

Calculating Risks and Number-Needed-to-Treat: A Method of Data Interpretation

Michael G. Kendrach, Timothy R. Covington,
Michelle W. McCarthy, and M. Catherine Harris

OBJECTIVE:

To define four measures of association—relative risk (RR), absolute risk reduction (ARR), relative risk reduction (RRR), and number-needed-to-treat (NNT), demonstrate the method of calculating these values, and discuss the clinical relevance of each value.

DATA SOURCES:

MEDLINE 1991 to present and bibliographies of pertinent articles.

STUDY SELECTION:

Articles designed to assess statistical data

interpretation and/or discuss measures of association were evaluated.

DATA SYNTHESIS:

The RR, ARR, RRR, and NNT were calculated for the study endpoints. For the primary endpoint (nonfatal MI or death from CHD), RR=0.70; 2.26% (ARR) of the patients receiving pravastatin were spared the event; 30% (RRR) of the baseline risk was removed as a result of pravastatin therapy and 44 patients need to be treated to prevent an adverse event.

CONCLUSION:

RR, ARR, RRR, and NNT can assist the reader in determining the association between treatment and a health outcome and can be calculated from results of studies reporting dichotomous endpoints. These values can assist in the interpretation of clinical trials to determine the magnitude of benefit from therapy.

KEY WORDS: Decision making, Risk, Treatment outcome, Data interpretation, Attitude of health personnel

J Managed Care Pharm 1997; 3:179-83.

Advertisements and other sources of information present the results of clinical trials evaluating drug therapy in a variety of ways, some of which may be misleading or misinterpreted. This especially is true when health care providers are unaware of the different methods of data analysis and presentation and their significance. Clearly, pharmacists are active in decision making regarding drug utilization and should be able to accurately and appropriately assess the validity of the biomedical literature. However, studies suggest that this isn't always the case. Furthermore, health care practitioners' willingness to select a therapy option may be influenced by the manner in which trial results were presented.¹⁻⁵

The results of clinical trials are key to managed care pharmacists and other practitioners in determining the efficacy, side effect profile, and other aspects of a drug product. When this data is misleading or misinterpreted, the repercussions may be serious and widespread. Thus, health care practitioners should be well versed in the inter-

pretation of clinical trial results prior to recommending therapy.

One technique to ascertain the clinical impact of drug therapy is to examine four measures of association: relative risk (RR), absolute risk reduction (ARR), relative risk reduction (RRR), and number-needed-to-treat (NNT). These values can assist health care practitioners in determining the association between treatment and outcome. This article will define these measurements, demonstrate the method of calculating these values, and discuss the relevance of each value.

STUDY SELECTION

The Pravastatin Primary Prevention (3P) Trial was selected as a model to demonstrate how to calculate the measures of association using endpoint adverse event data. This study was designed to evaluate the efficacy of pravastatin, an HMG-CoA reductase inhibitor, in preventing coronary



Authors

MICHAEL G. KENDRACH, PHARM.D., is Director, Drug Information Service, and Assistant Professor, McWhorter School of Pharmacy, Samford University, Birmingham, AL. TIMOTHY R. COVINGTON, PHARM.D., M.S., is Executive Director, Managed Care Institute, and Bruno Professor of Pharmacy, McWhorter School of Pharmacy, Samford University. MICHELLE W. MCCARTHY, PHARM.D., is Associate Director, Drug Information Service, and Assistant Professor, McWhorter School of Pharmacy, Samford University. M. CATHERINE HARRIS, PHARM.D., is Drug Information Specialist, University of Alabama at Birmingham Drug Resource Center, Birmingham, AL.

AUTHOR CORRESPONDENCE: Michael Kendrach, McWhorter School of Pharmacy, Samford University, 800 Lakeside Drive, Birmingham, AL 35229.

Copyright © 1997, Academy of Managed Care Pharmacy, Inc. All rights reserved.

Table 1. Selected Baseline Characteristics of Randomized Patients

Variable	Placebo (N=3,293)	Pravastatin (N=3,302)
Age (years)*	55.1 ± 5.5	55.3 ± 5.5
Total cholesterol*†	272 ± 22	272 ± 23
Low-density lipoprotein (LDL)*†	192 ± 17	191 ± 17
High-density lipoprotein (LDL)*†	164 ± 68	162 ± 70
Triglycerides*†	164 ± 68	162 ± 70
Ex-smoker (%)	1,127 (34)	1,138 (34)
Current smoker (%)	1,460 (44)	1,445 (44)

*Values expressed as mean ± standard deviation.

†Units equal mg/dL

events in men with moderate hypercholesterolemia (see Table 1) and no history of myocardial infarction (MI). A total of 6,595 patients were randomized to either pravastatin 40 mg/day (N = 3,302) or placebo (N = 3,293) and followed for a total of 32,216 subject-years (an average of 4.9 years/subject). The primary endpoint of the study was the occurrence of nonfatal MI or death from coronary heart disease (CHD) as a first event; these two categories were combined. In addition, the effect of treatment on death from any cause, the incidence of nonfatal MI, and the frequency of coronary revascularization procedures were analyzed. The baseline characteristics of the patients enrolled were similar (see Table 1).⁶ Further details of the study design are described elsewhere.^{6,7}

Table 2. Format of Outcome Data According to Therapy

	Outcome*	
	Yes	No
Pravastatin	A	B
Placebo	C	D

* Example: Pravastatin Primary Preventive Trial Primary Endpoint

Since dichotomous measurements ("yes" or "no", but not both) were collected as endpoints, these figures are used to calculate the measures of association. Table 2 shows the format in which the outcomes were arranged to perform the calculations. The 3P Trial endpoint data were entered into this table by adverse event and patient group to calculate the additional measures of association not provided by the study authors.

For instance, 174 pravastatin-treated patients and 248 placebo-treated patients experienced the primary endpoint, while 3,128 pravastatin and 3,045 placebo-treated patients did not. Thus, the corresponding cell letter for each group would be A, C, B, and D, respectively (see Table 3).

Once the outcome data was entered into Table 3, the measures of association were calculated. Relative risk (RR) is

the ratio of risk of an event occurring in one group compared to another group. This value indicates the risk of the event after the experimental treatment as a percentage of the original risk.^{8,9} The interpretation of this calculation focuses around the whole number one (1.0). A value of < 1.0 indicates that therapy lessened the risk of developing the adverse outcome in the treatment group compared to placebo. A RR value = 1.0 denotes no difference between treatments. However, a RR value of > 1.0 indicates that therapy increased the risk of developing the adverse outcome. The formula for calculating RR from the gathered data is $[A / (A + B)] / [C / (C + D)]$.^{8,9}

Table 3. Number of Patients with the Primary Endpoint Result

	Pravastatin Primary Prevention Trial Primary Endpoint	
	Yes	No
Pravastatin	174	3,128
Placebo	248	3,045

Absolute risk reduction (ARR) is defined as the difference in the risk of the outcome between patients who have received one therapy from those who have received another. This measure provides the percentage of patients spared the adverse outcome as a result of receiving the experimental rather than the control therapy, and changes with a change in baseline risk. The formula for calculating ARR is $[C / (C + D)] - [A / (A + B)]$.^{8,9}

Relative risk reduction (RRR), a term used in clinical studies and medication advertisements, estimates the percentage of baseline risk that is removed as a result of therapy. This measure is used to compare the efficacy of treatment to the "control." Two methods of calculating this measurement are as follows: $\{[C / (C + D)] - [A / (A + B)]\} / [C / (C + D)]$ or simply, $1 - RR$.^{8,9}

Number-needed-to-treat (NNT), indicates the number of patients who require treatment to prevent one event, and can provide valuable insight into the benefits of therapy. NNT can be calculated by taking the reciprocal of the ARR ($1 / ARR$).^{8,9}

DATA SYNTHESIS

Calculation examples of RR, ARR, RRR, and NNT for the primary endpoint of the 3P Trial are:

▲ Relative Risk: $[174 / (174 + 3,128)] / [248 / (248 + 3,045)] = 0.70$

▲ Absolute Risk Reduction: $[248 / (248 + 3,045)] - [174 / (174 + 3,128)] = 2.26\%$

▲ Relative Risk Reduction: $\{[248 / (248 + 3,045)] - [174 / (174 + 3,128)]\} / [248 / (248 + 3,045)] = 0.30$
or $1 - 0.70 = 0.30$

▲ Number-Needed-to-Treat: $1 / 0.0226 = 44$

Table 4. Relative Risk (RR), Absolute Risk Reduction (ARR), Relative Risk Reduction (RRR), and Number-Needed-to-Treat (NNT) According to Study Endpoint

Study Endpoint	RR	ARR (%)	RRR (%)	NNT
Nonfatal MI* or Death from CHD*	0.70	2.26	30	44
Nonfatal MI (silent MIs omitted) or Death from CHD	0.69	2.08	31	48
Nonfatal MI	0.70	1.86	30	54
Death from CHD	0.73	0.42	27	233
Coronary Angiography	0.70	1.16	30	86
PTCA† or CABG†	0.64	0.88	36	113
Death from all Cardiovascular Causes	0.68	0.70	32	142
Death from Any Cause	0.78	0.89	22	112

*MI = myocardial infarction; CHD = coronary heart disease

†PTCA = percutaneous transluminal coronary angioplasty; CABG = coronary artery bypass graft

Table 4 shows the respective values for RR, ARR, RRR, and NNT according to measured adverse event endpoints of the 3P Trial.

Using the primary endpoint to explain these values in a text format, patients randomized to the pravastatin group benefited compared to placebo-treated patients (indicated by the calculated measures of association). The risk of death or nonfatal MI in the pravastatin group is approximately two-thirds of the original risk (RR = 0.70). In addition, 75 pravastatin patients were spared a nonfatal MI or death from CHD compared to placebo (ARR = 2.26%). Furthermore, approximately one-third of the baseline risk was removed as a result of pravastatin therapy (RRR = 30%), and 44 patients need to be treated for an average of 4.9 years to prevent one nonfatal MI or death from CHD (NNT = 44).

DISCUSSION

Measures of association (RR, ARR, RRR, NNT) can be calculated from studies designed to measure dichotomous outcomes as endpoints. Since most studies in general discuss only the relative changes (RR and RRR), additional measures can be calculated to more fully evaluate the clinical significance of the results. The definitions and implications of these measures provide a better understanding of the effect of medical intervention.

The Influx and Impact of Aggressive Advertising

As the economics of health care continue to be a major consideration in the delivery of patient care, payers may not be willing to pay for services and/or goods at a premium price

when clinically equivalent, lower-priced alternatives are available. In response, some manufacturers compensate through more aggressive marketing techniques to maintain or increase market share. However, the medical literature has expressed some concern about pharmaceutical advertisements that present "misleading" information. In fact, one study by expert reviewers demonstrated that many pharmaceutical advertisements contained deficiencies in areas where the Food and Drug Administration (FDA) has established explicit standards of quality.¹⁰ This is of particular concern when one considers the data showing that advertisements influence clinical decisions and physician prescribing.¹¹⁻¹³ Clearly, health care professionals need to be aware of the limitations of advertisements and promotional materials. But they also must be prepared to analyze them thoroughly and accurately.

Deceptive Nature of Numbers

One method of enhancing promotional materials is to display relative differences or advantages. However, "benefits" presented as relative changes can be misinterpreted due to an absent or hidden (e.g., in small-print) baseline value. For instance, RR and RRR do not reflect the baseline risk in the control population.

A simple example from everyday life illustrates the deceptive nature of relative numbers. A long-distance phone company encourages customers to sign up for their plan by claiming calls will be connected 42% faster than other long-distance companies. Based on the television commercials and direct mail materials, the offer seems attractive. However, an analysis of the data may show otherwise.

First, the consumers should determine what is the baseline time and magnitude of difference for the first company to connect the call versus the second company. If the first company takes 3.5 seconds and the second company takes six seconds, the 42% difference may not be significant. However, if the first company takes 35 minutes and the second company takes 60 minutes (still a 42% difference), there is a significant time disparity. And the cost differences could be significant. This example is comparable to some situations in medical decision making. However, the clinical implications of the risk reductions (3% to 1% versus 60% to 20%) are very different.

Consider a therapy with a 5% incidence of severe side effects that also reduces the probability of an adverse outcome from 3% to 1% (RRR = 66%). This therapy may not be desirable due to the risk-benefit ratio. However, if the therapy reduces the probability of an adverse outcome from 60% to 20% (also a RRR of 66%), it may be acceptable since, out of 100 treated patients, 40 patients would benefit and only five would experience side effects.⁸

Presentation of Findings Can Influence Decisions

Results of several studies have indicated that the manner in which research findings are presented can influence decisions. For example, physicians have been surveyed to determine their willingness to prescribe drug therapy according to vari-

ous measures of association. Although each therapy "option" was obtained from one study (but presented as ARR, NNT, or RRR), the RRR was most favorable and selected most often.^{1,4}

Naylor et al.¹ and Bucher et al.² conducted separate, but similarly designed, studies to determine physician perception of the effectiveness of drug therapy. In the Naylor study, house staff and faculty (internal medicine and subspecialties) were included; the Bucher study focused on primary care physicians (internists and general practitioners). The physicians randomly received one of two separate questionnaires, each containing four statements enumerating the effects of therapy for three endpoint measures (fatal plus non-fatal MI, fatal MI, and total mortality) of the Helsinki Heart study.¹⁴ The first three statements were expressed as either RRR (form "A") or ARR (form "B"). The last statement on both forms was the same and expressed the effect of therapy in terms of NNT for fatal plus non-fatal MI. The physicians were asked to grade each therapy according to their perception of therapy efficacy using an 11-point scale (-5 to +5). The far left of the scale (-5) was labeled "therapy is harmful," while the opposite side of the scale (+5) was labeled "therapy is beneficial"; a zero value represented "no effect." Results demonstrated that ratings of therapeutic effectiveness were higher for the questionnaire displaying the results as relative reductions than absolute reductions ($p < 0.001$). Furthermore, physicians rated the effectiveness of therapy significantly higher for the endpoint represented as risk reduction (either RRR or ARR) than NNT ($p < 0.001$).

In addition, Forrow et al.³ conducted a study to determine if differences in presentation of research results (emphasizing relative or absolute change in outcome rates) may lead to differences in perceived treatment benefits. Physicians attending educational conferences at either a selected teaching or community hospital were surveyed; selected fellows and faculty received a questionnaire via mail. Two different questionnaires (one for hypertension and one for hypercholesterolemia) were developed, each having six brief statements summarizing information derived from published studies. Following each statement, the participants were asked how this information would affect their decision to treat their patients with this condition. Two of the six questions on each questionnaire reported the outcome of the same study in a different manner, either as RRR or ARR. Physicians were asked to rate the likelihood of starting treatment based on each statement using a seven-point scale (ranging from "definitely more likely to treat" to "definitely less likely to treat"). Results demonstrated that 46% of the physicians responded differently to the same question presented in the two different manners; of these, approximately 90% indicated a stronger desire to treat patients according to the relative, versus absolute, change in the outcome rate ($p < 0.0001$).

Bobbio et al.⁴ also surveyed general practitioners to determine whether reporting outcomes as various measures of association affected physician opinions on the treatments' usefulness. The questionnaire presented five different drug "thera-

pies," each expressed as either RRR, ARR, NNT, percentage of event-free patients, or RRR plus relative increase in total mortality. Although each drug "therapy" appeared to be a different drug regimen, all were derived from the same outcome (change in cardiac event incidence) of the Helsinki Heart Study.¹⁴ For each therapy, physicians were asked to place an "X" along a 10-cm visual analog scale labeled "I would definitely not prescribe this drug" on the left and "I would definitely prescribe this drug" on the right. The highest response rate for agreement to prescribe the drug was for RRR (78%) while the next highest response (37%) was percentage of event-free patients ($p < 0.001$).

A survey of authorities responsible for establishing health care policies⁵ produced similar results. Participants were presented with two disease state scenarios: breast cancer screening and cardiac rehabilitation. Following these scenarios, four disease prevention "programs" described the effectiveness of each one. The outcome of each "program" was presented as either RRR, ARR, proportion of event-free patients, or NNT. Afterwards, the individuals were asked to rate their willingness to implement the "program" by placing an "X" along a linear scale marked from 0 ("I would not support purchasing of this service") to 10 ("I would strongly support the purchasing of this service"). The highest mean score (95% confidence interval) for both "programs" was for RRR: 79% (76-83%) and 76% (72-80%), respectively ($p < 0.05$). The next highest mean score was for NNT, 51% (47-55%) and 62% (58-66%). The results indicate that the method of reporting trial results influenced the policy decisions.

Influencing Medical Therapy and Benefits

Survey results of pharmacists and patients have demonstrated that views of medical therapy can be influenced by how potential benefits are presented. Papay et al.¹⁵ designed a questionnaire to determine if pharmacists were able to recognize different presentation styles of the same study result. The 3P Trial⁶ primary endpoint was expressed as RRR, ARR, ARR represented as "rate decreased from 7.9 to 5.5%", or NNT as though each of these were four different "therapies." After reading a brief clinical scenario, pharmacists were asked to select one of the four "therapies" that they would most likely recommend to the medical staff. Preliminary results indicated that only 6.5% of the responders were able to identify that all four "therapies" were equivalent; the highest number of responses (39%) was for ARR represented as "rate decreased from 7.9 to 5.5%."

Hux et al.¹⁶ distributed a questionnaire to patients describing the effects of the same medication displayed in three different formats: RRR, ARR, and NNT. Eighty-eight percent of the patients were willing to take the medication when presented as RRR, while the next highest response (56%) was NNT ($p < 0.0001$). Malenka et al.¹⁷ reported similar results—when asked to select a therapy which appeared more beneficial, patients selected the therapy results displayed as RRR most often (56.8%), compared to ARR (14.7%).

Interpreting ARR and NNT

Very few studies include ARR, while even fewer present NNT. Both values incorporate the influence of any change in baseline risk. However, neither of these two measures provide the magnitude of baseline risk. Since NNT is related to the ARR, a change in the underlying risk will cause an alteration in NNT. Furthermore, NNT is directly related to the proportion of patients in the control group that suffer an adverse event. Thus, if the risk of an adverse event doubles, only half as many patients need to be treated to prevent the same number of adverse events, and vice versa.⁸

A preliminary method to assist in the interpretation of NNT has been proposed by calculating a threshold NNT (T-NNT). In simple terms, the method for generating a T-NNT focuses on the concept that the net cost of treating the number of patients to prevent one adverse event equals the net value of the adverse events prevented by treating that number of patients.¹⁸ As with most therapies, both positive and negative consequences need to be considered in the decision to treat or not to treat. For instance, patients with elevated cholesterol can develop cardiac morbidity and/or mortality if no therapy is initiated.¹⁹ If therapy is initiated, the disease could regress and the patient's risk of cardiac morbidity and/or mortality may be reduced.²⁰ However, the side effects and costs of treatment should be considered. Generating a T-NNT includes identifying the consequences of treating patients (e.g., side effects) versus not treating (e.g., adverse event). In addition, the associated costs (costs of treating, costs saved by preventing the adverse event, and costs of treating side effects caused by the treatment) are incorporated into the calculations.¹⁸ Although the calculation of the T-NNT is beyond the scope of this article, if a patient's risk of an adverse event without therapy is high and the calculated NNT is below the T-NNT, therapy should be considered, and vice versa.¹⁸

Otherwise, the decision to start therapy can be based on the patient's risk of the adverse event if left untreated versus the risk of harm from therapy (e.g., side effects) and/or costs.⁹ For example, if pravastatin produces myalgia in 3% of treated patients, the NNT to cause myalgia is 1/0.03 or 33. Thus, for every nonfatal MI or death from CHD prevented, 1.3 patients can experience myalgia. On the other hand, treating 233 patients to prevent one death due to CHD (see Table 4) can result in seven individuals developing myalgia. Despite this simple example, this approach cannot be applied to severe, episodic events. The number of adverse events per life saved (or, if the events are rare enough, the number of lives saved per adverse event) can provide a compelling picture of the trade-offs associated with the clinical intervention.⁹

CONCLUSION

Health care professionals need to remind themselves of the "limitations" of the medical literature and not be misled by the presentation styles. As demonstrated by the measures of association formulas, these values can be easily calculated to provide additional data that can be used to evaluate and interpret the

biomedical literature and assist in the decision making process.

Those in managed care, including managed care pharmacists, should exercise caution in evaluating review literature, studies, and other information and materials that they use in developing drug benefit programs, formularies, and other components of patient care. ■



References

- Naylor CD, Chen E, Strauss B. Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? *Ann Intern Med* 1992; 117: 916-21.
- Bucher HC, Weinbacher M, Gyr K. Influence of method of reporting study results on decision of physicians to prescribe drugs to lower cholesterol concentration. *BMJ* 1994; 309: 761-4.
- Forrow L, Taylor WC, Arnold RM. Absolutely relative: how research results are summarized can affect treatment decisions. *Am J Med* 1992; 92: 121-4.
- Bobbio M, Demichelis B, Giustetto G. Completeness of reporting trial results: effect on physicians' willingness to prescribe. *Lancet* 1994; 343: 1209-11.
- Fahey T, Griffiths S, Peters TJ. Evidence based purchasing: understanding results of clinical trials and systematic reviews. *BMJ* 1995; 311: 1056-60.
- Shepherd J, Cobbe SM, Ford I, Isles CG, Lorimer AR, Macfarlane PW, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. *N Engl J Med* 1995; 333: 1301-7.
- The West of Scotland Coronary Prevention Study Group. A coronary primary prevention study of Scottish men aged 45-64 years: trial design. *J Clin Epidemiol* 1992; 45: 849-60.
- Jaeschke R, Guyatt G, Shannon H, Walter S, Cook D, Heddle N. Basic statistics for clinicians. 3. Assessing the effects of treatment: measures of association. *Can Med Assoc J* 1995; 152: 351-7.
- Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Working Group. User's guides to the medical literature, II. How to use an article about therapy or prevention. B: what are the results and will they help me in caring for my patients? *JAMA* 1994; 271: 59-63.
- Wilkes MS, Doblin BH, Shapiro ME. Pharmaceutical advertisements in leading medical journals: experts' assessments. *Ann Intern Med* 1992; 116: 912-19.
- Avorn J, Chen M, Hartley R. Scientific versus commercial sources of influence on the prescribing behaviors of physicians. *Am J Med* 1982; 73: 4-8.
- McGavock H, Webb CH, Johnson GD, Milligan E. Market penetration of new drugs in one United Kingdom region: implications for general practitioners and administrators. *BMJ* 1993; 307: 118-20.
- Ferner RE, Scott DK. Whatalotwegot—the messages in drug advertisements. *BMJ* 1994; 309: 1734-6.
- Frick MH, Elo O, Haapa K, Heinonen OP, Heinsalmi P, Helo P, et al. Helsinki Heart Study: primary-prevention trial with gemfibrozil in middle-aged men with dyslipidemia: safety of treatment, changes in risk factors, and incidence of coronary heart disease. *N Engl J Med* 1987; 317: 1237-45.
- Papay JI, Kendrach MG, McCarthy MW, Fancher MG, Henderson RP, Woolley T. The influence of different reporting methods on pharmacists' therapeutic recommendations (Abstract). 1997 Spring Practice and Research Forum, American College of Clinical Pharmacists.
- Hux JE, Naylor CD. Communicating the benefits of chronic preventative therapy: does the format of efficacy data determine patients' acceptance of treatment? *Med Decision Making* 1995; 15: 152-7.
- Malenka DJ, Baron JA, Johansen S, Wahrenberer JW, Ross JM. The framing effect of relative and absolute risk. *J Gen Intern Med* 1993; 8: 543-8.
- Guyatt GH, Sackett DL, Sinclair JC, for the Evidence-Based Working Group. User's guides to the medical literature, IX. A method for grading health care recommendations. *JAMA* 1995; 274: 1800-4.
- Anonymous. Consensus conference. Lowering blood cholesterol to prevent heart disease. *JAMA* 1985; 253: 2080-6.
- Law MR, Wald NJ, Thompson SG. By how much and how quickly does reduction in serum cholesterol concentration lower risk of ischaemic heart disease? *BMJ* 1994; 308: 367-72.