Calculating measures of treatment effect for use in clinical practice

Efstathia Andrikopoulou, MD,[a] and Charity J. Morgan, PhD[b]

[a] Division of Cardiovascular Disease, University of Alabama at Birmingham, Birmingham, AL
[b] Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL

Received Dec 24, 2015; accepted Dec 24, 2015
doi:10.1007/s12350-016-0394-6

Over the past few decades, an ever-growing number of clinical trials and studies have provided physicians with a wealth of data and results. Some of these findings may play a significant role in clinical practice and even help define clinical guidelines. It is up to physicians to determine which studies are useful in their own practice and patient population and which are not. The ability to translate results from studies and trials into clinical practice relies primarily on understanding of statistical analysis. There is a wide range of statistical measures available to assess different types of studies. In this editorial, we will present easily calculable markers of treatment effect—both benefit and harm—that can be used to critically evaluate randomized prospective studies.

For simplicity, we will only consider trials in which results are presented in rates or proportions, e.g., the proportion of subjects responding to treatment or the rate of an adverse event. An example of this type of endpoint would be the mortality rate, which is one of the most common outcomes used in cardiovascular studies to compare the effect of a new drug or intervention to that of either standard of care or placebo. If we calculate the difference in mortality rates between the treatment under study to that of either standard of care or placebo, then this will give us the absolute risk reduction (ARR) or risk difference. In this case, the ARR reflects how much the new treatment lessens the risk of death. Figure 1A gives a visual representation of the ARR.

The reciprocal of the ARR yields the number needed to treat (NNT). First proposed by Laupacis et al[1], the NNT provides the average number of patients that would need to receive a treatment in order to have one patient benefit from that treatment. It therefore expresses treatment benefit in terms of patients, rather than proportions or probabilities, which can make it more appealing to clinicians to interpret and patients to understand.[2] The NNT is usually rounded up to the next integer.

A limitation of both the ARR and NNT is that they are not sensitive to changes to the underlying event rates. Practically, this means that an ARR of 2% may be clinically significant if the event rates are 1% and 3%, but less so if the rates are 40% and 42%.

The relative risk (RR) is the ratio of the risk of the event for treated subjects to the risk for controls. For example, an RR of 0.5 (sometimes also expressed as a percentage, e.g., 50%) would mean that the risk for subjects receiving treatment is half of that for controls. A RR of 1 would indicate that the treatment offers no benefit compared to control. Smaller values of RR reflect greater treatment benefit.

It is crucial to understand that studies with the same ARR may have dramatically different RR. In the previous example, if the event rates are 1% for treated subjects and 3% for controls, then the RR is 0.33; that is, the treated subjects are subject to only a third of the risk of controls. In contrast, if the rates are 40% and 42% for treated and control, respectively, then the RR is 0.95 and the treated subject have very little decrease in risk compared to the controls.

The relative risk can also be expressed as the relative risk reduction (RRR), which equals the RR subtracted from 1. The RRR is interpreted in a similar way as the RR, with the exception that a RRR of zero indicates no treatment benefit and larger values of RRR are associated with greater benefit.

Similar to measures quantifying the benefit of a particular treatment, there are measures assessing the harm from a treatment. Just as the ARR measures how much the treatment reduces risk compared to the control group, the absolute risk increase (ARI) is a measure of

Reprint requests: Charity J. Morgan, PhD, Department of Biostatistics, University of Alabama at Birmingham, 1720 Second Avenue South, Birmingham, AL, 35294-0022; *cjmorgan@uab.edu*
J Nucl Cardiol 2017;24:188–90.
1071-3581/$34.00

the additional risk incurred by treated subjects. Like the ARR, the ARI is the difference in event rates. The only difference between the two is the direction of the effect. A graphical depiction of the ARI is shown in Figure 1B.

Furthermore, just as the NNT is the reciprocal of the ARR, the reciprocal of the ARI is the number needed to harm (NNH).[3] The interpretation of the NNH is similar to that of the NNT, except the NNH is the average number of patients that would need to be treated in order to see an adverse outcome.

The RR may also be used to describe treatment harm. If the risk of an adverse event is higher under a new treatment than with standard of care or placebo, then the RR will be greater than 1. For example, a RR of 1.5 would indicate that the risk of the event with treatment is 1.5 times that of standard of care. Larger values of RR are associated with greater treatment harm. As mentioned above, a RR of 1 means that there is no difference in risk between treated subjects and controls.

Similar to how the reduction in the RR can be used to assess treatment benefit, we may also consider the relative risk increase (RRI) to quantify treatment harm. The RRI is computed by subtracting 1 from the RR. For example, a RR of 1.5 would translate to a RRI of 0.5, or a 50% increase in the risk of the event for patients receiving treatment. A RRI of 0 means that there is no increase in risk associated with the treatment; larger values of RRI are associated with greater harm.

The formulas for each of these measures are summarized in Table 1. Several measures of treatment effect are discussed in further detail in Cook and Sackett.[2]

We demonstrate these calculations with a simple example. Table 2 depicts the results of a hypothetical randomized clinical trial that compares a new treatment (Treated) to standard of care (Control). One of the efficacy endpoints is the all-cause mortality rate, which was significantly higher in controls (2.4%) than in treated subjects (1.0%). Based on the data from Table 2, we can calculate the RR for all-cause mortality: 0.01/0.024 = 0.42 (or 42%). The RRR is 1-0.42 = 0.58, which means that there is a 58% decreased risk of dying in patients receiving the new treatment compared to those receiving standard of care. The ARR is 0.024-0.01 = 0.014, or 1.4%. The NNT is 1/0.014 = 71.4, or 72 rounding to the next integer. That is, on average, we would need to treat 72 patients in order for one patient to benefit.

We illustrate the calculation of measures of harm using as an example the adverse event of palpitations. From Table 2, we see that in our hypothetical trial, treated subjects had a significantly higher rate of palpitations than control subjects (5.5% vs 3.1%, respectively). The RR for palpitations is 0.055/0.031 = 1.77. The RRI is 1.77-1 = 0.77, which means

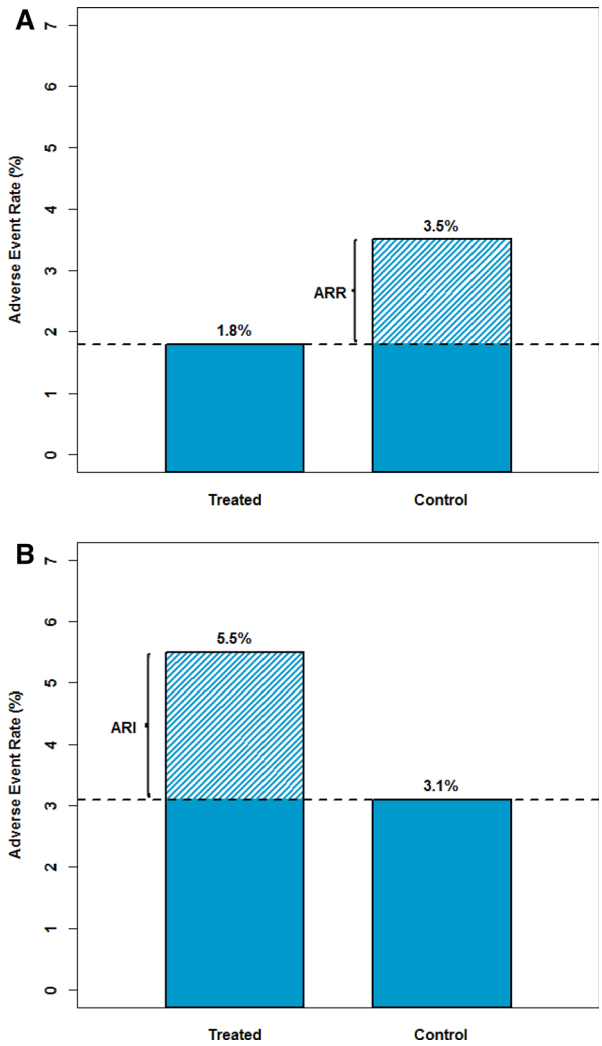

Figure 1. Examples of (**A**) the absolute risk reduction (ARR) and (**B**) the absolute risk increase (ARI) for treated subjects relative to controls.

that there is a 77% higher likelihood of experiencing palpitations in patients receiving the new treatment. The ARI for treated subjects relative to controls is the difference between the rates of palpitations: 0.055-0.031 = 0.024, or 2.4%. The NNH is 1/0.024 = 41.7, or 42 rounding to the nearest integer. On average, 42 patients would need to be treated in order to see one additional case of palpitations.

Each of these methods discussed here has its own advantages and limitations. Treatment benefit (and harm) cannot be adequately summarized by a single measure. It is recommended to report the baseline rate (i.e., the event rate for controls or standard of care) along with whichever measures of treatment effect are used.[4] The goal of this editorial was to present methods of estimating treatment effects that are frequently reported

Table 1. Formulas for measures of treatment benefit and harm

Benefit	
Absolute risk reduction (ARR)	$\mathrm{ARR} = p_C - p_T$
Number needed to treat (NNT)	$\mathrm{NNT} = \frac{1}{\mathrm{ARR}} = \frac{1}{p_C - p_T}$
Relative risk (RR)	$\mathrm{RR} = \frac{p_T}{p_C}$
Relative risk reduction (RRR)	$\mathrm{RRR} = 1 - \mathrm{RR} = 1 - \frac{p_T}{p_C}$
Harm	
Absolute risk increase (ARI)	$\mathrm{ARI} = p_T - p_C$
Number needed to harm (NNH)	$\mathrm{NNH} = \frac{1}{\mathrm{ARI}} = \frac{1}{p_T - p_C}$
Relative risk increase (RRI)	$\mathrm{RRI} = \mathrm{RR} - 1 = \frac{p_T}{p_C} - 1$

p_T is the event rate for treated subjects: $p_T = \frac{\#\text{ of treated subjects with event}}{\text{Total }\#\text{ of treated subjects}}$

p_C is the event rate for controls: $p_C = \frac{\#\text{ of control subjects with event}}{\text{Total }\#\text{ of control subjects}}$

Table 2. Results for a hypothetical trial

	Treated (*N* = 1557)	**Control (*N* = 781)**	***P***
Clinical outcomes			
Major adverse cardiovascular effects (MACE)	28 (1.8%)	27 (3.5%)	.019
All-cause mortality	16 (1.0%)	19 (2.4%)	.014
Adverse effects			
Palpitations	85 (5.5%)	24 (3.1%)	.013
Memory impairment	20 (1.3%)	5 (0.6%)	.224
Increase in serum creatinine (≥ 0.5 mg/dL from baseline)	78 (5.0%)	12 (1.5%)	<.001

in published studies and explain their meaning via simple examples. Even when not directly reported, the above-mentioned indices can be easily calculated by clinicians themselves. Thus, physicians and patients alike can readily interpret study results and use these measures of treatment benefit and harm to make informed decisions.

Disclosure

The authors have no conflicts of interest to disclose.

References

1. Laupacis A, Sackett D, Roberts R. An assessment of clinically useful measures of the consequences of treatment. N Engl J Med 1988;318:1728–33.
2. Cook R, Sackett D. The number needed to treat: A clinically useful measure of treatment effect. BMJ 1995;310:452–4.
3. Altman DG. Confidence intervals for the number needed to treat. BMJ 1998;317:1309–12.
4. Hutton JL. Misleading statistics: The problems surrounding number needed to treat and number needed to harm. Pharm Med 2010;24:145–9.