Webinar Transcript
Measuring Social Inequalities in Health: Measurement and Value Judgments
Sam Harper of McGill University
NAACCR Seminar, May 24, 2011

Original recording is available on the Health Disparities Calculator Webinar web site.

[Slide1] Title Page

[Slide2] Thank you for the opportunity to speak to you today about measuring social inequalities in health. This is a vast topic but in the time we have today I'd like to consider the role that value judgments play in measuring health inequalities—and why I think it is important.

[Slide3] Why are we interested in monitoring health inequalities? There are multiple reasons, one of which is because we want to know whether they are getting better or worse over time. As most of you know the United States, like many countries, has two overarching population health goals. The first is to improve average health and life expectancy. The second is to eliminate health inequalities, and we actually have specific goals for health inequality that have been cornerstones of the "Healthy People Goals" that are put forth each decade. Healthy People 2010 first made eliminating health inequalities across a range of social groups, including categories such as race or ethnicity and socioeconomic position, a public health goal, which has been carried through to Healthy People 2020. These goals are laudable, but they also imply that we have methods for measuring and monitoring whether or not we are making progress toward actually eliminating health inequalities.

[Slide4] But measuring inequality is sometimes tricky business. The Nobel Prize-winning economist Amartya Sen, who has written widely on measuring economic inequality, noted that inequality is an ambiguous concept involving multiple dimensions. And he notes that "if a concept has some basic ambiguity, then a precise representation of that ambiguous concept [meaning, an actual measurement] must preserve that ambiguity." I think what Sen is saying here is that being clear about what we mean when we say 'inequality' is crucial for measuring it accurately. Okay, now that we recognize the importance of measuring health inequalities, and that the notion of inequality may be ambiguous, how should we actually measure health inequalities?

[Slide5] As it turns out, there are a large number of potential measures of health inequality. A few years ago some colleagues and I conducted a review of selected measures of inequality, and found a multitude of measures in use, and these measures differed across a number of different dimension such as scale and which reference group was used to measure departures from health equality. One consequence of having so many choices is that we may need to consider some of these issues before selecting an inequality measure.

[Slide6] So, today I'd like to talk about five issues to consider when choosing a measure of health inequality: it's simplicity, it's scale, how it weights individuals, how it weights different parts of the health distribution, and what reference level we use to measure health differences. Each of these choices, we will see, has consequences for our judgments about the magnitude and direction of changes in health inequality over time.

[Slide7] First issue: simple vs. complex measures of inequality.

[Slide8] This graph shows the proportion of black and white individuals <65 years of age without health insurance from 1998 to 2009. In the context of comparing only a few groups, simple inequality measures, such as the difference or ratio of rates make sense.

[Slide9] However, in some cases we are not necessarily interested in comparing two specific groups, but how much health inequality exists across an entire category like race/ethnicity. And recall that the Healthy People goals specified eliminating inequalities not just between blacks and whites, but across the entire category of race-ethnicity. If we start to include important Hispanic and Asian subgroups, the sheer number of comparisons makes using simpler measures more difficult. In these cases, we may want to use summary measures of inequality that can tell us how much variation exists across the entire range of a social group category.

[Slide10] Second issue: absolute and relative inequality. This has been one of the more thorny issues to deal with in the monitoring of health inequalities, and here is an example of why.

[Slide11] Here we show the incidence of esophageal cancer for blacks and whites from 1993 to 2004. This is an easy case, because it is obvious from the graphs that the lines are converging, and that inequality is decreasing (though it looks like this may partially be due to increasing incidence among whites, which may not be how we'd like to reduce inequality).

[Slide12] In any case, it is easy to confirm that inequality is decreasing, and we can calculate either the rate difference, shown here as the burgundy line, or the rate ratio, the blue line, and see that both measures are declining, which unambiguously suggests that racial inequalities in esophageal cancer incidence are going down.

[Slide13] But, unfortunately, this is not a typical case. Far more typical is the situation shown for prostate cancer mortality. Here we see both rates increasing until about the early 1990s, after which rates decline for both groups. But what's happening to inequality? It's a bit more difficult to see from the graph.

[Slide14] Delancey and colleagues looked at this issue in a recent paper, measured inequality using the rate ratio, and concluded that "racial disparities in mortality…increased over most of the interval since 1975."

[Slide15] That is true, but one could just as easily measure inequality using the black-white difference in rates. If we do that, we actually see a much different picture. According to the rate difference, inequality decreased by roughly 25%, but the rate ratio actually increased by roughly 10%. So, which answer is correct? Are black-white inequalities in prostate cancer increasing or decreasing? Unfortunately, it depends on which measure you choose. The underlying data are exactly the same, but the choice of inequality metric leads to completely opposing conclusions, a fact which was not mentioned in the Delancey article.

[Slide16] Decisions about whether to measure inequality in relative or absolute terms also have consequences for thinking about the impacts of new treatments, screening tools, or medical discoveries. Levine and colleagues looked at whether the introduction of highly active antiretroviral therapy in the 1990s increased black-white inequalities in mortality from HIV/AIDS, and concluded that "disparities widened significantly after the introduction of HAART."

[Slide17] How did they measure inequality? Using a relative measure, the mortality rate ratio comparing blacks to whites. This table shows that for most age groups the rate ratio nearly doubled after the introduction of HAART.

[Slide18] However, if one uses a measure of absolute inequality, the black-white mortality rate *difference*, once again we find exactly the opposite answer. After the introduction of HAART, which had huge effects on decreasing mortality in both blacks and whites, we see that absolute black-white inequality in HIV/AIDS mortality decreased precipitously, whereas relative inequality increased. These two examples show that measures of inequality are not value neutral, and the measures themselves include judgments about what aspects of inequality matter. For absolute inequality, what matters is by how much each group improved on the absolute scale—inequality went down because the black mortality rate declined by a greater absolute amount than the white rate. For relative inequality, what matters is how much each group improves on the relative scale, that is, the proportionate decline. Relative inequality increased because the white rate declined more in percentage terms than the black rate.

[Slide19] This issue is not specific to measuring health inequality. Economists struggle with the same issue when measuring economic inequality, which can also be measured on absolute and relative terms. This has led some economists to note that "there is no economic theory that tells us that inequality is relative, not absolute. It is not that one concept is right and the other wrong. Nor are they two ways of measuring the same thing. Rather, they are two different concepts." To achieve clarity in measuring health inequality, we should have a frank and open discussion about which of these two concepts we think matters more for population health.

[Slide20] Third issue: population weighting.

[Slide21] We often are measuring inequality across social groups like race-ethnic or geographic groups that have different population sizes. Should we consider population size? This diagram shows 3 potential ways of conceptualizing inequality. We have 3 social groups whose health is proportional to their height and who differ in population size. We want to measure the amount of inequality across the groups. Concept 1 treats each social group as exactly equal regardless of their population size. Concept 2 takes each group's size into consideration, meaning that because the "purple coat guy" group is 50% of the population we should weight their health 50%. Concept 3 inequality goes further and not only weights each group by their population size, it also accounts for the fact that health differs within each social group. As such, it is more like a global measure of inequality.

[Slide22] Now, this matters not only because social groups often differ by population size, but also because the size of social groups often changes over time. This graph shows the proportion of the US population in different education groups, and we can see that, as we all know, the proportion of the population with less than 12 years of education has declined dramatically in the past few decades, whereas the proportion of the population with greater than 12 years of education has increased dramatically? Should this matter when we are measuring trends in educational inequalities in health? One argument for doing so would be that health inequalities that negatively impact those with <12 years of education would have a much larger population health impact in 1965 than 2005, and those that negatively impact those with greater than 12 years a much greater impact in 2003 than 1965. If we think such considerations matter, then we want to weight by population size.

[Slide23] Similarly, we know that the United States has become and will become more racially diverse in the coming decades. If we are concerned about health inequalities affecting the Hispanic population, these are going to have a much greater population health impact in the future, and we may want to include population size in our measure of health inequality. Now, it just so happens that some measures of health inequality weight social groups by their population size and some don't and, as in the case of absolute and relative inequality, this has important consequences for our judgments about the magnitude and trend of health inequalities.

[Slide24] This graph shows changes over time in geographic inequalities across the 50 US states in stomach cancer mortality from 1950 to 2000. The blue line shows the Index of Disparity, which weights each state equally, and the red line shows the Mean Log Deviation, which weights each state by its population size. You can see that over time the two measures do not agree as to whether inequality is even increasing or decreasing. The Mean Log Deviation goes down, whereas the Index of Disparity increases. Again, the important message here is that the statistical measures of inequality impose certain value judgments on what aspects of inequality are thought to matter. The Mean Log Deviation weights every individual equally, and this means that California gets roughly 70 times the weight of Wyoming. The Index of Disparity weights each state equally, but this means that individuals in California count roughly 1/70th as much as individuals in Wyoming.

[Slide25] However, we often do not see the consequences of such value judgments in the literature on health inequalities. Here is another example. Ezzati and colleagues measured whether inequalities in life expectancy across US counties was increasing or decreasing over time. They used an unweighted measure of inequality, the standard deviation, and concluded that “there was a steady increase in mortality inequality across the US counties between 1983 and 1999.”

[Slide26] However, we performed a similar analysis and used both unweighted and population-weighted measures of inequality, and we found that the results were highly sensitive to this choice. We also found that unweighted inequality in county life expectancy increased by roughly 20%, but weighting counties by their population size led to the opposite conclusion, a roughly 10% decrease in inequality. Again, these opposing conclusions have nothing to do with the data, but rather with the choice of how to measure and express inequality.

[Slide27] So, in thinking about weighting we can see a rationale for both weighted and unweighted measures of inequality. Weighting individuals equally is consistent with what we already do when we estimate population average health such as overall life expectancy, and this also allows for inequality measures to be responsive to demographic change. On the other hand, weighting social groups equally (and therefore individuals *unequally* in most cases) may make sense if one is concerned with disproportionate impacts on small or marginalized social groups.

[Slide28] Fourth issue: sensitivity to different parts of the health distribution. What do I mean by the health distribution? By that I mean whether we care about which social group's health improves or worsens over time.

[Slide29] This graph shows hypothetical data on rates of smoking across socioeconomic position, measured here by education. You can see that the groups with lower education tend to have the highest rates of smoking.

[Slide30] Again, let's think about two possible ways of measuring inequality in health across education. I am showing formulas for two different measures of health inequality. The specific formulas are not important, but I put them up here just to show you how one can see that statistical measures of inequality actually carry value judgments. On the left you can see the Index of Disparity, a measure that does not weight by population size and takes each group's rate and subtracts it from the rate in the best group, here the group with more than 16 years of education. On the right we have the Mean Log Deviation, which also takes the difference between each group's rate and a reference level, but is population weighted and uses the difference in logarithm of the rates. These are both measures of relative inequality, so let's ignore the absolute/relative issue, and for this example I have fixed the population sizes to be equal, so we can also ignore that, but it turns out that take the difference in rates versus the difference in log rates makes an important difference.

[Slide31] To see why, let's suppose that we observe this population at two time points. Time 1 is the distribution I just showed you, but suppose 2 years later we that the smoking rate in the group with 12 years of education has declined from 30% to 25%, with no other changes. By how much did inequality decrease? According to the Index of Disparity, inequality decreased by 8% and by 5% according to the Mean Log Deviation. So far, so good.

[Slide32] But suppose, rather than the group with 12 years of education showing a decrease in smoking, it was the least-educated group—those with less than 12 years of education—where smoking declines. Suppose that their rate of smoking decreased from 40% to 35%, so we still have the only change from Time 1 to Time 2 being a 5-percentage point decline in smoking in one education group. Now what do the inequality measures say? Well, you can see that the Index of Disparity again shows exactly an 8% decrease in inequality—same as last time. However, the Mean Log Deviation, which you'll recall showed a decrease of 5% in the previous scenario, now shows a 15% decrease in inequality. Why? Well, because the Mean Log Deviation measures the difference in log rates, it is more sensitive to improvements in smoking that occur among the group with higher initial rates of smoking. On the other hand, the Index of Disparity simply takes the difference between each group's rate and the reference rate, and does not care whether the 5-percentage point decrease comes from the group with the highest or lowest initial rate of smoking. You might go so far as to say that, all things being equal, the Mean Log Deviation places additional weight on the health of those least healthy to begin with, whereas the Index of Disparity is agnostic about where health improvements come from. Again, the important point here is that statistical measures of inequality actually contain important value judgments that we should be aware of.

[Slide33] Okay, last issue: what is the right standard from which we should measure departures from health equality. What is the right reference group?

[Slide34] Let's take again our hypothetical data on smoking. From where should we measure differences in health? Some inequality measures use the population average rate as the reference point, but others do not. And some may allow us to specify the reference point. Does this make a difference? Let's measure inequalities in smoking at two time points using the Index of Disparity. We can calculate the Index of Disparity using either the population average rate, about 30%, as the reference group, or we can also use the Index of Disparity and use the best observed rate, the rate in Group D (about 10%), as the standard. What difference does this make? Suppose that between Time 1 and Time 2 smoking actually increases for Group C from 30% to 40%. If we use the population average as the reference group, this means that Group C is now actually closer to the population average, so inequality actually goes down by 7%. On the other hand, if we use the best rate (Group D’s rate) as the reference point, inequality increases, since Group C has now moved further away from Group D. Again, the point is that a seemingly innocuous choice about which reference group to use can have important consequences on whether we think inequality is going up or going down.

[Slide35] And this may have important consequences especially for health conditions that are actually worsening. Take obesity, for example. This graph shows obesity trends by education group among women, along with the Health People target rate of 10%. We can see what we all know to be true, that obesity has increased among all social groups, and all groups are now further from the target.

[Slide36] However, if we measure inequalities in obesity, we can see that inequalities across these 3 education groups are going down. This graph shows 3 absolute and 4 relative inequality measures, and regardless of whether which measure we use, or whether or not we weight by population size, inequalities are going down. The uncomfortable conclusion from this is that, from the standpoint of the Healthy People goals of eliminating health inequalities, this is an unambiguous success! We may want to consider how much value we place on reducing health inequalities if it comes at the expense of worsening overall health for everyone.

[Slide37] Now, it is easy for me to construct hypothetical examples of disagreement among these measures of inequality, but I would suggest to you that these are not isolated examples. My collaborators and I measured trends in health inequalities using multiple measures across 22 different cancer-related outcomes, and we found that in nearly half of all cases, substantive judgments about inequality trends could not be made without some sort of a priori decision about population weighting or absolute and relative inequality.

[Slide38] Thus, for health, as for income, implicit values—about things like scale or weighting—in empirical work matter greatly to the conclusions drawn about the distributive justice of income, and of health. And arguments can be made both ways.

[Slide39] And just to show you that this problem is not specific to health, but to the study of an ambiguous concept like inequality, these graphs show exactly the same issues affecting the measurement of global economic inequality—which is decreasing if you weight countries by population size but increasing if you don't, and increasing if you measure relative inequality but decreasing if you measure absolute inequality.

[Slide40] So, to sum up: Measures of health inequality are not value neutral. Things like the scale of measurement, weighting, and reference points have an important impact on our judgments of both the magnitude of health inequality and whether health inequalities are worsening or improving. We need to consider these issues carefully. Monitoring health inequalities requires both precise measurement and value judgments—they are inseparable. Finally, because inequality is such an ambiguous concept, it seems likely that a suite of health inequality measures is necessary to provide a complete description of the magnitude of inequality.

[Slide41] Resources, Methods, and Empirical Examples slide.

[Slide42] Acknowledgements slide.

[Slide43] Sam Harper's email address.