Nurse Educator

Nurse Educator
Vol. 32, No. 1, pp. 16-20

How to Read, Interpret, and Understand Evidence-Based Literature Statistics

Dorette Sugg Welk, PhD, RN

To implement best practices through research utilization, nurses need to read, interpret, and understand literature formatted with evidence-based practice language and statistics. Hypothetical examples highlight 7 terms and their formulas. A scenario to personally calculate such evidence-based practice statistics can be used to enhance personal effectiveness and to teach others.

Consider the following hypothetical scenario and results that are formatted in evidence-based practice (EBP) language such as those that you might see in the Cochrane Collaboration,[1] a primary resource for evidence-based systematic reviews.

Scenario: How effective is a daily dose of 500 mg of vitamin C in the prevention of ulcers on the heels of bedridden elderly clients? Results: With an NNT of 5, vitamin C is effective (OR, 0.10; 95% CI, 0.05-0.20).

If there are some abbreviations or values in this situation that are unfamiliar to you, you will find explanations and examples in this article that will help you in reading, interpreting, and understanding them as you use evidence-based literature for your best practices. Nurses always rise to the occasion to learn the latest research information that may improve patient care and outcomes.

An obstacle to involvement in EBP is lack of skill in understanding the "bottom line" of systematic reviews and accompanying risk-related numbers.[2-4] Content and research experts conduct systematic reviews using strict criteria for inclusion of primary research studies and statistical analysis.[5] The Cochrane Collaboration is a major resource for more than 1,000 systematic reviews of randomized clinical trials for the effects of healthcare interventions created through collaboration of more than 50 worldwide review and methods teams.[6]

The systematic review teams basically seek the response to 1 question: how many people have a bad outcome in the experimental group compared with the control group? Bad outcomes refer to the undesirable outcomes in a study, such as development of a heel ulcer. Noteworthy in EBP statistics is the simplicity of using a head count rather than group averages. Even when individual study results are not statistically significant, if the experimental group has fewer bad outcomes than the control group, the nurse or other provider might want to apply the results anyway. Seven terms and their abbreviations and formulas are common in the reported results,[7,8] as summarized in Table 1. In this article, hypothetical examples and their derivations describe these 7 terms. At the end, you can derive these values for a clinical scenario toward a better understanding when teaching these terms to others.

Author Affiliations: Nursing Faculty Online, University of Phoenix, Phoenix, Ariz; Faculty Emeritus, Bloomsburg University, Bloomsburg, Pa.
Correspondence: 94 Hemlock Lane, Bloomsburg, PA 17815 (welk@email.phoenix.edu).

Absolute Risk Reduction

Absolute risk reduction (ARR) is the absolute arithmetic difference (absolute means that one ignores plus and minus signs) in percentages of bad outcomes between the experimental and control groups. Absolute risk reduction means that more people in the control group than in the experimental group develop a bad outcome. To calculate the ARR, you need to know just 2 things: the experimental event rate (EER), or the percentage of the bad outcome in the experimental group; and the control group event rate (CER), or the percentage of the bad outcome in the control group. Let us look at an example: 13% of patients with diabetes receiving intensive insulin therapy (the experimental group) develop neuropathy (the bad outcome), whereas 38% of the patients in the control group receiving usual insulin therapy develop neuropathy. ARR = |EER − CER| or |13% − 38%| = 25%. Therefore, the ARR of giving patients with diabetes intensive insulin therapy in terms of development of the bad outcome of neuropathy is 0.25, or 25%. By simply having a head count of the incidence of occurrence of an outcome when you know the number of the total involved, there is some tangible evidence that a practitioner might consider using in the development of practice guidelines.

Number Needed to Treat

Number needed to treat (NNT) is the number of patients who need to be treated to obtain 1 additional good outcome for a patient that would not have occurred without the treatment. The best NNT would be 1, where every patient in the experimental group treated had a good outcome while no one in the control group had a good outcome. However, this NNT is not very realistic, and NNTs of 2 to 5 are indicative of effective treatments. The formula for NNT is 1/ARR. For the insulin therapy example, the NNT is 1/0.25, or 4. This NNT means that the practitioner would need to treat 4

Table 1. 7 Evidence-Based Practice Terms, Abbreviations, and Formulas

Term and Definition	Abbreviation	Formula
Absolute risk reduction: the absolute arithmetic difference (ignore plus and minus signs) in rates of bad outcomes between the experimental and control groups	ARR	|EER − CER|
Number needed to treat: the number of patients who need to be treated to obtain 1 additional good outcome for a patient that would not have occurred without the treatment	NNT	1/ARR
Absolute risk increase: occurs when more patients in the experimental group develop the bad outcome compared with the control group	ARI	CER − EER
Number needed to harm: the number of patients that, if they were given the experimental treatment, would result in 1 additional patient being harmed compared with patients in the control group	NNH	1/ARI
95% Confidence interval: the interval of values between which one is 95% sure that the true value of the population lies	95% CI	±1.96 × the square root of [CER × (1 − CER)]/*N* of the control group + [EER × (1 − EER)]/*N* of the experimental group
Risk or relative ratio: the ratio of risk in the experimental group to the risk in the control group	RR	EER/CER
Odds ratio: the ratio of odds in favor of having the bad outcome in the experimental group to the odds in favor of having the bad outcome in the control group	OR	For EER, bad outcomes/good outcomes; for CER, bad outcomes/good outcomes. Calculate the ratio of the above 2 answers

EER indicates experimental event rate; CER, control group event rate.

patients with intensive insulin therapy to achieve 1 additional patient with a positive outcome (does not develop neuropathy).

Evidence-based practice language is more about risk rather than solid numbers. If a practitioner had a large diabetic population, an NNT of 4 would be good news, and the practitioner might choose to use the intensive therapy with more patients. When reading evidence-based reports, the lower the NNT, the lower the risk based on the review team's synthesis of numerous research studies involving, for example, the intensive versus routine insulin therapy. In studies, for example, of disease prevention or the use of educational strategies, an NNT of 20 or more might be good news for the practitioner because the preciseness of a "bad outcome" is less controllable than in direct disease situations. Context for interpretation is always an important consideration.

Absolute Risk Increase

The absolute risk increase (ARI) is the opposite of risk reduction and occurs when more patients in the treatment group compared with the control group develop a bad outcome. In the insulin therapy example, ARI would mean that more people in the intensive insulin therapy group developed neuropathy (the bad outcome) than those in the control group. The formula for ARI is |CER − EER|, and application of the formula follows in the example below, where it is used in determining the number needed to harm (NNH).

Number Needed to Harm

Number needed to harm is the number of patients that, if given the experimental treatment, would result in the harm of 1 additional patient compared with patients in the control group. "Harm" sounds rather alarming, but it relates to an outcome as simple as a sore throat or the ultimate negative outcome of a patient death despite a treatment anticipated improving health.

Let us look at an example: adults in the experimental group (n = 100) receive vitamin C tablets to determine if the vitamin C reduces work absences related to common cold symptoms; the control group (n = 100) receives placebo candy only. The "bad outcomes" of sore throat and nausea occur in 90% of the adults in the experimental group who develop at least 1 adverse effect and in 80% of adults in the control group who develop at least 1 adverse effect.

Because the bad outcome occurred more often in the experimental group than in the control group, the ARI calculation applies: |CER − EER| = |80% − 90%| = 0.10, or 10%. There is a 10% increased risk of a sore throat or nausea when taking the vitamin C tablets. The formula for NNH is 1/ARI, or 1/0.10 = 10. In other words, the practitioner needs to treat 10 adults with vitamin C tablets to lead to 1 additional person who develops adverse symptoms. If the researcher found that overall work absences decreased, the researcher might construe that this NNH of 10 is worth the risk of a sore throat or nausea; if work absences were not different between the groups, the researcher may not recommend an NNH of 10 as acceptable in practice. The practitioner interprets both NNT and NNH in the context of the actual practice setting.

95% Confidence Intervals

A 95% confidence interval (CI) is an attempt to quantify the uncertainty in measurement. In the vitamin C study, it is possible that nausea incidents were related to a gastrointestinal virus and

not the vitamin C. A 95% CI is calculated to give the interval of values between which one is 95% sure that the true value of the population lies. In a repeat of the study, the 95% CI provides the range in which the researcher would expect the results to fall.[9] Calculation requires that one knows the EER, the CER, and the number of persons in each group. The formula[8] is ±1.96 × the square root of [CER × (1 − CER)]/*N* of the control group + [EER × (1 − EER)]/*N* of the experimental group. Online calculators quickly assist in determining CIs and ranges for NNH and NNT.[10] Many calculators show an NNT label, but if the experimental group exhibits more bad outcomes, the results present as NNH accordingly. In the vitamin C example, the 95% CI is 0.20% to 19.80%. The 95% CI range for NNH is 1/0.20 to 1/19.8, or 5 to 505.

The wide range of 5 to 505 occurs because both groups developed an adverse effect at a fairly equal rate (80% in the control group compared with 90% in the experimental group). Such a wide range for NNH makes it difficult to feel "confident" about the effectiveness of an experimental treatment. If the experimental group responses had been 80% for an adverse effect in the vitamin C group compared with only 40% in the placebo group, the 95% CI would have been much narrower, and an NNH of 10 would seem more useful to the practitioner in deciding about using the tablets. It is important to look at the 95% CI value for any EBP statistic when considering the credibility of the evidence it further defines.

Risk or Relative Ratio

The risk or relative ratio (RR) is the ratio of the percentage of the bad outcome in the experimental group to the percentage of the bad outcome in the control group. The use of ratios contributes important information to EBP because it takes into consideration that some bad outcomes occur in both groups where considering these proportions adds to the evidence. The formula is RR = EER/CER.

What are the odds of patients with diabetes older than 65 years developing heel ulcers when padded devices are applied for their prevention compared with those not using padded devices? The sample has 100 patients, with 50 in each group. The numbers are designed to reflect some ease in calculation and not any true incidence of the outcomes.

	Developed Heel Ulcers	Did Not Develop Heel Ulcers	
Used padded devices (Experimental)	10 (A)	40 (B)	50 (C)
Did not use padded devices (Control)	15 (D)	35 (E)	50 (F)
	25	75	100

To develop the odds ratios, complete these steps:

1. What are the odds of the experimental group developing heel ulcers? (A/B)
2. What are the odds of the control group developing heel ulcers? (D/E)
3. What are the odds of patients with diabetes older than 65 years developing heel ulcers in patients with diabetes when padded devices are applied for their prevention compared with those not using padded devices? (A/B)/(C/D)

In addition, consider these calculations:

4. What is the experimental event rate (EER)? (A/C)
5. What is the control event rate (CER)? (D/F)
6. What is the absolute risk reduction when wearing padded devices? (ARR = |EER-CER|= |(A)/(C)–(D)/(F)|
7. What is the number needed to treat (NNT) with padded devices for 1 additional patient to have a good outcome (not developing heel ulcers)? (NNT = 1/ARR)
8. What is the risk ratio (RR) of developing a heel ulcer in this study? (RR = EER/CER)

Figure 1. A calculation example for EBP terms.

Table 2. Answers to the Practice Scenario

Question	Formula	Calculation and Interpretation
1. What are the odds of the experimental group developing heel ulcers?	A/B	10/40 = 0.25, or a 25% odds of developing a bad outcome of an ulcer compared with a good outcome (non-ulcer)
2. What are the odds of the control group developing heel ulcers?	D/E	15/35 = 0.43, or a 43% odds of developing a bad outcome of an ulcer compared with a good outcome (non-ulcer)
3. What are the odds of patients with diabetes older than 65 years developing heel ulcers when padded devices are applied for their prevention compared with those not using padded devices?	(A/B)/(C/D)	10/40 divided by 15/35 = 0.25/0.43 = 0.58, or a 58% odds of developing a heel ulcer for patients with diabetes older than 65 years wearing a padded device.
4. What is the experimental event rate (EER)?	A/C	10/50 = 0.20, where the bad outcome (ulcer) occurs 20% of the time in this study
5. What is the control event rate (CER)?	D/F	15/50 = 0.30, where the bad outcome (ulcer) occurs 30% of the time in this study
6. What is the absolute risk reduction (ARR) when wearing padded devices?	(ARR = |EER − CER| = |(A)/(C) − (D)/(F)|	|0.20 − 0.30| = |0.10|, where the use of padded devices reduces the risk of developing a heel ulcer by 10%
7. What is the number needed to treat with padded devices for 1 additional patient to have a good outcome (not developing heel ulcers)?	1/ARR	1/0.10 = 10, which means that 10 patients need to be treated to achieve 1 additional positive outcome where no ulcer occurs with the use of padded devices
8. What is the risk ratio of developing a heel ulcer in this study?	EER/CER	0.2/0.3 = 0.667 = 0.67, or a 67% risk of developing a heel ulcer when using a device compared with the control group, who did not use any device

Note: Because the 2 groups in this hypothetical study were fairly close in their incidence of the bad outcome (10 compared with 15), we are not surprised to see that the risk of heel ulcer development is fairly high. Situations in which the control group has a much greater incidence of a bad outcome compared with the experimental group will result in a much reduced RR.

Let us look at an example in the intensive versus regular insulin therapy study with neuropathy as the bad outcome: 13% of the patients in the experimental group developed neuropathy compared with 38% in the control group. Therefore, RR = 0.13/0.38 = 0.34, or 34%. This RR means that the person receiving intensive insulin therapy still has a 34% chance (34 times out of 100) or risk of developing neuropathy when compared with the control group. The fact that only 13% did develop the neuropathy with intensive insulin therapy (the ARR) does not give the practitioner the same evidence as also considering the higher incidence in the control group and its likeliness in the population. When discussing possible treatment options, it would be the 34% relative risk value that the practitioner would share with the patient.

Odds Ratio

An odds ratio (OR) is equivalent to saying, "taking all things into consideration…." Odds ratio is the ratio of the odds in favor of having the bad outcome in the experimental group to the odds in favor of having the bad outcome in the control group. There are 3 steps: determine the experimental group odds (bad outcome occurs/good outcome occurs), determine the control group odds (bad outcome occurs/good outcome occurs), and calculate the ratio of these 2 ratio values.

Let us look at an example: for simplicity of calculations, let us say that N = 10 and there are 5 patients in each group. The researcher finds that the experimental group odds are 1 bad outcome/4 good outcomes = 0.25. The control group odds are 4 bad outcomes/1 good outcome = 4. The OR = 0.25/4 = 0.0625 = 6.25%. This OR means that the odds of developing a bad outcome are 6.25% when given the experimental treatment. By adding the component of good outcomes, the OR goes beyond the RR and gives the broadest picture. Odds ratios greater than 1 are indicative of a harmful exposure to the experimental group treatment, whereas ORs less than 1 are indicative of a nonharmful exposure (eg, in the latter case, the patients do not develop neuropathy).[9]

Interpreting the Opening Scenario

Let us look again at the scenario presented at the beginning of this article and interpret it.

> *How effective is a daily dose of 500 mg of vitamin C in the prevention of ulcers on the heels of bedridden elderly clients? Results: With an NNT of 5, vitamin C is effective (OR, 0.10; 95% CI, 0.05-0.20).*

The NNT of 5 means that the practitioner would need to treat 5 patients with 500 mg of vitamin C daily for 1 additional patient to have a positive outcome of not developing a heel ulcer. Although we do not directly see the original research studies, we know that the systematic review team found more studies where the vitamin C was effective in the experimental groups. The OR of 0.10 means

that those receiving vitamin C still have a 10% chance of developing heel ulcers. The small range for the 95% CI of 5% to 20% gives the reader good confidence that the OR of 0.10 credibly reflects the true value for the population.

A Scenario to Calculate

Consider the hypothetical scenario and related questions in Figure 1. The grid cells contain hypothetical numbers of incidence of individuals who met the category criteria in the research study. Table 2 contains the answers.

Conclusion

Evidence-based practice statistics are commonly reported in terms of RRs and risk reductions and increases. They come from the simple knowledge of the proportion of individuals who develop a bad outcome in the experimental group compared with those who do so in the control group. A review of EBP terms, abbreviations, and applied examples gives the reader practice skills for research utilization and decisions about best practices. Working through a hypothetical example enhances the reader's personal effectiveness to read, interpret, and apply evidence-based literature in any teaching, mentoring, or practice setting.

References

1. The Cochrane Collection. Available at: http://www.cochrane.org. Accessed March 11, 2006.
2. Gerrish K, Clayton J. Promoting evidence-based practice: an organization approach. *J Nurs Manag*. 2004;12:114-123.
3. Crawford P, Brown B, Anthony P, Hicks C. Reluctant empiricists: community mental health nurses and the art of evidence-based praxis. *Health Soc Care Community*. 2002;10(4):287-298.
4. Melynk BM. Strategies for overcoming barriers in implementing evidence-based practice. *Pediatr Nurs*. 2002;28(2):159-161.
5. Mulrow CD, Cook DJ, Davidoff F. Systematic reviews: critical links in the great chain of evidence. *Ann Intern Med*. 1997;126(5):389-391.
6. Melynk BM, Fineout-Overholt E, Stone P, Ackerman M. Evidence-based practice: the past, the present, and recommendations for the millennium. *Pediatr Nurs*. 2000;29(1):77-80.
7. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine: How to Practice and Teach EBM*. Edinburgh: Churchill Livingstone; 2000.
8. Gibbs LE. *Evidence-Based Practice for the Helping Professions: A Practical Guide With Integrated Multimedia*. Belmont, Calif: Wadsworth Publishing; 2002.
9. NSW Therapeutic Assessment Group, Inc. Evidence-based practice. Part 3: Applying research evidence to decision making. *Tagnet Bulletin*. Available at: http://www.ciap.health.nsw.gov.au/nswtag/publications/newsletter/tagnet_6.ptd. Accessed March 11, 2006.
10. GraftPad Software. QuickCalcs: Online calculators for scientists. Available at: http://graphpad.com/quickcalcs/NNT1.cfm. Accessed March 12, 2006.