Interpreting Differential Effects in Light of Fundamental Statistical Tendencies

James P. Scanlan

2009 Joint Statistical Meetings
Washington, DC
August 1-6, 2009

Oral Presentation

[The PowerPoint presentation accompanying this oral presentation may be found [here](). Issues discussed in the presentation have been further developed on the [Subgroup Effects]() sub-page of the [Scanlan's Rule]() page of jpscanlan.com and materials made available by means of that sub-page[1]].

My name is James Scanlan. I am an attorney here in Washington. My title is "Interpreting Differential Effects in light of Fundamental Statistical Tendencies"

SLIDE 2 – SUMMARY

I want to show two things here today: first, that factors that similarly affect two groups with different baseline rates of an outcome will tend to show a larger proportionate effect on the outcome for the group with the lower base rate but a larger proportionate effect on the opposite outcome for the other group. Second, that true subgroup effects can only be identified by determining the degree to which a factor shifts each group's risk distribution

SLIDE 3

---

[1] Such materials include the following on-line comments:

(1) Interpreting effects on referral inequalities of policies that increase overall referral rates. BMJ May 6, 2011 (responding to McBride D, Hardoon S, Walters K, et al.. Explaining variation in referral from primary to secondary care: Cohort study. BMJ 2010;341:c6267):
http://www.bmj.com/content/341/bmj.c6267.full/reply#bmj_el_260111;

(2) Subgroup analyses should not be undertaken without consideration of methodological issues. Journal Review Nov. 27, 2010 (responding to Mullins CJ, Onukwugha, Cooke JL, et al. The potential impact of comparative effectiveness research on minority populations. Health Affairs (Millwood) 2010;29(11):2098-2104): http://content.healthaffairs.org/content/29/11/2098.abstract/reply#healthaff_el_457426;

(3) Problems in identifying interaction where groups have different base rates. *BMJ* Sept. 21, 2010 (responding to Altman DG, Bland JM. Interaction revisited: the difference between two estimates. BMJ 2003;326:219): Altman DG, Bland JM. Interaction revisited: the difference between two estimates. BMJ 2003;326:219): http://www.bmj.com/content/326/7382/219/reply#bmj_el_241943

(4). Assessing heterogeneity of treatment effects in light of fundamental statistical tendencies. Trials May 26, 2011(responding to Kent DM, Rothwell PM, Ionnadis JPA, et al. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. Trials 2010,11:85): http://www.trialsjournal.com/content/11/1/85/comments#498686

(5) Systematizing the analysis of effect heterogeneity requires rethinking some fundamentals. Trials June 1, 2011 (responding to Gabler NB, Naihua D, Liao D, et al. Dealing with heterogeneity of treatments effects: is the literature up to the challenge. Trials 2009,10:43): http://www.trialsjournal.com/content/10/1/43/comments#512684

Because my topic is complex, I like provide few references at the outset. These will take you to 100 or places where I have illustrated the central point here – that is, that standard measures of differences between rates are problematic because all are affected by the prevalence of an outcome. These are all available on jpscanlan.com. This presentation will also be there in a day or two. The materials on the site should clarify things that I fail to address in sufficient detail here today.

SLIDE 4

I term as an "illogical expectation" the tendency to regard it as somehow normal that a factor that similarly affects two groups' susceptibilities to an outcome will cause the same proportionate change in the outcome rates for each group and to regard anything else as a differential effect (subgroup effect, interaction)

SLIDE 5 – ONE REASON IT IS ILLOGICAL

One reason the expectation is illogical is that where two groups have different base rates of experiencing an outcome, a factor that has the same proportionate effect on each group's rate of experiencing the outcome necessarily has a different proportionate effect on each group's rate of avoiding the outcome. The figures shown at the bottom are are outcome rates and opposite outcome rates (i/e. 100 minus the first rate) for two groups with different base rates. Without my having to go through the math, these figures ought to show why it is impossible to have the same proportional change for each group for both outcomes. But that impossibility is one reason that it is illogical to expect it for either.

SLIDE 6 - A LOGICAL EXPECTATION

What it is logical to expect is that a factor that similarly affects two groups with different baseline rates of an outcome will tend to show a larger proportionate effect on the outcome for the group with the lower base rate but a larger proportionate effect on the opposite outcome for the other group.

Now, why is that expectation logical?

SLIDE 7 – SCANLAN'S RULE

I have here what I have sometimes called heuristic rule x or interpretive rule 1. Researchers in the UK have lately called it Scanlan's rule. I am adopting that usage, slightly modified, and call it Scanlan's Rule 1 (or SR1).

When two groups differ in their susceptibility to an outcome, the rarer the outcome, the greater tends to be the relative difference in experiencing it and the smaller tends to be the relative difference in avoiding it.

2

Scanlan's Rule 2, incidentally, involves the ways absolute differences and odds ratios are correlated with overall prevalence. But that subject is well covered on the website and we don't need to get into that here.

Let's look at a few illustrations of SR1.

SLIDE 8 – FIGURE 1 – FAIL RATIOS

Figure 1 is based on two test score distributions, where the distributions have the same standard deviation and where the advantaged group (AG) has an average that is one half a standard deviation greater than the average for the disadvantaged group (DG). The numbers along the bottom are AG's failure rates, which are used as benchmarks for overall prevalence. The blue line plots the ratio of DG's failure rate to AG's failure rate at each point. And as we move from left to right we observe the effects on that ratio of lowering cutoffs such as to serially enable the population between each point to pass the test. And we observe that as cutoffs are lowered, and test failure becomes less common, the ratio of AG's failure rate to DG's failure rate increases.

SLIDE 9 - FIGURE 2 – PASS RATIOS

Figure 2 shows the opposite side of the picture – the relative difference in pass rates – here presented in terms of the ratio of AG's pass rate to DG's pass rate. As cutoffs are lowered, and failure becomes less common (passing becomes more common), the relative difference in pass rates decreases.

Thus do we observe the way that relative differences in experiencing an outcome and relative difference in avoiding an outcome tend to move systematically in opposite directions as the overall prevalence of an outcome changes.

SLIDE 10 – FIGURE 3 – INCOME DATA

Figure 3 illustrates the same thing with income data for blacks and whites. It shows, for example, how lowering poverty tends to increase relative differences in poverty while reducing relative differences in rates of avoiding poverty.

SLIDE 11 – FIGURE 4 – NHANES

Figure 4 illustrates the same thing with NHANES data on systolic blood pressure. It shows how lowering hypertension will tend to increase racial differences in hypertension rates and reduce racial differences in rates of avoiding hypertension.

SLIDE 12 – IMPLICATIONS OF S2 1

This slide lists some of the implications of SR1. The first two I have discussed in many places. Here we are concerned with the third:

Factors the decrease/increase prevalence of an outcome will tend to have a greater proportionate effect on group with lower baser rate but a greater proportionate effect on the opposite outcome for the other group

SLIDE 13 – PROBLEMS ARISING FROM SR1

When we observe standard patterns of changes in differences between group rates as prevalence of an outcome changes, how do we determine whether the difference between the groups' situations actually changed in a meaningful way and how do we quantify the difference at each point in time (in a way that is unaffected by overall prevalence)?

Of particular interest today, how do we identify genuine differential effects?

SLIDE 14 – ESTIMATED EFFECT SIZE

Estimated effect size (EES) = estimated difference, in terms of percentage of a standard deviation, between means of hypothesized underlying, continuously-scaled normal distributions of factors associated with experiencing an outcome, derived from adverse outcome rates of AG and DG

The EES was the subject of my JSM 2008 presentation and it is treated under the Solutions and Solutions Database tabs of the Measuring Health Disparities page of jpscanlan.com. From there you can download an Access database to implement this approach.

SLIDE 15 – TABLE 2 – EXAMPLE

Table 2 presents a situation where adverse outcomes decline and the relative difference changed in the standard way – that is, the relative difference in the adverse outcome increased while the relative difference in the favorable outcome decreased.

Was there a real change one way or the other?. The EES, which we derive from each group's failure rate in each year, indicates that the difference declined slightly – from .50 to .47 standard deviations.

SLIDE 16 – TWO PERSPECTIVES

I am going to present a few examples. But I first note that all of these issues can be examined from two perspectives – (a) in terms of the differential effect on two groups or (b) in terms of the size of the group difference in two settings.

SLIDE 17 – TABLE 3 – EFFECT ON

The data in table 3 are drawn from online calculators showing heart attack risks based on the Framingham studies. It is based on men and women who are similar with regard to each of the risk factors (as indicated in parentheses).

The first data row shows heart attack risk for women without and with hypertension. I arbitrarily used systolic blood pressure rates of 120 for non-hypertensive and 150 for hypertensive. The fourth column show the reduction in risk for non-hypertensive compared with hypertensive and that the fifth column shows the increase in likelihood of not having a heart attack for the non-hypertensive versus the hypertensive.

The second data row shows the same thing for men.

And we see from those two columns what I have suggested one should expect in the circumstances – i.e., a greater reduction in the risk of the adverse outcome for the group with the lower base rate (women) but a greater increase in the likelihood of the opposite outcome for men. But is there any real difference in the effect of hypertension (or control of it) on women and men? The EES – which here is derived from the hypertensive and non-hypertensive rates – suggest that effect is somewhat greater for women

SLIDE 18 – TABLE 4

Table 4 presents a similar analysis from what I called Perspective 2 – and it shows that the male-female difference is somewhat greater among the non-hypertensive than the hypertensive.

SLIDE 19 – TABLE 5

Table 6 involves a 2006 NEJM study[2] that received a lot of attention for its finding that among light smokers blacks were much more likely to contract lung cancer than whites, but that the difference was much reduced or eliminated among smokers of more then 29 cigarettes.   Here the second perspective is more illuminating.

The relative differences observed were again what one ought to expect in the circumstances – that is the racial difference in the adverse outcome was large was larger among the light smokers (where the outcome was rarer), but the relative difference in the favorable outcome was smaller in that group. In such circumstances, only a measure like the EES can offer useful information on where the disparity was larger.   And it seems in fact to have been larger among the lighter smokers.

Slide 20 – TABLE 6

---

[2] Haimon CA, Stram DO, Wilkens LR, *et al.* Ethnic and racial differences in the smoking related risk of lung cancer. *N Engl J Med* 2006;354:333-42.

Table 6 offers Perspective 1. We see the expected patterns of differences between light and heavy smokers for each race – as to the favorable and the adverse outcome. But the EES tells us the difference was larger among whites.

SLIDE 21 AND 22 – TABLES 7 AND 8 – GOTTLIEB

Tables 7 and 8 offer us the same two perspectives from a study exploring differential effects of beta blockers on mortally among high and low risk heart patients.[3] Age groups are used for the two risk groups. Table 7, providing Perspective 1, shows the standard pattern of relative differences with regard to the effects of beta blockers on low and high risk groups – that is, a greater reduction in the adverse outcome for the low-risk group, but a greater increase in avoidance of the outcome for the high-risk group. The final column indicates that there is a slightly greater effect for the low-risk group.

Table 8, providing perspective 2, shows the standard patterns of relative differences between groups for those with and without beta blockers – that is, greater relative difference by age for the adverse outcome for those with beta blockers (where mortality is lower), but greater relative difference by age with regard to avoiding mortality for among those without beta blockers. And the final column indicates a slightly greater difference by age in beta blocker group.

SLIDE 23 – TABLE 23 – ISSUES

I do not mean to suggest that we will observe the described patterns in every case. In fact, there is a body of literature – mainly meta analyses – where which groups with higher base adverse outcome rates tend to show larger proportionate reductions in those outcomes (such as in the sources summarized in Sackett [4]). Possibly absolute minimums have some role in this, but I'll leave the discussion to my web page. In any case, I am inclined to think that the common result will be one where there exists a sound theoretical basis for it. And the expectation I describe is the one with a sound theoretical basis.

Clinical significance issue: It might be observed that if a factor reduces an adverse outcome for one group proportionately more than for another, as far as clinical decision-making goes, what difference does it make if these patterns are to be expected in the circumstances? I think the answer here is that the decision-make ought to be looking at absolute effects in any event.

---

[3] Gottlieb SS, McCarter RJ, Vogel RA. Effect of bet-blaockade on mortalitly among high-risk and low-risk patients with myocardial infarction. *N Engl. J Med* 1998;339:489-97.

[4] Sackett DL. Why randomized controlled trials fail but needn't: Failure to employ physiological statistics, or the only formula a clinician-trialist is ever likely to need (or understand!). *JAMC* 2001:165(9): 1226-1237

Statistical significance: Interestingly most discussions about spurious subgroup effects involve statistical significance issues. With regard to what I term spurious subgroup effects, when the observed pattern is in the direction that the statistical forces will tend to drive it, the points I make provide a reason why one ought to regard the pattern to be other than chance-driven. That is, while the pattern should not be regarded as chance-driven, it still may be spurious.

Statisticians who are intrigued by my solutions/EES approach frequently raise the issue of confidence intervals. That issue is addressed a bit on the website, but it remains unresolved. But at least one epidemiologist is working on it.