

April 6, 2016

**ELECTRONICALLY TRANSMITTED**

Felice J. Levine, Chair of the Board of Directors  
Wendy A. Naus, Executive Director  
Officers and Directors of the Board of Directors  
Consortium of Social Science Associations  
1701 K Street NW, Suite 1150  
Washington, DC 20006

Re: Recommendation That COSSA (a) Form a Committee to Address the Unsoundness of Statistical Analyses of Demographic Differences That Fail to Consider Ways That Measures Tend to Be Systematically Affected by the Frequency of an Outcome and (b) Advise Arms of the U.S. Government That Reducing the Frequency of an Outcome Tends to Increase Relative Differences in Outcome Rates and the Proportions More Susceptible Groups Comprise of Persons Experiencing the Outcome

Dear Chair Levine, Executive Director Naus, and other Officers and Directors of the Board of Directors of the Consortium of Social Science Associations:

This letter has two purposes. One purpose, which involves what could be a long-term undertaking by the Consortium of Social Science Associations (COSSA), is to urge the organization to form a committee to examine the ways analyses by social scientists and others of demographic differences are undermined as a result of the failure to recognize patterns by which standard measures of differences between outcome rates tend to be systematically affected by the frequency of an outcome. A second purpose of the letter, which involves a matter COSSA can address immediately, is to urge COSSA to formally advise arms of the United States Government that a statistical belief underlying important civil rights law enforcement policies – specifically, that relaxing standards and otherwise reducing the frequency of an adverse outcome will tend to reduce relative differences in rates of experiencing the outcome and reduce the proportions groups most susceptible to the outcome make up of persons experiencing the outcome – is incorrect. As explained below, reducing the frequency of an outcome in fact tends to increase both relative differences in rates of experiencing the outcome and the proportions the most susceptible groups make up of persons experiencing the outcome.

This letter is similar in purpose to an October 8, 2015 letter to the [American Statistical Association](#) (ASA)<sup>1</sup> and a March 29, 2016 letter to the [Population Association of America and Association of Population Centers](#) (PAA & APC). The ASA letter is currently being considered by a subcommittee of ASA's Science and Public Affairs Advisory Committee. I have yet to be advised of any action by the PAA or APC.

Section A provides a brief summary of the measurement issues pertinent to the recommendations in this letter, with a focus principally on the pattern whereby the rarer an outcome the greater tend to relative differences between rates at which advantaged and disadvantaged experience the outcome and the smaller tend to be relative differences between rates at which such groups avoid the outcome. This is a pattern recognized by the National Center for Health Statistics (NCHS) more than a decade ago and that has been discussed in ASA publications since 1994, but that remains largely unknown among persons examining demographic differences in outcome rates, including leaders and members of COSSA's Governing Associations, as well as leaders, members, and researchers of COSSA's affiliated Membership Organizations, Centers and Institutes, and Colleges and Universities.

Section B describes materials apart from the ASA and PAA/APC letters discussing the measurement issues in detail and demonstrating that to date little research into demographic differences in outcome rates has been sound.

Section C describes recent developments that reflect on the urgency of the recommended COSSA actions. The developments include (a) a lawsuit recently brought by the United States Government against the City of Ferguson, Missouri that is based on an understanding of elementary statistics that is the exact opposite of reality and (b) recent actions of NCHS and the Centers for Disease Control and Prevention (CDC) regarding the measurement of health and healthcare disparities that, among other things, repudiate a decade of healthcare disparities research relying on earlier NCHS guidance.

Section D addresses the measurement issues in somewhat greater detail than Section A, including discussion of the patterns by which absolute differences between rates and odds ratios tend to be affected by the frequency of an outcome and illustrates a measure of the differences in the circumstances of advantaged and disadvantaged group reflected by their favorable or adverse outcome rates that is unaffected by the frequency of the outcome. The section also discusses implications of recent NCHS actions.

---

<sup>1</sup> To facilitate consideration of issues raised in letters such as this I include links to referenced materials in electronic copies of the letters. I do not include links with every mention of an item, but attempt to include them often enough to make online items readily available without the reader's having to search for an earlier link. Electronic copies of letters such as this are available by means of the [Institutional Correspondence](#) subpage of the [Measuring Health Disparities](#) page of [jpscanlan.com](#). In this case, electronic copies of the letter are being emailed to the addressees and a link to the letter will be posted on the ASA Connect section of the American Statistical Association website. Since the online copy of this letter may be corrected, I include here a [link](#) to the most recent version (which, if containing corrections, will indicate on the first page when it was last corrected).

Felice J. Levine, Chair, et al.  
Consortium of Social Science Associations.  
April 6, 2016  
Page 3

## A. Summary

### 1. This Letter's Broader Recommendation

For reasons related to features of the risk distributions underlying rates at which advantaged and disadvantaged groups experience favorable or adverse outcomes, standard measures of differences between such rates tend to be systematically affected by the frequency of outcome. The pattern most pertinent to analyses of demographic differences in the law and the social and medical science is that whereby the rarer an outcome, the greater tend to be relative differences between rates of experiencing the outcome and the smaller tend to be relative differences between rates of avoiding the outcome.<sup>2</sup> The pattern has been recognized by all scholars who have considered my articulations of it in ASA publications and elsewhere,<sup>3</sup> including NCHS statisticians in a number of official and official papers between 2004 and 2009.<sup>4</sup>

A simple illustration of this pattern is set out in Table 1 below. The table is based on a situation where the means of normal test score distributions of an advantaged group (AG) and a disadvantaged group (DG) differ by half a standard deviation and both distributions have the same standard deviation. The table presents pass and fail rates for each group at two cutoff points, along with the ratio of AG's pass rate to DG's pass rate and the ratio of DG's fail rate to AG's fail rate at each cutoff. It also shows the absolute (percentage point) difference between rates and the ratio of DG's failure odds to AG's failure odds, though I will defer substantial discussion of those measures until Section D.

---

<sup>2</sup> A more precise description of the pattern would state, rather than "the rarer an outcome," "the more the outcome is restricted toward either end of the overall distribution." But I have characterized the pattern in the manner done in the text above for some time and those discussing it have not been confused by the usage. Thus, I am not at this time inclined to depart from the usage in the text. Similar considerations apply to my use of the term "frequency" used here or the term "prevalence" that I have used in other places. See Section A.8 of the [Scanlan's Rule](#) page of [jpscanlan.com](#).

<sup>3</sup> See (a) Lambert PJ, Subramanian S. [Group inequalities and "Scanlan's Rule": Two apparent conundrums and how we might address them](#). Working Paper 84/2014, Madras School of Economics (2014); (b) Lambert PJ, Subramanian S. Disparities in Socio-Economic outcomes: Some positive propositions and their normative implications. *Soc Choice Welf* 2014;43:565-576); (c) [Eikemo TA, Skalicka V, Avendano M. Variations in health inequalities: are they a mathematical artifact?](#) *International Journal for Equity in Health* 2009;8:32; (d) Houweling TAJ, Kunst AE, Huisman M, Mackenbach JP. [Using relative and absolute measures for monitoring health inequalities: experiences from cross-national analyses on maternal and child health](#). *International Journal for Equity in Health* 2007;6:15; and (e) Carr-Hill R, Chalmers-Dixon P. [The Public Health Observatory Handbook of Health Inequalities Measurement](#). Oxford: SEPHO; 2005.

<sup>4</sup> See (a) Keppel KG, Percy JN. Healthy People 2010: Measuring Disparities in Health. *Chance* 2009;22(1):6-7; (b) Keppel KG, Pamuk E, Lynch J, et al. [Methodological issues in measuring health disparities](#). *Vital Health Stat* 2005;2 (141); (c) Keppel KG, Percy JN. Measuring relative disparities in terms of adverse events. *J Public Health Manag Pract* 2005;11(6):479-483.; (d) Keppel KG, Percy JN, Klein RJ. [Measuring progress in Healthy People 2010](#). Healthy People Statistical Notes. No. 25. Hyattsville, MD: National Center for Health Statistics. 2004.

The table shows that when the cutoff is lowered from the higher point to the lower point, the ratio of DG’s failure rate to AG’s failure rate increases from 1.85 to 2.60, while the ratio of AG’s pass rate to DG’s pass rate decreases from 1.27 to 1.09.<sup>5</sup> One would observe the same pattern if, instead of lowering the cutoff, we were able to improve test performance sufficiently to enable everyone with scores falling between the two points to reach the higher cutoff. The table thus illustrates how generally decreasing failure rates and generally increasing pass rates, tends to increase relative differences in rates of test failure (the decreasing outcome), while reducing relative differences in rates of test passage (the increasing outcome).<sup>6</sup>

**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

The following, somewhat overlapping points should be borne in mind regarding the illustration in Table 1. First, researchers associated with COSSA and its member entities commonly rely on relative differences in favorable outcomes or relative difference in adverse outcomes while not being aware that the relative difference in the opposite outcome could, or commonly would, yield a different conclusion. So far as the published record reveals, overwhelmingly, if not universally, researchers at such entities are unaware that lowering a test cutoff tends to increase relative differences in failure rates while reducing relative differences in pass rates.

<sup>5</sup> While I commonly refer to patterns of relative differences in this letter, the tables I use actually present rate ratios. The relative difference is the rate ratio minus 1 where the rate ratio is above 1 (in which case the larger the rate ratio the larger the relative difference) and 1 minus the rate ratio where the rate ratio is below 1 (in which case the smaller the rate ratio the larger the relative difference). One should be careful not to mistakenly refer to the rate ratio (which can also be termed “risk ratio” or “relative risk”) as the relative difference. But the distinction between the rate ratio and the relative difference is not pertinent to the discussion here of patterns by which relative differences tend to be affected by the frequency of an outcome, save with respect to the fact that one may not make the same points about patterns of changes in rate ratios that one may make about patterns of changes in relative differences. In recent years I commonly present the rate ratios for both outcomes with the larger figure in the numerator. Thus, as to both outcomes, the larger the rate ratio, the larger the relative difference. Choice of numerator in the ratio, however, has no bearing on the described patterns by which the two relative differences tend to be affected by the frequency of the outcome.

<sup>6</sup> Though I discuss failure differences first, the table presents pass rates first. I do so partly out of convention and partly because many recent discussions of the issues employing the example in Table 1 discuss pass differences first. See “[Things DOJ doesn’t know about racial disparities in Ferguson](#),” *The Hill* (Feb. 22, 2016), “[Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies](#),” *Amstat News* (Dec. 2012), and other articles referenced *infra* at page 13.

Second, as reflected in Table 1 and as will be made clear throughout this letter, relaxing a standard and thereby reducing the frequency an outcome tends to increase relative differences between rates at which advantaged and disadvantaged groups experience the outcome. Yet, as discussed further with regard to this letter's narrower recommendation, many civil rights law enforcement policies are based on the belief that reducing the frequency of an outcome tends to reduce relative differences in experiencing it. That mistaken belief among civil rights enforcement agencies is able to persist because it is shared by a substantial part of the social science research community including researchers associated with COSSA and its members. See the ASA letter at 36-40.

Third, one premise of Table 1 is that there is no rational basis for distinguishing between the rows with respect to the strength of the forces causing DG's pass and fail rates to differ from AG's pass and fails rates and thus any measure that does distinguish between the rows in that regard is an unsound measure. Yet researchers associated with COSSA and its members commonly employ these measures to quantify demographic differences. So far as the published records reveal, such researchers rely on these measures overwhelmingly, if not universally, not because they have considered the implications of the fact that the measure tends to change solely because the frequency of an outcome changes and have concluded both that the measure is sound and that there is no reason to discuss the effects of the frequency of the outcome. Rather, they do so without recognizing that the measures tend to change solely because the frequency of an outcome changes.

In light of these failures of understanding among researcher associated COSSA member entities, those entities would benefit greatly from having the issues raised in this letter brought to their attention. Thus, I suggest, COSSA has a responsibility to bring those issues to the attention of its members regardless of whether it chooses to undertake the particular tasks recommended in the letter.

Table 2 contains another simple illustration of the above-described pattern by which the two relative differences tend to be affected by the frequency of an outcome. The table is a portion of Table 1 of my guest editorial in the Spring 2006 issue of the ASA magazine *Chance* titled "[Can We Actually Measure Health Disparities?](#)" The table presents the percentages of whites and blacks living on incomes falling below and above 125%, 100%, and 75% of the poverty line in 2004, along with the same measures of difference presented in Table 1. The table also presents a figure termed EES (for estimated effect size), which is a measure of the difference between the circumstances of two groups reflected by their rates of experiencing some outcome that is theoretically unaffected by the frequency of the outcome that I will discuss further in Section D.

Table 2 shows the effects on various measures of differences in the circumstances of blacks and whites vis a vis poverty, or its avoidance, of (a) a reduction in poverty such as to enable everyone with incomes above 75% of the poverty line to escape poverty (as reflected by comparison of the bottom row with the middle row) and (b) an increase in poverty such as to pull into poverty everyone with incomes below 125% of the poverty line (as reflected by comparison

of the top row with the middle row). The table thus shows the way that general reductions in poverty will tend to increase relative differences in poverty rates while reducing relative differences in rates of avoiding poverty. That is, just as shown in Table 1, the relative difference in the decreasing outcome (poverty) increases, while the relative difference in the increasing outcome (avoidance of poverty) decreases. The table similarly shows that general increases in poverty will tend to have the opposite effect on each relative difference. The EES reflects the fact that all that has occurred in either situation is a change in the frequency of the outcome (though, if it did so perfectly, the EES value would not change at all). Again, I will defer discussion of absolute differences and odds ratios until later.

**Table 2. Rates at which white and blacks fall above and below 125%, 100%, and 75% of the poverty line (2004), with measures of differences between rates**

Perc of Poverty Line	White Perc Below	Black Perc Below	White Perc Above	Black Perc Above	B/W Below Ratio	W/B Above Ratio	Abs Df (Perc Pnts)	Odds Ratio	EES
125%	14.9%	31.0%	85.1%	69.0%	2.08	1.23	16	2.56	.54
100%	10.8%	24.7%	89.2%	75.3%	2.29	1.18	14	2.70	.55
75%	7.2%	17.8%	92.8%	82.2%	2.48	1.13	11	2.82	.54

In "[Race and Mortality Revisited](#)," *Society* (July/Aug. 2014) (at 329-330, 343), with reference to a version of the above table, I argue that it is not possible to justify the expenditure of any resources to examine changes in demographic difference in poverty rates (which is the same thing as examining the comparative changes in the poverty rates of different groups) without consideration of the patterns reflected in the table. Yet, to my knowledge, no study of changing demographic differences in poverty by researchers associated with COSSA members or elsewhere has ever considered the implications of these patterns. Thus, as discussed with regard to Table 1, COSSA members would benefit greatly from being made aware of the patterns in Table 2 and the implications of those patterns.

Table 3 presents a third simple illustration, while providing an additional perspective. Tables 1 and 2 involved situation where the only factor causing changes in measures from row to row was a change in the frequencies of the adverse and favorable outcomes. By contrast, Table 3 involves a situation where the changes in the measures from row to row are functions of both (a) changes in frequency of the outcome and (a) any changes in the strength of the forces causing the outcome rates to differ between the points in time for which the rates are presented. This illustration is relevant to the second of the subsequent developments discussed in Section C.

The table is an excerpted and slightly modified portion of Table 1(at 21) of my July 1, 2015 letter to the [Agency for Healthcare Research and Quality](#), which was a reproduction of Table 5 of my [presentation](#) at the Federal Committee on Statistical Methodology 2013 Research Conference. The table is based on data underlying the 2012 *National Healthcare Disparities Report* pertaining to situations that the report highlighted in its Table H.2 (at 14) as among the fastest decreasing healthcare disparities.

Table 3 presents rates at which non-Hispanic white and Hispanic hospital patients over age 65 with pneumonia received pneumococcal screening or vaccination in the years 2005 and 2010 (leaving the rates of failure to receive screening or vaccination to be inferred). The table also presents each of the measures presented in Table 2.

**Table 3. Rates at which whites and Hispanics over 65 with pneumonia received pneumococcal screening or vaccination in 2005 and 2010 (from 2012 NHDR), with measures of difference**

Year	White	Hisp	W/H Receipt Ratio	H/W Non-Receipt Ratio	Abs Df Perc Pts	Odd Ratio	EES
2005	63.90%	44.70%	1.43	1.53	19.3	2.19	0.49
2010	94.50%	88.30%	1.07	2.13	6.2	2.28	0.40

The principal point of the illustration in the table for purposes of this section of the letter is that the relative difference for the decreasing outcome (non-receipt of appropriate care) increased substantially while the relative difference for the increasing outcome (receipt of appropriate care) decreased substantially. Moreover, the pattern occurred even though, as shown in the EES column, to the extent that the forces causing the outcome rates to differ can be effectively measured, it decreased. That decrease caused the increase in the relative difference for the adverse outcome to be somewhat smaller, and the decrease in the relative difference for the favorable outcome to be somewhat larger than, than either change would have been had there solely occurred a change in the frequency of the outcome.

In addition to illustrating the pattern that is the focus of this section, the table is pertinent to matters discussed in subsequent sections, including the second recent development discussed in Section C, in the following respects. The National Center for Health Statistics, which at the time of the 2012 *National Healthcare Disparities Report* was measuring healthcare disparities in terms of relative differences in adverse outcomes, would have regarded the pattern in the table to reflect a substantial increase in disparity (which NCHS would have quantified as a 60 percentage point increase in the relative difference). As discussed in the above-mentioned letter to AHRQ, until at least 2010, AHRQ would have regarded the matter in the same way as NCHS. But commencing with the 2010 report (and continuing at least until the 2012 report), even though AHRQ intended to measure disparities in accord with NCHS guidance, the *National Healthcare Disparities Report* measured changes in disparities based on the comparative size of absolute (percentage point) changes in each group's rate. And it was because the Hispanic rate increased by 43.7 percentage points while the white rate increased by only 30.6 percentage points that the NHDR regarded the situation to warrant mention as one of the fastest decreasing disparities. AHRQ would nevertheless have continued to quantify the size of the disparity in terms of the relative difference in the adverse outcome, thus finding a much larger disparity in 2010 than 2005, notwithstanding the substantial decrease in disparity highlighted in the 2010 NHDR. As I discuss in Section C, under the NCHS approach to measuring healthcare disparities for Healthy People 2020, NCHS would regard the disparity to have decreased substantially, rather than increased substantially.



The pattern illustrated in Tables 1 through 3 whereby the rarer an outcome the greater tend to be relative differences in experiencing it and the smaller tend to be relative differences in avoiding it can be found in virtually any data showing points on a continuum of factors associated with likelihood of experiencing an outcome (and its opposite) as well as in countless situations of differing overall frequencies for an outcome, whether at different points in time, in setting differentiated other than temporally, or with respect to different types of outcomes. Examples of the pattern with respect to issues examined in the law and the social and medical sciences include: (a) reductions in mortality, morbidity, and poverty tend to be accompanied by increasing relative differences in mortality, morbidity, and poverty but decreasing relative differences in rates of avoiding those outcomes (as illustrated with respect to poverty in Table 2 *supra*); (b) improvements in healthcare tend to be accompanied by increasing relative differences in rates of failure to receive appropriate care but decreasing relative differences in rates of receipt of appropriate care (as illustrated in Table 3 *supra*); (c) relaxing mortgage lending and school discipline standards tends to increase relative differences in adverse borrower and discipline outcomes but reduce relative differences in rates of avoiding those outcomes; (d) relative differences in survival tend to increase as survival declines with age while relative difference in mortality tend to decrease as mortality increases with age; (e) relative differences in adverse outcomes tend to be larger, while relative differences in the corresponding favorable outcomes tend to be smaller, within populations/subpopulations and geographic areas where the adverse outcomes are less common (*e.g.*, the young, whites, persons with high income or high education, British civil servants,<sup>7</sup> the countries of Norway and Sweden, and the states of Minnesota and Massachusetts), than within populations/subpopulations and geographic areas where the adverse outcomes are more common.

A corollary to the above-described pattern of relative differences is a pattern whereby a factor that affects the likelihood of experiencing an outcome rate will tend to cause larger proportionate changes for the group with the lower baseline rates for the outcome while causing larger proportionate changes in the opposite outcome rates for the other group (which is the group with the lower baseline rate for that outcome). In the situation reflected in Table 1, lowering the cutoff caused the failure rate to decrease by 75% for AG but only 65% for DG, while causing the pass rate to increase by 38% for DG but only 19% for AG. Thus, for example, having high income or education will tend to cause larger proportionate reductions in mortality and self-rated health of less-than-good for advantaged groups while causing larger proportionate

---

<sup>7</sup> Readers unfamiliar with the Whitehall Studies or their findings may be puzzled by mention of British civil servants. Thus, I note that the Whitehall Studies of health disparities among British civil servants play an important role in the history of health disparities research, and interpretations of implications of the studies' findings that relative socioeconomic/occupational differences in adverse outcomes are greater among British civil servants than in the United Kingdom population at large exemplify the impossibility of soundly interpreting data on demographic differences without understanding the patterns described in this letter. See my "[Race and Mortality Revisited](#)," *Society* (July/Aug. 2014) (at 340) and 2006 British Society for Populations Studies paper "[The Misinterpretation of Health Inequalities in the United Kingdom](#)." See also the [Whitehall Studies](#) subpage of the [Measuring Health Disparities](#) page of [jpscanlan.com](#).

increases in survival and self-rated health of good-or-better for disadvantaged groups; having a chronic condition will tend to cause larger proportionate increases in rates of self-rated health of less-than-good for advantaged groups while causing larger proportionate decreases in self-rated health of good-or-better for disadvantaged groups; having fair or poor health (or aging or experiencing chronic conditions) will tend to cause a larger proportionate increases in mortality for advantaged groups while causing larger proportionate decreases in survival for disadvantaged groups.

A second corollary to the above-described pattern of relative differences is a pattern whereby reducing the frequency of an outcome tends to increase the proportions groups most susceptible to the outcome make up of (a) persons experiencing the outcome and (b) persons avoiding the outcome. According to the illustration in Table 1, assuming DG makes up 50% of test takers, lowering the cutoff would cause the proportion DG makes up of persons failing the test to increase from 65% to 72% and the proportion DG makes up of persons passing the test to increase from 44% to 48%. See Table 1 (at 11) of the ASA letter. Thus, for example, reductions in poverty tend to cause disadvantaged groups like racial minorities and female-headed families to make up higher proportions of the poor, and higher proportions of the non-poor, than they previously did. Increases in poverty will tend to have the opposite effect, that is, to cause disadvantaged groups to make up smaller proportions of the poor, and smaller proportions of the non-poor, than they previously did.

Appraisals of differences in the circumstances of advantaged and disadvantaged groups reflected by their outcome rates in terms of absolute (percentage point) differences between rates and differences measured by odds ratios are unaffected by whether one examines the favorable outcome or the corresponding adverse outcome. But in order for a measure to effectively quantify the strength of the forces causing outcome rates to differ, it must remain unchanged as there occurs a general change in the frequency of an outcome akin to that effected by the lowering of a test cutoff. And, like the two relative differences, absolute differences and odds ratios also tend to be affected by the frequency of an outcome, though in a more complicated way than the two relative differences. Because of that greater complexity, I defer discussion of these measures to Section D.

Failure to understand patterns described or referenced above has undermined virtually all efforts by social and medical scientists (and courts and administrative agencies) (a) to determine whether differences in the circumstances of advantaged and disadvantaged groups are increasing or decreasing over time, (b) to draw inferences about underlying processes based on perceptions about the comparative size of demographic differences or the comparative size of effects of a factor on different groups, or (c) even to determine whether a particular difference should be deemed large or small. In fact, with the exception of NCHS, none of government entities that COSSA has lately supported in testimony to Congress has shown any understanding of these patterns, and even NCHS has failed to understand the implications of the pattern it has recognized (a matter given further attention in Sections C and D). Thus, none of these entities has yet produced sound analyses of demographic differences in outcome rates or effectively overseen government-funded research regarding such differences.

COSSA should therefore form a committee to examine the patterns by which measures commonly employed in analyses of demographic differences tend to be affected by the frequency of an outcome, determine the extent to which standard demographic analyses are undermined by failure to understand these patterns, and explore methods for appraising the differences in the circumstances of advantaged and disadvantaged groups reflected by their outcome rates that are not affected by the frequency of an outcome.

## **2. This Letter's Narrower Recommendation**

Many federal civil rights enforcement policies are based on the belief that reducing the frequency of adverse outcomes will tend to reduce relative differences between rates at which advantaged and disadvantaged groups experience those outcomes and reduce the proportions disadvantaged groups make up of persons experiencing those outcomes. For example, for more than twenty years the federal government has encouraged lenders to relax lending standards and otherwise reduce the frequency of adverse borrower outcomes in order to reduce relative racial/ethnic differences in rates of experiencing those outcomes. For at least several years, the federal government has encouraged public schools to relax discipline standards and otherwise reduce the frequency of adverse school discipline outcomes in order to reduce relative racial and other demographic differences in rates of experiencing those outcomes and the proportions disadvantaged groups make up of persons experiencing the outcomes. In March 2015, in a report on the police and court practices of the City of Ferguson, Missouri, the Department of Justice (DOJ) attributed the high black representation among persons experiencing adverse interactions with the City's police and courts to over policing and unduly harsh court procedures. On February 10, 2016, DOJ brought a civil rights suit against the City of Ferguson based on the same theory.

As explained above, however, reducing the frequency of an outcome tends to increase, not decrease, relative differences between rates of experiencing the outcome and the proportions groups most susceptible to the outcome make up of persons experiencing the outcome. Unaware of this pattern, the government continues to monitor the fairness of practices on the basis of relative differences in adverse outcomes and the proportions disadvantaged groups make up of persons experiencing the outcomes. Thus, by complying with government encouragements to reduce the frequency of adverse outcomes entities covered by federal civil rights laws tend to increase the chances that the government will accuse them of discrimination.

For a COSSA committee to examine the problems with standard analyses of demographic differences arising from the failure to understand the patterns discussed in this letter, and to provide a sound foundation for analyses of demographic differences in outcome rates, may take some time. But even without forming a special committee, COSSA leadership should recognize that that the government is mistaken in its beliefs regarding the consequences of reducing the frequency of adverse outcomes. COSSA leadership should also recognize the urgency of making the government understand that actions it is currently taking and encouraging others to take are based on beliefs regarding the effects of those actions that are the opposite of reality.

I note, however, that larger undertaking recommended above has a similar urgency. As suggested in the PAA/APC letter, numerous presentations given at the 2016 Population Association of America Conference are based on studies (including ongoing studies) that are undermined by failure to recognize patterns by which the measures employed tend to be affected by the frequency of an outcome. Without action from an authoritative entity like, COSSA, ASA, or PAA/APC, many presentations at the 2017 Population Association of America Conference are likely to involve studies that are similarly undermined. But similar issues apply to activities of all COSSA members, including upcoming conferences, ongoing research, doctoral dissertations, and classroom instruction. Thus, unless COSSA concludes that one can effectively analyze demographic differences without consideration of the extent to which observed patterns are functions of the frequency of the outcome – and that analyses of demographic differences cannot even be improved by an understanding of the patterns by which measures tend to be affected by frequency of an outcome – it should take immediate action to alert its members to the potential implications of the matters addressed in this letter.

## **B. Materials Addressing the Pertinent Issues in Detail**

The statistical patterns underlying the recommendations in letter are addressed at much greater length in the above-mentioned letter to ASA. Other recent extended treatments of the issues may be found in my “[Race and Mortality Revisited](#),” *Society* (July/Aug. 2014); “[The Perverse Enforcement of Fair Lending Laws](#),” *Mortgage Banking* (May 2014); “[Measuring Health and Healthcare Disparities](#),” Federal Committee on Statistical Methodology 2013 Research Conference (March 2014) (FCSM paper); “[The Mismeasure of Discrimination](#),” Faculty Workshop, University of Kansas School of Law (Sept. 2013) (Kansas Law paper); *amicus curiae* [brief](#) in *Texas Department of Housing and Community Affairs et al. v. The Inclusive Communities Project, Inc.*, Sup. Ct. No. 13-1371 (Nov. 2014) (TDHCA brief).

My forthcoming commentary titled “The Mismeasure of Health Disparities,” *Journal of Public Health Management and Practice* (July/Aug. 2016), addresses certain recent developments regarding health and healthcare disparities research, some of which are also discussed in Sections C and D *infra*. The extended discussions in “Race and Mortality Revisited” and the FCSM paper of the NCHS recommendation that all healthcare disparities be measured in terms of relative differences in adverse outcomes should be read with recognition that, as discussed in Sections C and D *infra* and the forthcoming *Journal of Public Health Management and Practice* commentary, NCHS has recently reversed that recommendation.

The above items, and most other materials I have created since December 2007, include descriptions of a method for quantifying differences in the circumstances of two groups reflected by their rates of experiencing some outcome or its opposite (which may also be characterized as the strength of the forces causing the outcome rates to differ) that is unaffected by the frequency of an outcome (also discussed in Section D *infra*).

Earlier extended treatments of the problems with standard measure of differences between outcome rates (with many examples of fundamentally unsound analyses in the law and

the social and medical sciences), but without suggesting a solution to those problems beyond understanding the implications of the effects of the frequency of an outcome with respect to various measures, may be found in "[Can We Actually Measure Health Disparities?](#)," *Chance* (Spring 2006); "[The Misinterpretation of Health Inequalities in the United Kingdom](#)," British Society for Populations Studies Conference 2006; "[Race and Mortality](#)," *Society* (Jan./Feb. 2000); "[Divining Difference](#)," *Chance* (Fall 1994); "[Comment on 'McLanahan, Sorensen, and Watson's 'Sex Differences in Poverty, 1950-1980''](#)," *Signs* (Winter 1991); "[The Perils of Provocative Statistics](#)," *Public Interest* (Winter 1991); and "[The 'Feminization of Poverty' is Misunderstood](#)," *Plain Dealer* (Nov. 11, 1987).

Extensive graphical and tabular illustrations of the pertinent patterns may be found in methods workshops given at arms of American universities between 2012 and 2015,<sup>8</sup> as well as in more than a score of presentations given at statistics, epidemiology, public health, and demography conferences in North America and Europe between 2001 and 2011 (available [here](#)). Treatments of the patterns with respect to particular outcomes, with tabular illustrations, may also be found on the pages and subpages of [jpscanlan.com](#) devoted to measurement issues.<sup>9</sup> Over 140 online comments (available [here](#)) address the problems in various studies in medical, epidemiological, and health policy journals that attempted to measure aspects of health and healthcare disparities – or to provide guidance on the interpretation of subgroup effects or the calculation of number-needed-to-treat – without consideration of the patterns by which measures tend to be affected by the frequency of an outcome.<sup>10</sup>

---

<sup>8</sup> See "[The Mismeasure of Health Disparities in Massachusetts and Less Affluent Places](#)," Department of Quantitative Health Sciences, University of Massachusetts Medical School (Nov. 18, 2015) "[The Mismeasure of Discrimination](#)," Center for Demographic and Social Analysis, University of California, Irvine (Jan. 20, 2015); "[The Mismeasure of Demographic Differences in Outcome Rates](#)," Public Sociology Association of George Mason University (Oct. 18, 2014); "[Rethinking the Measurement of Demographic Differences in Outcome Rates](#)," Maryland Population Research Center of the University of Maryland (Oct. 10, 2014); "[The Mismeasure of Association: The Unsoundness of the Rate Ratio and Other Measures That Are Affected by the Prevalence of an Outcome](#)," Minnesota Population Center and Division of Epidemiology and Community Health of the School of Public Health of the University of Minnesota (Sept. 5, 2014); "[The Mismeasure of Group Differences in the Law and the Social and Medical Sciences](#)," Institute for Quantitative Social Science at Harvard University (Oct. 17, 2012); "[The Mismeasure of Group Differences in the Law and the Social and Medical Sciences](#)," Department of Mathematics and Statistics of American University (Sept. 25, 2012).

<sup>9</sup> The principal measurement pages are: [Measuring Health Disparities](#), [Scanlan's Rule](#), [Mortality and Survival](#), [Statistical Reasoning](#), [Immunization Disparities](#), [Immunization Disparities](#), [Educational Disparities](#), [Disparate Impact](#), [Discipline Disparities](#), [Lending Disparities](#), [Employment Discrimination](#), [Feminization of Poverty](#). The pages have close to 100 subpages.

<sup>10</sup> With respect to interpretations of subgroup effects, there also exist issues regarding the fact that, irrespective of the issues I raise about the ways measures tend to be affected by the frequency of an outcome, it is illogical to expect that in the usual course a factor affecting an outcome rate will cause equal proportionate changes in different baseline rates, given that it is not possible for a factor to cause equal proportionate changes in different rates of experiencing an outcome while causing equal proportionate changes in the opposite outcome rates (or that it is similarly illogical to regard the rate ratio as a sound measure of association, given that anytime the rate ratio is the same for two different pairs of outcome rates the rate ratio will necessarily will be different for the corresponding two pairs of rates for the opposite outcome). See "Race and Mortality Revisited" at 339-41 and the ASA letter at

Recent, fairly succinct treatments of the issues in the context of the government's encouraging lenders and public schools to relax lending and discipline standards under the mistaken belief that doing so will tend to reduce relative racial differences in adverse borrower and discipline outcomes, as well as the DOJ actions regarding Ferguson, Missouri, may found in my "[Things DoJ doesn't know about racial disparities in Ferguson](#)," *The Hill* (Feb. 22, 2016); "[Things government doesn't know about racial disparities](#)," *The Hill* (Jan. 28, 2014); "[The Paradox of Lowering Standards](#)," *Baltimore Sun* (Aug. 5, 2013); and "[Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies](#)," *Amstat News* (Dec. 2012). Older, somewhat longer treatments of the misunderstanding of the effects of relaxing standards on measures of disproportionality, which involve a variety of settings and illustrate both how longstanding and how universal is that misunderstanding, may be found in my "[Mired in Numbers](#)," *Legal Times* (Oct. 12, 1996); "[When Statistics Lie](#)," *Legal Times* (Jan. 1, 1996); "[Getting it Straight When Statistics Can Lie](#)," *Legal Times* (June 23, 1993); "[Bias Data Can Make the Good Look Bad](#)," *American Banker* (Apr. 27, 1992); and "[An Issue of Numbers](#)," *National Law Journal* (Mar. 5, 1990).

Like this letter and the letters to ASA and PAA/APC, available online are letters of varying length advising governmental and nongovernmental entities (including COSSA members) of the ways analyses of demographic or other differences in outcome rates that the entities conduct, fund, or provide guidance on – or that in some manner pertain to the entities' activities – are undermined by the failure to recognize patterns by which measures of differences between outcome rates tend to be affected by the frequency of an outcome.<sup>11</sup> The March 8, 2016

---

12-13. See also the [Illogical Premises](#), [Illogical Premises II](#), [Subgroup Effects](#), [Subgroup Effects – Nonclinical](#), and [Inevitability of Interaction](#) subpages of the [Scanlan's Rule](#) page of [jpscanlan.com](#) and the [Comment on Hingorani BMJ 2013](#). I do not address these issues at length here, though COSSA should consider them carefully.

<sup>11</sup> See letters to [Council of Economic Advisers](#) (Mar. 16, 2016), [City of Madison, Wisconsin](#) (Mar. 12, 2016), [Stanford Center on Poverty and Inequality](#) (Mar. 8, 2016), [City of Boulder, Colorado](#) (Mar. 5, 2016), [Houston Independent School District](#) (Jan. 5, 2016), [Boston Lawyers' Committee for Civil Rights and Economic Justice](#) (Nov. 12, 2015), [House Judiciary Committee](#) (Oct. 19, 2015), [Chief Data Scientist of White House OSTP](#) (Sept. 8, 2015), [McKinney, Texas Independent School District](#) (Aug. 31, 2015), [Department of Health and Human Services and Department of Education](#) (Aug. 24, 2015), [Agency for Healthcare Research and Quality](#) (July 1, 2015), [City of Minneapolis, Minnesota](#) (June 8, 2015), [Texas Appleseed](#) (Apr. 7, 2015), [Senate Committee on Health, Education, Labor and Pensions](#) (Mar. 20, 2015), [United States Department of Justice and City of Ferguson, Missouri](#) (Mar. 9, 2015), [Vermont Senate Committee on Education](#) (Feb. 26, 2015), [Portland, Oregon Board of Education](#) (Feb. 25, 2015), [Wisconsin Council on Families and Children's Race to Equity Project](#) (Dec. 23, 2014), [Financial Markets and Community Investment Program, Government Accountability Office](#) (Sept. 9, 2014), [Education Law Center](#) (Aug. 14, 2014), [IDEA Data Center](#) (Aug. 11, 2014), [Institute of Medicine II](#) (May 28, 2014), [Annie E. Casey Foundation](#) (May 13, 2014), [Education Trust](#) (April 30, 2014), [Investigations and Oversight Subcommittee of House Finance Committee](#) (Dec. 4, 2013), [Mailman School of Public Health of Columbia University](#) (May 24, 2013), [Senate Committee on Health, Education, Labor and Pensions](#) (Apr. 1, 2013), [Federal Reserve Board](#) (March 4, 2013), [Harvard University et al.](#) (Oct. 26, 2012), [Harvard University](#) (Oct. 9, 2012), [United States Department of Justice](#) (Apr. 23, 2012), [United States Department of Education](#) (Apr. 18, 2012), [The Commonwealth Fund](#) (June 1, 2010), [Institute of Medicine](#) (June 1, 2010), [National Quality Forum](#) (Oct. 22, 2009), [Robert Wood Johnson Foundation](#) (Apr. 8, 2009)

letter to the [Stanford Center on Poverty and Inequality](#) (at 3-4) (an arm of a COSSA affiliate member) discusses recent work by others that independently reaches conclusions about the measurement of group differences in outcome rates that, in my view, are similar to those in the materials discussed above with respect to both the problem and potential solutions. Pages 5-6 discuss what may be a recurring anomaly absent action of the type suggested in this letter and one that is related to the second recent development discussed in Section C – specifically, that government funded research will rely on NCHS recommendations to measure healthcare disparities in terms of relative differences in adverse outcomes (and do so in circumstances where relative differences in favorable outcomes will yield substantially different conclusions), while unaware that NCHS had already reversed that recommendation.

Letters to other COSSA affiliate members or their arms regarding measurement issues at the recipient entities include those to [Harvard University et al. \(including Harvard Medical School and Massachusetts General Hospital\)](#). (Oct. 26, 2012), [Harvard University](#) (Oct. 9, 2012), and [Mailman School of Public Health of Columbia University](#) (May 24, 2013). The exchanges with arms of Harvard University are discussed in "Race and Mortality Revisited" at 343-344.

Warranting your particular attention, given the nature of COSSA and COSSA member support of and interactions with the recipients, are the July 1, 2015 letter to the [Agency for Healthcare Research and Quality](#) (AHRQ) and the August 24, 2015 letter to the [Department of Health and Human Services and Department of Education](#) (HHS & DOE). The AHRQ letter explains that the health and healthcare disparities research the agency conducts or funds has been substantially undermined by the failure, on the parts of both AHRQ and recipients of AHRQ funding, to recognize the ways measures employed in such research tend to be affected by the frequency of an outcome. The letter also explains that, apparently as a result of the agency's failure to understand certain measurement issues, AHRQ has failed to measure health and healthcare disparities in the manner it believed it was doing (as discussed somewhat above with regard to Table 3). The failure resulted in an anomaly whereby the 2010 *National Healthcare Disparities Report* highlighted as some of the largest reductions in healthcare disparities over a certain period situations where AHRQ would nevertheless find the disparities to be substantially larger at the end of the period than they were at the beginning of the period. The HHS/DOE letter discusses the fact that, even though a decade earlier NCHS recognized that reducing the frequency of an outcome will tend to increase relative differences between rates of experiencing the outcome, in December 2014, HHS joined DOE in a policy statement reflecting the view that reducing the frequency of an outcome will tend to reduce those differences. The letter also touches on the matter I discuss in the next section regarding recent NCHS actions effectively repudiating the *National Healthcare Disparities Reports* and other healthcare research that relied on earlier NCHS recommendations.

### C. Recent Developments

This letter suggests that the reader should rely on the longer treatments of the subject in places like the October 2015 ASA letter and the July/August 2014 "Race and Mortality Revisited." For that reason, I set out below two recent developments that bear on the urgency of the recommended COSSA actions.

The first recent development involves a relatively simple matter that is touched upon in Sections A and B. On February 10, 2016, the DOJ filed a suit against Ferguson, Missouri, that, like the underlying March 2015 DOJ report on police and court practices in Ferguson, is based on a belief about the relationship between the frequency of an outcome and measures of disproportionality that is the opposite of reality. The suit also raises issues about the perception of the magnitude of disparities that DOJ has so far failed to understand. See my "[Things DOJ doesn't know about racial disparities in Ferguson](#)," *The Hill* (Feb. 22, 2016). Meanwhile, all across the country, jurisdictions are examining racial disproportionality in arrests with the same mistaken understanding of the relationship between the frequency of an outcome and measures of disproportionality in experiencing the outcome that underlies the DOJ's actions against Ferguson and that has underlain a range of federal law enforcement practices for more than two decades. See my letters to [City of Madison, Wisconsin](#) (Mar. 12, 2016) and [City of Boulder, Colorado](#) (Mar. 5, 2016). There exist similar pressing issues with regard to racial disproportionality in school discipline. See, e.g., the letters to [Houston Independent School District](#) (Jan. 5, 2016) and [Boston Lawyers' Committee for Civil Rights and Economic Justice](#) (Nov. 12, 2015) with respect to the fact that neither the entities whose practices are challenged or appraised, nor the entities purporting to provide sound analyses of those practices, understand the relationship between the frequency of an outcome and measures of disproportionality.<sup>12</sup>

The second recent development involves actions of NCHS and the Centers for Disease Control and Prevention (CDC) regarding health and healthcare disparities research.<sup>13</sup> Most of my extended discussions of health and healthcare disparities research issues since 2006 address the fact that in 2004-05 NCHS statisticians recognized that determinations of whether health and healthcare disparities were increasing or decreasing would commonly turn on whether one examined relative differences in favorable outcomes or relative differences in the corresponding

---

<sup>12</sup> See the following web pages regarding situations where recent reductions in discipline rates were accompanied by increased relative racial/ethnic differences in discipline rates in the jurisdictions indicated in the title of the page: [California Disparities](#), [Colorado Disparities](#), [Connecticut Disparities](#), [Maryland Disparities](#), [Minnesota Disparities](#), [Beaverton, OR Disparities](#), [Denver Disparities](#), [Henrico County, VA Disparities](#), [Los Angeles SWPBS](#), [Minneapolis Disparities](#), [Montgomery County, MD Disparities](#), [Portland, OR Disparities](#), [St. Paul Disparities](#). The pages principally address the patterns of increasing relative racial differences in jurisdictions. But, like many of the letters listed in note 7, many of the pages also reflect the failure of officials of the various jurisdictions, and of entities examining discipline disparities in those jurisdictions, to understand that relaxing standards tend to increase relative differences in adverse outcomes.

<sup>13</sup> These developments are the subject of my forthcoming "The Mismeasure of Health Disparities," *Journal of Public Health Management and Practice* (2016).



favorable outcomes. That is, they recognized that reductions in mortality tend to be associated with increased relative differences in mortality but reduced relative differences in survival and that improvements in healthcare tend to be associated with increased relative differences in failure to receive appropriate care but reduced relative differences in receipt of appropriate care. See especially "Race and Mortality Revisited" at 331-335.

The forces that cause adverse outcome rates of advantaged and disadvantaged groups to differ are exactly the same forces that cause the corresponding favorable outcome rates to differ. Thus, NCHS should have regarded the fact that the two relative differences commonly (or ever) yield opposite conclusions regarding whether those forces are increasing or decreasing as calling into question the value of either relative difference for quantifying the differences in the circumstances of advantaged and disadvantaged groups. Instead, however, the agency simply determined that in appraising progress on the disparities reduction goals in Health People 2010, both health and healthcare disparities would be measured in terms of relative differences in adverse outcomes (that is, in the case of healthcare, non-receipt rather than receipt of appropriate care). See references in note 4 *supra*.

Prior to the NCHS recommendation disparities in health outcomes usually (or almost invariably) were measured in terms of relative differences in adverse outcomes (and thus improvements in health tended to be associated with increasing disparities),<sup>14</sup> while disparities in healthcare outcomes usually were measured in terms of relative differences in favorable outcomes (and thus improvements in healthcare tended to be associated with reduced disparities). The NCHS recommendation to measure healthcare disparities in terms of relative differences in adverse outcomes, which was adopted by AHRQ for the *National Healthcare Disparities Reports*, reversed the situation as to healthcare disparities. But NCHS left entirely unaddressed the fundamental unsoundness of the relative difference as a measure of either health or healthcare disparities.

The NCHS belief that it could arbitrarily choose which relative difference to rely on reflected a fundamental misunderstanding of the purpose of studying health disparities. The unsoundness of its position is especially evident when observers attempt to draw inferences about processes on the basis of the comparative size of the relative differences in experiencing some outcome without consideration that the comparative size of relative differences for the opposite outcome would commonly support a very different inference. See "Race and Mortality Revisited" at 339-341 and the March 8, 2016 letter to the [Stanford Center on Poverty and Inequality](#) at 5-6.

---

<sup>14</sup> See the [Mortality and Survival](#) page of [jpscanlan.com](#) regarding the fact that what titles of articles describe as analyses of demographic disparities in cancer survival are usually analyses of relative differences in cancer mortality, without the authors' recognizing such things as (a) that general improvements in survival tend to increase relative differences in mortality while reducing relative differences in survival or (b) that more survivable cancers tend to show larger relative differences in mortality, but smaller relative differences in survival, than less survivable cancers. On occasion, however, what are described as analyses of cancer survival disparities are analyses of relative differences in survival, that is, are what they are represented to be. See also the FCSM paper at 21-22.

NCHS usually made it clear enough that it was examining healthcare disparities in terms of relative differences in adverse outcomes even when the subject was cast in terms of the favorable outcome. But subsequent to 2004-05 the agency did little to inform the public or researchers (or other arms of HHS) of the substantial implications of measuring healthcare disparities in terms of relative differences in adverse, rather than favorable, outcomes.<sup>15</sup> Thus, so far as the published record reveals (including the *National Healthcare Disparities Reports* of AHRQ and the 2011 and 2013 *Health Disparities and Inequalities Reports* of the CDC) no arm of HHS other than NCHS has ever indicated an awareness that is even possible for relative differences in favorable health and healthcare outcomes, and relative differences in the corresponding adverse health and healthcare outcomes, to change in opposite directions as the frequency of an outcome changes, much less that NCHS statisticians found that this tends to occur systematically. See also the discussion in Section B *supra* regarding the December 2014 joint HHS/DOE policy statement reflecting an understanding of the effects of reducing the frequency of an outcome on relative differences in rates of experiencing the outcome that is the opposite of the conclusion NCHS reached more than a decade ago.

As discussed in the HHS/DOE letter (at 24) some time prior to August 2015, NCHS published online a “[Health Disparities Tool: A User’s Guide](#)” (and associated materials) indicating that for purposes of appraising progress in Healthy People 2020 health and healthcare disparities reduction goals, healthcare disparities will be measured in terms of relative differences in favorable outcomes. In February 2016, NCHS confirmed this approach in a Statistical Note titled “[Measuring Progress Toward Target Attainment and the Elimination of Health Disparities in Healthy People 2020](#).”<sup>16</sup> In indicating that measurement of relative differences in healthcare outcomes will now be calculated on the basis of the favorable outcome, the Note explains that the approach differs from the methodology of Health People 2010. But few readers would understand the implications of the change. And fewer still would understand that the change of approach effectively repudiates more than a decade of healthcare disparities research, including that in the *National Healthcare Disparities Reports*, that, in reliance on NCHS guidance, measured healthcare disparities in terms of relative differences in adverse outcomes.

In the 2011 and 2013 *Health Disparities and Inequalities Reports* the CDC evidenced little awareness that different measures might yield different conclusions about directions of

---

<sup>15</sup> See (1) Keppel KG, Bilheimer L, Gurley G. Improving population health and reducing health care disparities. *Health Affairs*. 2007;26(5):1281–1292; (2) Keppel KG. Ten largest racial and ethnic health disparities in the United States based on Health People 2010 objectives. *Am J Epidemiology*. 2007;166(1):97-103. See page 32 of the [presentation](#) associated with the FCSM paper regarding the way the Healthy People 2010 Technical Appendix advises that favorable healthcare outcomes are re-expressed in terms of the adverse outcome for purposes of calculating the relative difference, without providing any indication of the implications of doing so save by references to the 2004-05 papers of NCHS statisticians.

<sup>16</sup> Talih M, Huang DT. Measuring progress toward target attainment and the elimination of health disparities in Healthy People 2020. Healthy People Statistical Notes, no 27. Hyattsville, MD: National Center for Health Statistics. 2016.

changes in disparities over time and, as noted, no awareness that relative differences in favorable outcomes and relative differences in the corresponding adverse outcomes could yield different conclusions. While CDC (apart from its arm NCHS) had previously not offered guidance on the measurement of health and healthcare disparities save as reflected in the approaches of its own analyses, CDC scientists/statisticians in conjunction NCHS statisticians, recently published a substantial guide on the measurement of health disparities.<sup>17</sup>

Nothing in the document (save, possibly, for a reference to discussion of “other issues” in the NCHS monograph listed as item (b) in note 4 of this letter) suggests any awareness of the NCHS recognition that the two relative differences would commonly change in opposite directions. The article gives significant attention to choosing between the absolute difference and a relative difference when the two yield opposite conclusions about such things as the direction of changes in disparities over time. But its discussion of such issue contains nothing to suggest that the unmentioned relative difference might offer a different perspective, even though (as discussed in "Race and Mortality Revisited" and the ASA letter and addressed somewhat in Section D below) anytime it is mentioned that a relative differences and the absolute difference yield opposite conclusions about the directions of changes in disparities over time, the unmentioned relative difference will necessarily have changed in the opposite direction of the mentioned relative differences and the same direction as the absolute difference. Most important, however, the document shows no awareness of the patterns whereby the measures it discusses tend to be affected by the frequency of an outcome or the implications of such patterns with respect to the utility of the measures for appraising differences in the circumstances of advantaged and disadvantaged groups.

In sum, recent developments heighten the urgency of both the broader and narrower recommendations in this letter and the letters to ASA and PAA/APC.

#### **D. Illustrations of Patterns by Which Standard Measures of Differences Between Outcome Rates Tend to be Affected by the Frequency of an Outcome**

There are four principal measures by which demographers and others appraise differences in rates at which advantaged and disadvantaged groups experience favorable or adverse outcomes: (1) relative (percentage) differences between rates of experiencing the outcome; (2) relative differences between rates of avoiding the outcome (*i.e.*, experiencing the opposite outcome); (3) absolute (percentage point) differences between outcome rates; and (4) odds ratios. None of these measures provides a sound basis for quantifying differences in the circumstances of advantaged and disadvantaged groups reflected by their outcome rates (or, otherwise put, the strength of the forces causing the groups' outcome rates to differ) because each measure tends to be systematically affected by the frequency of an outcome.

---

<sup>17</sup> Penman-Aguilar A, Talih M, Huang D, et al. Measurement of health disparities, health inequities, and social determinants of health to support the advancement of health equity. *J Public Health Manag Pract.* 2016;22(1 Supp):S33–S42.

The pattern by which the two relative differences tend to be affected by the frequency of an outcome – specifically, the rarer an outcome the greater tend to be relative differences in experiencing it and the smaller tend to be relative differences in avoiding it – was discussed in Section A.

The fact that absolute differences between rates and differences measured by odds ratios tend to be affected by the frequency of an outcome, though in more complicated than the two relative differences was briefly mentioned in Section A. The patterns by which those measures tend to be affected by the frequency of an outcome are addressed further below.

Roughly, as uncommon outcomes (below 50% for both groups) become somewhat more common, absolute differences between rates tend to increase; as common outcomes (above 50% for both groups) become even more common, absolute differences tend to decrease. The frequency-related directions of change of the absolute difference are harder to predict when the outcome is neither common nor uncommon or changes from being uncommon to common during a period examined.

The absolute difference and both relative differences may all change in the same direction as the frequency of an outcome changes (in which case the difference measured by the odds ratio will also change in the same direction as the other measures). In such circumstances, one may infer that there has been a change in the strength of the forces causing the outcome rates to differ even if the standard measures do not effectively quantify that change.<sup>18</sup>

But in the common situation where all measures do not change in the same direction as the frequency of an outcome changes, the absolute difference will tend to change in the same direction as the smaller relative difference. Since observers who rely on relative differences to appraise differences in the circumstances of two groups reflected by their differing outcome rates usually rely on the larger of the two relative differences, there exists a tendency for observers relying on the absolute difference to reach opposite conclusions about directions of changes in the size of demographic differences, or about the comparative size of the differences in different settings, from observers relying on relative differences.

Further, anytime an observer notes that the absolute difference has changed in a different direction from the relative difference the observer happens to be examining, the unmentioned relative difference will necessarily have changed in the opposite direction of the mentioned relative difference and the same direction as the absolute difference. Thus,

---

<sup>18</sup> For simplicity I cast the matter in terms of an outcome that is increasing in frequency and without reference to the opposite outcome. But it should be obvious that when the outcome is decreasing the absolute difference tends to change in the opposite direction of that described. Further, in the case of an uncommon outcome that is increasing while remaining uncommon (with corresponding increase in the absolute difference), the opposite outcome is necessarily a common outcome that is decreasing while remaining common (also with corresponding increase in the absolute difference).

observers maintaining that one must make a value judgment in choosing between the relative difference and the absolute difference with respect to the appraisal of a particular change in some demographic disparity over time have already made a choice (usually without thought) to rely on the relative difference that yields an opposite conclusion from the absolute difference rather than the relative difference that yields the same conclusion as the absolute difference.

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 and in the same direction as the larger relative difference.

Graphical depictions of the relationship of the two relative differences, the absolute difference, and the odds ratio according to the same specifications underlying Table 1 of this letter may be found in Figures 2 and 3 of the ASA letter (at 21-22). Table 4 below, which is based on the same specifications as Table 1 of this letter, presents favorable outcome rates for an advantaged and disadvantaged group (with corresponding adverse outcome rates implied) at four frequency levels benchmarked by the advantaged group's favorable outcome rate. The table also presents rate ratios for the favorable and adverse outcomes and the absolute differences and odds ratios (with the larger odds in the numerator). The parenthetical numbers indicate the ways one would rank the size of the disparity between the situation of the advantaged and disadvantaged groups according to each of the four measures. Those numbers show that rankings according to relative differences in favorable outcomes are the opposite of rankings according to relative differences in adverse outcomes. They also show that rankings according to absolute differences and odds ratios, while different from the rankings according to either relative difference, are the opposite of each other.

**Table 4. Favorable outcome rates of advantaged group (AG) and disadvantage group (DG) at four settings with different favorable outcome frequencies, with measures of difference**

Setting	AG Fav Rate	DG Fav Rate	AG/DG Fav Ratio	DG/AG Adv Ratio	Abs Df (Perc Pnts)	Odds Ratio
A	20.0%	9.0%	2.22 (1)	1.14 (4)	11.0 (4)	2.53 (1)
B	40.0%	22.6%	1.77 (2)	1.29 (3)	17.4 (2)	2.28 (3)
C	70.0%	51.0%	1.37 (3)	1.63 (2)	19.0 (1)	2.24 (4)
D	80.0%	63.4%	1.26 (4)	1.83 (1)	16.6 (3)	2.31 (2)

A version of this table appears as Table 5 (at 335) of "Race and Mortality Revisited," where, as in many other places, it is used to refute the notion that choice of a measure in analyses of demographic differences in health or healthcare outcomes involves a value judgment. In the main, I will rest here on the treatment of that issue in "Race and Mortality Revisited" (at 335-336) and other places. The table, however, usefully illustrates a number of other points pertinent to the subject of this letter.

To being with, the key point of the table is that there is no rational basis for asserting that the strength of the forces causing the outcome rates to differ varies among the rows and that any

measure that suggests that that they do vary is an unsound measure (a point, as discussed above, that applies as well to the less complicated Table 1). The table thus implies a means of appraising the strength of those forces reflected by any pair of AG and DG rates. Just as knowledge of the difference between the means of the underlying distributions enables one to derive the rate for DG associated with any favorable or adverse outcome rate of AG (and vice versa), knowledge of the properties of normal distributions enables one to derive from any pair of rates an estimate of the difference between the means of underlying normal distributions, in terms of percentage of a standard deviation. I illustrate that method in the tables discussed below.

Table 4 also illustrates how inferences based on the comparative size of some relative difference will tend to be countered by an inference based on the relative difference in the opposite outcome. For example, assuming that Rows A and B reflect the situation of job applicants with low qualifications while Rows C and D reflect the situation of job applicants with high qualifications, any inference based on the smaller relative differences in the favorable outcome among the latter than among the former would be no more valid than a contrasting inference based on the smaller relative difference in the adverse outcome among the former than among the latter. The same point applies to any appraisal of heterogeneity of effects involving groups with different baseline rates. That is, for example, in the case of the employment setting just posited, the table shows that improvements in qualifications tends to increase favorable outcome rates proportionately more for DG (the group with the lower baseline favorable outcomes rate) than AG, while reducing adverse outcomes rates proportionately more for AG than DG. As discussed with regard to the first corollary to the pattern whereby the two relative differences tend to be affected by the frequency of an outcome, in any situation where the two relative differences yield opposite conclusions about the comparative size of a disparity, a factor that affects the outcome rates is necessarily causing a larger proportionate change in one outcome rate for one group while causing a larger proportionate change in the opposite outcome rate for the other group. See "Race and Mortality Revisited" at 339-341. See also the [Interactions by Age](#) subpage of the [Scanlan's Rule](#) page of [jpscanlan.com](#).

Further with respect to the interpretation of data on employment outcomes, the table illustrates that in an employment tester context, choice of the level of qualifications to fabricate for tester job applicants, and choice of labor market in which to have them apply, will have a substantial effect on the various measures one might employ to appraise the results of the studies. And, in the prevailing environment where either relative difference is regarded as a legitimate measure, there may even be situations where favorable outcome rates reach a point where observers decide to examine relative differences in adverse outcomes rather than relative differences in favorable outcome rates, thus reversing the effects of further enhancement of fabricated qualifications. See "Race and Mortality Revisited" at 341, [University of Massachusetts Medical School Seminar](#) at 143-150. See also the [Four-Fifths Rule](#) subpage of the [Disparate Impact](#) page of [jpscanlan.com](#) (discussing that the Uniform Guidelines for Employee Selection Procedures recommend that users measure disparate impact in terms of the relative difference in the favorable outcome until the point where favorable outcome rates are so

high that the disadvantaged group's rate cannot be less than 80% of the advantaged group's rate, in which case users should rely on the relative difference in the adverse outcome).

Table 4 also illustrates that observers relying on absolute differences to appraise healthcare disparities regarding uncommon outcomes will find that (a) improvements in care tend to increase disparities and (b) disparities tend to be greater at higher-performing institutions than lower-performance institutions (as reflected by comparison of Row A with Row B). It similarly illustrates that observers relying on absolute differences to appraise disparities regarding common outcomes will find that (a) improvements in care tend to decrease disparities and (b) disparities tend to be smaller at higher-performing institutions than lower-performance institutions (as reflected by comparison of Row C with Row D). See "Race and Mortality Revisited" at 337-339 regarding the substantial implications of the failure to understand these patterns with respect to pay-for-performance issues.

Tables 5 to 7 illustrate various points made above, while including the measure described above (termed EES for estimated effect size) that is theoretically unaffected by the frequency of an outcome. Each is a version of a table from "[Race and Mortality Revisited](#)" or the 2013 [FCSM paper](#). I urge the reader to examine all of the tables in those items for the perspectives they offer regarding the problems in standard measures of differences between outcome rates and the plausibility of the measure I suggest or one similar to it.

Two things should be borne in mind in appraising the points I make with these tables as well as many of the points made above. First, while patterns in the table that follow are in most respects generally consistent with the frequency-related patterns described, one will of course find many situations where observed patterns do not accord with those patterns. Observed patterns are functions of (a) the described frequency-related patterns and (b) the strength of the forces causing the outcome rates to differ in the settings being compared. The strength of those forces may vary greatly from setting to setting and may change substantially over time (especially longer periods of time) particularly during periods of dramatic changes in the frequency of an outcome. Such factors may cause the frequency-related patterns not to be observed even though they are having a substantial effect. The crucial consideration, however, is that so long as measures tend to be affected by the frequency of an outcome in any manner, one cannot rely on the measures to appraise the strength of the forces causing rates to differ without consideration of the role of the frequency of the outcome.

Second, in "Race and Mortality Revisited"(at 337) and many other places, I describe the imperfections of the measure I suggest. The [Kansas Law paper](#) discusses some complexities of the matter that are particularly pertinent to the analyses of discrimination issues.<sup>19</sup> But the

---

<sup>19</sup> Section C (at 23-26) of the Kansas Law paper also treats the impossibility of appraising the size of a disparity based solely on information regarding the proportion a group makes up of persons potentially experiencing an outcome and the proportion it makes up of persons actually experiencing the outcome. While I do not otherwise discuss the matter here, it is an important issue. See also the [TDHCA brief](#) (at 32-27), the [IDEA Data Center Disproportionality Guide](#) subpage of the Discipline Disparities page, and [slides 97 to 108](#) of the [University of Maryland Workshop](#).

measure at least provides a rational basis for appraising the strength of the forces causing the outcome rates to differ, while the standard methods do not.<sup>20</sup> An important role of the committee that I suggests COSSA form will be to appraise the utility of that measure and to determine whether there may be a superior measure that is also unaffected by the frequency of an outcome but that better quantifies the strength of the forces causing the outcome rates to differ.

Table 5 presents results of a study I have used in many places to show how researcher unaware of the NCHS recommendation that healthcare disparities be measured in terms of relative differences in adverse outcomes would tend to reach opposite conclusions from NCHS.<sup>21</sup> Relying on relative differences in vaccination rates as a measure of disparity (and evidencing no awareness that NCHS would do otherwise), the authors found that a school-entry Hepatitis B vaccination requirement, which dramatically increased vaccination rates, dramatically reduced racial disparities for both fifth and ninth graders. NCHS, relying on relative differences in rates of failure to be vaccinated, would have found substantially increased disparities for both grades. CDC, which commonly measures vaccination disparities in terms of absolute differences between rates, would have found substantially increased disparities for fifth graders (where initial rates were quite low) and substantially decreased disparities for ninth graders (where initial rates were much higher).

The EES column indicates that, to the extent that the disparities can be effectively measured, there occurred notable reductions in both grades. That, it warrants note, is the type of change one would expect in the case of a formal requirement, which, if rigidly enforced, ought to eliminate any disparity.

**Table 5. White and black Hepatitis B vaccination rates in grades 5 and 9 before and after imposition of school-entry vaccination requirement, with measures of difference**

Grade	Year	Requirement Situation	White Vac Rt	Black Vac Rt	W/B Ratio Vac	B/W Ratio No Vac	Abs Df (Perc Pnts)	Odds Ratio	EES
5	1996	Pre	8%	3%	2.67	1.05	5	2.81	.47
5	1997	Post	46%	33%	1.39	1.24	13	1.73	.34
9	1996	Pre	46%	32%	1.44	1.26	14	1.81	.37
9	1997	Post	89%	84%	1.06	1.45	5	1.54	.24

---

<sup>20</sup> A qualification to this point is that the odds ratio (more usefully represented in terms of the standardized mean difference) could be a superior measure to the one I suggest depending on actual shapes of the underlying distributions. For middle ranges of frequency of an outcome, the standardized mean difference differs little from the EES.

<sup>21</sup> Morita JY, Ramirez E, Trick WE. Effect of school-entry vaccination requirements on racial and ethnic disparities in Hepatitis B immunization coverage among public high school students. *Pediatrics*. 2008;121:e547.



As mentioned above, under the healthcare disparities measurement approach of Healthy People 2020, NCHS would now agree with the authors that there occurred a substantial decrease in disparities in both grades.

Table 6 is based on a study that relied on NCHS guidance for measuring healthcare disparities in terms of relative differences in adverse outcomes.<sup>22</sup> The study's abstract highlighted as a very large increase in relative area-socioeconomic differences in mammography what was in fact a very large increase in relative area-socioeconomic differences in failure to receive mammography. As shown in the table, there was actually a very large decline in the relative difference in mammography rates. The EES indicates that, to the extent the disparity can be measured, it increased negligibly.<sup>23</sup>

**Table 6. Mammography rates of highest and lowest socioeconomic groups in 1987 and 2004, with measures of difference**

Year	High Mamm Rt	Low Mamm Rt	H/L Ratio Mamm	L/H Ratio No Mamm	Abs Df Perc Pnts	Odds Ratio	EES
1987	36.3%	17.2%	2.11	1.30	19	2.74	0.60
2004	77.4%	55.2%	1.40	1.98	22	2.78	0.62

Though the authors of the study relied on NCHS guidance, in this case NCHS would now reach an opposite conclusion from that of the authors.

Table 7 is based on data from a World Health Organization (WHO) [Handbook on Health Inequality Monitoring](#) (2013) that, while citing NCHS guidance that recommended measuring healthcare disparities in terms of relative differences in adverse outcomes (item (b) in note 4 of this letter), measured income disparities in skilled attendance at birth and other healthcare outcomes in terms of relative difference in favorable outcomes. The document did not reflect awareness that NCHS would analyze the matter differently or that it could make a difference whether one examined the relative difference in the favorable outcome or the relative difference in the corresponding adverse outcome. The table shows the starkly contrasting interpretations of the income-related disparities for skilled attendance at birth for Columbia and Bangladesh according to the WHO approach and the NCHS approach. These countries, which WHO found to have the smallest and largest disparities of the countries it studied, would have the largest and smallest disparities according to the NCHS approach. See FCSM paper at 18-19. The NCHS Healthy People 2020 approach would now agree with the WHO interpretation.

<sup>22</sup> Harper S, Lynch J, Meersman SC, et al. Trends in area-socioeconomic disparities in breast cancer screening, mortality, and survival among women ages 50 years and over (1987-2005). *Cancer Epidemiol Biomarkers Prev.* 2009;18(1):121-131.

<sup>23</sup> See the FCSM paper (at 16) regarding my inability to replicate exactly what the study described as a 161 percent increase in relative area-socioeconomic differences in mammography (meaning failure to receive mammography) and regarding the authors explanation for examining relative differences in rates of failure to receive mammography rather rates if receipt of mammography.

**Table 7. Rates of births attended by skilled measures for highest and lowest income quintiles in Columbia and Bangladesh, from WHO Handbook on Health Inequality Monitoring, with measures of difference**

Country	Highest Q Attend Rt	Lowest Q Attend Rt	H/L Ratio Attend	L/H Ratio No Attend	Abs Df Perc Pnts	Odds Ratio	EES
Columbia	99.4%	83.7%	1.2	27.2	15.7	32.3	1.34
Bangladesh	50.6%	4.9%	10.3	1.9	45.7	19.9	1.67

The FCSM paper and "Race and Mortality Revisited" present a number of examples like this, with reference to the then-current NCHS approach when pertinent, and with some attention to the disarray among federal agencies and other entities regarding the measurement of health and healthcare disparities. The recent reversal of the NCHS approach adds to the confusion in this area, and it remains to be seen whether and how recognition of that reversal will be reflected in the next *National Healthcare Disparities Report*.

The fundamental problem, however, is that none of the measures currently employed is sound and neither the government agencies involved with health and healthcare research, nor the nongovernmental entities and persons conducting such research, understand that the measures are unsound. And, as noted, the same problem exists with regard to all research into demographic differences in outcome rates.

Given the missions of COSSA and its member organizations, as well as the roles COSSA and those organizations have played in promoting sound research and influencing government funding of research into demographic differences, the organizations have a responsibility to take an active role in addressing the problems discussed above. Untold numbers of studies of demographic differences are ongoing or about to be undertaken without an understanding of the issues discussed above and in the materials listed in Section B. Similarly, countless entities are basing actions on the results of such studies and on the mistaken beliefs of federal civil rights law enforcement agencies regarding the implications of generally reducing adverse outcomes. Thus, consideration of the issues raised in this letter ought to be a priority of COSSA and its member organizations.

Please contact me if you have questions about any matter discussed in this letter.

Sincerely

/s/ James P. Scanlan

James P. Scanlan