

**THE MISINTERPRETATION OF HEALTH INEQUALITIES IN THE UNITED
KINGDOM
(Oral Presentation)¹**

**James P. Scanlan
Attorney at Law
Washington, DC, USA
jps@jpscanlan.com**

*(British Society for Population Studies Annual Conference 2006, University of
Southampton, September 18-20, 2006)*

I am going to be talking about health inequalities in the UK and maintaining that, by and large, such inequalities have been looked at the wrong way ever since they were first studied. This is nothing peculiar to the UK (though certainly health inequalities are studied a good deal more in the UK than in most other places). The points I'll be making here apply to the way racial differences in health outcomes are studied in the US and the way racial, socioeconomic, and geographic differences in health outcomes are studied around the world.

Much of what I will be saying turns on a particular statistical tendency, and explaining that tendency has never been easy. So I am going to start out by summarizing my principal points and then turning to explaining that tendency.

[SLIDE 2]

My first slide sets out the main background points to my presentation. They are:

- During the almost 30 years in which socioeconomic inequalities in health have been studied in the UK, the perception has been that, at least as to mortality, those inequalities have been increasing.

¹ Presented in conjunction with PowerPoint presentation found at:
http://www.jpscanlan.com/images/Scanlan_BSPTS_Presentation.ppt.
Complete paper may be found at: http://www.jpscanlan.com/images/BSPTS_2006_Complete_Paper.pdf.

- Generally these inequalities have been measured in terms of relative differences in mortality rates, that is, the ratio of the mortality rate of the disadvantaged group to the mortality rate of the advantaged group.
- Similar perceptions exist with respect to racial, socioeconomic, and geographic health inequalities throughout the industrialized world, with the matter frequently being characterized in terms that “despite declining mortality, inequalities in mortality have increased.”

[SLIDE 3]

The next two slides present what I have termed “thesis,” even though that’s not quite the right word for what is just a description of the way numbers tend usually to work. In any case, what I call thesis here is the following:

- Almost all health inequalities research is suspect for failure to recognize the way relationships between the rates at which two group experience (or avoid) an outcome are influenced by the prevalence of the outcome.

Most notably, researchers have failed to recognize the tendency whereby, when an outcome like mortality declines, relative differences in experiencing it increase, while relative differences in avoiding it decline.

[SLIDE 4]

- Thus, relative differences in mortality have tended to increase, *because of declining mortality*, not despite declining mortality, and to be accompanied by declining relative differences in survival rates.
- But all measure of differences in experiencing or avoiding an outcome tend to be affected by changes in the prevalence of an outcome, raising questions of whether

we can determine whether inequalities are increasing or decreasing in any meaningful sense.

In the next slide I present what I call Interpretive Rule 1 or IR1 for short.

[SLIDE 5]

- When two groups differ in their susceptibility to an outcome, the rarer the outcome
- (a) the greater tends to be the relative difference in rates of experiencing it, and
- (b) the smaller tends to be the relative difference in rates of avoiding it.

These patterns can be illustrated with virtually any data that allows one to examine the rates at which two groups fall above or below any point on a continuum of factors associated with experiencing or avoiding some outcome. I will illustrate these patterns with hypothetical test data. I have elsewhere illustrated them with actual income data. But one should expect to find these patterns to exist generally wherever two groups have different risk distributions with respect to experiencing or avoiding some outcome.

[SLIDE 6]

Jumping ahead a little once again, the next slide provides a list of references, most of which are available on my web site. And to the extent I leave you puzzled or unpersuaded, I would hope that these references might make my case a little better than I am able to do here. In fact, there may be 20 or so references on my web site where I discuss the implications of IR1 with respect to the interpretation of group differences in the law and the social and medical sciences. Incidentally, I am not a statistician or a demographer. I am a lawyer in the US and the things I say about statistics flow

generally out of issues that arise from the use of statistics in litigation in the US, mainly employment discrimination litigation.

[SLIDE 7]

The next slide contains two additional references. I cite the SEPHO Handbook of Health Inequalities Measurement (which was issued late last year) because it essentially accepted the principal point I am making today on the way differences between rates of experiencing or avoiding an outcome will change simply as a result of changes in the prevalence of an outcome. However, it is a pretty comprehensive document that discusses a wide variety of measures. And, in my view, it fails to consider the full extent of the implications of the acceptance of what I call IR1. The second reference is a major report of the US National Center for Health Statistics, which was also issued last year. It for the first time recognizes that one might appraise changes over time differently depending on whether one examines favorable or adverse outcomes. But it fails to acknowledge that relative differences in favorable and adverse outcomes tend to change systematically in opposite directions when the prevalence of an outcome changes, and it merely recommends that all differences be measured in terms of adverse outcomes. So NCHS has simply avoided for the time being an issue that I think it eventually will have to confront.

The figures that follow are based on the normal distribution of two groups of equal size with respect to factors associated with experiencing or avoiding some outcome. I term them simply the advantaged and the disadvantaged groups, or AG and DG for short. The distributions have the same standard deviation and the means differ by one half a standard deviation. Conceptually, one might regard these figures to be based

on distributions of test scores and imagine that we are observing the implications of lowering a cutoff on the test. But the patterns we observe would hold as well if, rather than lowering a cutoff from one point to another, we improved test performance such that everyone currently scoring between the two points was enabled to score at the higher cutoff.

[SLIDE 8]

Along the bottom of Figure 1 we see various letters and figures. The letters are merely to afford ease of reference. The numbers reflect the failure rates of the advantaged group (the higher scoring group), used here as reference points. Thus, as we move from left to right we observe the implications of lowering the cutoff score, from Point A (where 99 percent of the AG fails the test), to point H (where 50 percent of AG fails the test), to point 0 (where 1 percent of Group A fails the test) – and at various points in between.

The blue line in Figure 1 then shows the proportion that DG (the disadvantaged group) comprises of the part of the total population falling below each cutoff. And we observe that, as the cutoff is lowered, and test failure becomes rarer, the disadvantaged group comprises a higher proportion of the total continuing to fail the test.

Recognizing this aspect of the matter is important to understanding why ratios of rates of experiencing some adverse outcome tend almost invariably to increase as the outcome declines. For progress in virtually every area of human well-being, including reductions in mortality, is generally a matter of serially restricting adverse outcomes to the points where only the most susceptible segments of the overall population continue to experience those outcomes – until, in an ideal world, the adverse outcomes disappear

entirely. And disadvantaged groups tend to comprise larger proportions of each increasingly more susceptible segment of the overall population. Thus, the closer a society comes to eliminating an adverse outcome, the more the outcome will be concentrated within disadvantaged groups, and the greater will be the relative difference between the rates at which advantaged and disadvantaged groups experience it.

[SLIDE 9]

Figure 2 now shows the relative difference in rates of failing the test. And we observed that as the cutoff is lowered, the difference in failure rates increases. Figure 2 is just a corollary to Figure 1. Again, I am using a hypothetical distribution of test scores as an illustration. As I noted, I have elsewhere used income data. But the basic pattern observed is virtually universal in settings when the distributions are other than highly irregular. And it is always for the same reason: that is, as we saw in Figure 1, reduce an adverse outcome and it will tend to be increasingly concentrated in disadvantaged groups. Correspondingly, the relative difference in experiencing the outcome increases.

[SLIDE 10]

Now let us examine the other side of the picture – the relative difference in experiencing the favorable outcome. Figure 3 shows the ratio of the rate at which the advantaged group falls above the line to the rate at which the disadvantaged group falls above the line. Conceptually, we can regard these as the relative pass rates. That ratio declines – *i.e.*, the difference grows smaller – as the overall failure rate declines. Thus, we observe that the size of relative inequalities in experiencing an outcome and in avoiding the outcome tend to move systematically in opposite directions as the prevalence of the outcome changes.

This has the following important implication with respect to the evaluation of changes in the size of inequalities. Some might be inclined to maintain that an increase in the difference between rates of experiencing an adverse outcome reflects some true worsening of the relative status of the disadvantaged group, even when the increase results solely from a general decline in the prevalence of the outcome. Even allowing the validity of the point for a moment, one would have to regard such a change as a much different occurrence – and a far less consequential occurrence – than a change that went beyond the usual consequences of the overall decline in the outcome. But it becomes difficult even to maintain that an increase in the relative difference in adverse outcomes that flows solely from a decrease in prevalence somehow reflects a true worsening of the relative situation of the disadvantaged group when one recognizes that, if one appraises the same matter in terms of the favorable outcome, one has to conclude that the inequality has declined.

As it happens, relative differences in many indicators have traditionally been measured in terms of the favorable outcome. In the US, where laws limit the use of employment tests on which minorities or women do not perform as well as whites or men, relative performance on tests has generally been examined in terms of pass rates. And because the lowering of cutoffs tends to reduce relative differences in pass rates, the lowering of cutoffs has been universally regarded as reducing the disproportionate impact of such tests on minorities or women, even though lowering cutoffs increases differences in failure rates. Beneficial health procedures (*e.g.*, prenatal care, immunization, mammography) have traditionally been evaluated in terms of differences in rates of receiving the procedure. Thus, the increased availability of such procedures has led to a

perception that inequalities are declining, even as that same increased availability, by reducing certain types of mortality, has led to the perception that racial differences in those types of mortality are increasing.

I mentioned that in the US the NCHS has recommended that all disparities be measured in terms of adverse outcomes, including evaluating differences in failing to receive mammography rather than in rates of receiving mammography. To the extent that the recommendation is followed, differences that once were thought to be shrinking now will be thought to be widening. But, as I mentioned, the NCHS has yet to straightforwardly deal with the fact that differences in adverse outcomes and favorable outcomes move systematically in opposite directions when the prevalence of an outcome changes.

[SLIDE 11]

Now consider odds ratio, the ratio of one group's odds (*i.e.*, the rate of experiencing an outcome divided by the rate of avoiding it) divided by the odds of the other group. Some commentators favor the use of odds ratios because one gets the same result whether one focuses on the adverse or the favorable outcome. And, given that odds ratios are functions of rates of both experiencing and avoiding an outcome, it warrants consideration whether the odds ratio might offer a useful means of evaluating changes in the relative well-being of two groups vis-à-vis experiencing and avoiding some outcome.

But in order to determine whether inequalities in outcomes like mortality (or survival) are changing in ways that are not solely functions of changes in the prevalence of the outcome, one needs a measure that does not change when there occurs a simple

across-the-board change in prevalence – “across the board” meaning a change akin to lowering the cutoff from Point J to Point K. But we see in Figure 3 that the odds ratio is very large when the failure rate is very large; grows smaller as the failure rate declines towards the area where a majority of AG passes the test; then grows large again as failure becomes rare. So the odds ratio does not provide a useful means of identifying changes between the relationship of the rates at which two groups experience some outcome that are not solely the result of changes in the prevalence of an outcome.

[SLIDE 12]

Some commentators favor the use of absolute differences rather than relative differences to measure health inequalities. Reasons for this preference include (1) that the absolute difference is the same whether one examines the adverse or the favorable outcome and (2) that the absolute difference gives a better picture than the relative difference of the proportion of the disadvantaged group that is harmed by its greater susceptibility to an adverse outcome. But we see in Figure 4 that, as with each of the other measures just described, absolute differences also change when there occurs an across-the board change in the prevalence of the outcome. The absolute difference is small at the point where almost everyone from both groups experiences the adverse outcome, grows larger as the adverse outcome becomes less common, and then grows small again as the adverse outcomes becomes rare. Thus, absolute differences cannot provide an efficient and reliable means of identifying changes in the relative status of two groups with respect to some outcome that are not solely a function of changes in prevalence of the outcome.

There are varied other measures of health inequalities (including some that attempt to take into account changes in the size of the groups being compared). Whether or not there would otherwise be something to be said for such measures, so far as I can tell, they all seem to suffer from the problem that, like the measure just discussed, they change when there occurs a change in the prevalence of an outcome. Hence, they cannot provide a ready means of identifying changes that are not solely consequences of changes in the prevalence of an outcome.

So is there way that by taking these expected patterns into account, we can identify meaningful changes between the rates at which two groups experience or avoid some outcome? Well, I refer you to the recent *Chance* article and the Athens paper for rather pessimistic appraisal of such prospects. I will add one thing, however. I am sure some of you have been thinking that my examples are unrealistic in assuming complete normality of the risk distributions. And that's true (though it's not true as to the income examples I have used elsewhere). But I do acknowledge that in reality there is going to be a lot of departure from true normality. Yet the patterns I have described are nevertheless, in general, going to be present most, if not all, of the time. That prevents us from sensibly analyzing changes over time the way we have been doing to date. At the same time, however, the very fact that we really cannot see the extent to which the underlying distributions depart from the perfectly normal interferes with our ability to try and interpret changes while taking the above-described tendencies into account.

All that said, in my time remaining I will go through a few implications of IR1 and the failure to recognize it. They are not particularly well organized, but they provide some focus for further thinking on these issues.

[SLIDE 13 - IMPLICATIONS 1

First, we don't really know whether inequalities have been increasing or decreasing in any meaningful sense in the UK or elsewhere. Second, it is not clear that there exist effective tools for measuring changes over time.

Third, a few words on morbidity: As to acute morbidity, the patterns just described generally obtain. That is, most acute morbidity has been declining; hence we should expect increasing relative differences in experiencing these conditions and declining relative differences in rates of avoiding them. But the issues are somewhat different with respect to non-acute morbidity, which is usually studied in terms of self-assessed health less than good. For one thing, inequalities in health less than good are generally studied in terms of odds ratios. As shown earlier, changes in odds ratios can be even more difficult to interpret than changes in relative differences. But, in general, rates of health less than good are not declining and hence inequalities usually are not deemed to be increasing. Morbidity may nevertheless offer some possibilities for determining whether inequalities are changing in ways that are not simply functions of changes in prevalence. One approach would be to examine changes in differences between blood pressure levels and other factors such as are discussed in a 2002 article by Ferrie *et al.* in the *Journal of Epidemiology and Community Health*.

[SLIDE 14 - IMPLICATIONS 2]

Policy Implications: It is not clear that IR1 has many policy implications with respect to how we address health inequalities. That is, fully appreciating IR1, it still will make sense to implement the same types of measures that are currently implemented with the expectation that they would reduce inequalities. It simply needs to be recognized

that, by reducing adverse outcomes, such measures are likely to increase relative differences in rates of experiencing those outcomes, and that such increase need not reflect a meaningful worsening of the relative situation of disadvantaged groups.

As to the Whitehall Studies, that the social gradient in mortality was greater among British civil servants than in the UK population at large has been the basis for many inferences about health inequalities. Most of those inferences are suspect, however, because it was not appreciated that relative differences in mortality were comparatively large among British civil servants simply because mortality was low in that group.

Similarly, much has been made of the comparatively large mortality differentials in egalitarian countries like Sweden and Norway. Yet such differentials are hardly surprising simply because mortality is low in these countries. The same holds with respect to large racial differences in infant mortality where parents are highly educated. Large relative differences in mortality (and small relative differences in survival) are to be expected among the highly educated simply because infant mortality is so low among the highly educated.

[SLIDE 15 – IMPLICATIONS 3]

Ameliorative interventions; exacerbating factors: The same tendencies underlying the patterns described above suggest that ameliorative interventions will reduce mortality more among groups with lower mortality but increase survival more among groups with high mortality. Exacerbating factors, like smoking or obesity, will tend to increase mortality more among groups with lower mortality but reduce survival more among groups with higher mortality.

Legal settings (mortgage rejection rate disparities, termination rate disparities): I will just observe here that, because the tendency I have described is almost universally misunderstood, we observe anomalous results in a host of settings where legal implications arise from the way one measures differences between the rates at which two groups experience or avoid some outcome.

The final issue is: just what is a large difference? Once we appreciate that a three-fold difference may not be greater than a two-fold difference in any meaningful sense, knowing just what we should regard as a particularly large disparity is not so easy. But certainly it is impossible to draw useful conclusions on such issues without appreciating the ways the sizes of differences are affected by the prevalence of an outcome.