

from month 1 after implementation and using the nurse administrator's ESI score as the second assessment, I asked residents to compare high-risk and lower-risk triage scores between the triage nurse and the nurse administrator. The resulting 2×2 table is completed as shown in Fig. 1, and calculation of chance agreement proceeds as follows:

$\text{kappa} = [(\text{observed agreement} - \text{expected agreement}) / (1 - \text{expected agreement})]$

High-risk assessments by nurse administrator: $11/25 = 0.44$

High-risk assessments by triage nurse: $10/25 = 0.40$

Lower-risk assessments by nurse administrator: $14/25 = 0.56$

Lower-risk assessments by triage nurse: $15/25 = 0.60$

Observed agreement = $(9 + 13)/25 = 0.88$

Expected agreement = (chance of high-risk assessment) + (chance of lower-risk assessment)

Chance of high-risk assessment = $0.44 \times 0.40 = 0.176$

Chance of lower-risk assessment = $0.56 \times 0.60 = 0.336$

Expected agreement by chance alone = $0.176 + 0.336 = 0.512$

$\text{kappa} = (0.88 - 0.512) / (1 - 0.512) = 0.368 / 0.488 = 0.75$

Table 1 in both the teachers¹ and learners^{2,4} versions of this article references Maclure and Willett⁵ as a source of the qualitative classification of kappa. My own review of that paper did not reveal any attempt to qualitatively assess kappa, but at least 3 other sources have done so.⁶⁻⁸ In my experience the most widely used classification for kappa is

the last of these,⁸ which proposed the guidelines for interpreting kappa values as outlined in Table 1 in this letter.

Christopher R. Carpenter

Assistant Professor
Division of Emergency Medicine
Washington University School
of Medicine
St. Louis, Mo.

References

1. McGinn T, Wyer PC, Newman TB, Keitz S, Leipzig R, Guyatt G, for Evidence-Based Medicine Teaching Tips Working Group. Tips for teachers of evidence-based medicine: 3. Understanding and calculating kappa. *CMAJ* 2004;171(11):Online-1 to Online-9. Available: www.cmaj.ca/cgi/data/171/11/1369/DC1/1 (accessed 2005 Jan 10).
2. Dawson B, Trapp RG. *Basic and clinical biostatistics*. 3rd ed. New York: McGraw-Hill; 2001. p. 66-7, 115-7.
3. Wuerz RC, Milne LW, Eitel DR, Travers D, Gilboy N. Reliability and validity of a new five-level triage instrument. *Acad Emerg Med* 2000;7:236-42.
4. McGinn T, Wyer PC, Newman TB, Keitz S, Leipzig R, Guyatt G, for Evidence-Based Medicine Teaching Tips Working Group. Tips for learners of evidence-based medicine: 3. Measures of observer variability (kappa statistic). *CMAJ* 2004;171(11):1369-73.
5. Maclure M, Willett WC. Misinterpretation and misuse of the kappa statistic. *Am J Epidemiol* 1987;126:161-9.
6. Altman DG. *Practical statistics for medical students*. London: Chapman and Hall; 1991.
7. Fleiss JL. *Statistical methods for rates and proportions*. 2nd ed. New York: John Wiley and Sons; 1981. p. 218.
8. Byrt T. How good is that agreement? [letter]. *Epidemiology* 1996;7:561.

DOI:10.1503/cmaj.1041742

In their excellent overview of a common statistical measure of agreement, Thomas McGinn and colleagues¹ suggest in Table 1 that values for the kappa statistic range from 0 to 1. However, negative values of kappa are also possible.² Although unusual in practice, a

negative kappa statistic results when agreement occurs less often than predicted by chance alone. This may indicate genuine disagreement, or it may reflect a problem in the application of a diagnostic test. Readers and researchers who encounter a negative kappa statistic should be aware of its implications, rather than blaming mathematical or typographic errors or computer "gnomes."

David N. Juurlink

Sunnybrook and Women's College
Health Sciences Centre
Toronto, Ont.

Allan S. Detsky

Mount Sinai Hospital
Toronto, Ont.

References

1. McGinn T, Wyer PC, Newman TB, Keitz S, Leipzig R, Guyatt G, for Evidence-Based Medicine Teaching Tips Working Group. Tips for learners of evidence-based medicine: 3. Measures of observer variability (kappa statistic). *CMAJ* 2004;171(11):1369-73.
2. Kramer MS, Feinstein AR. Clinical biostatistics. LIV. The biostatistics of concordance. *Clin Pharmacol Ther* 1981;29:111-23.

DOI:10.1503/cmaj.1041744

As a teacher of basic skills in evidence-based medicine (EBM), I have appreciated the *CMAJ* articles that have been appearing in the EBM "tips" series. In particular, I was happy to see the discussion of the kappa statistic,¹ specifically the calculation of chance agreement (e.g., Table 3 in the article).

Unfortunately, discussions of kappa tend to focus on dichotomous variables, such as positive or negative results on mammography or the presence or absence of Murphy's sign. In cases of con-

| | | Triage nurse | | Total |
|---------------------|------------|--------------|------------|-------|
| | | High risk | Lower risk | |
| Nurse administrator | High risk | 9 | 2 | 11 |
| | Lower risk | 1 | 13 | 14 |
| Total | | 10 | 15 | 25 |

Fig. 1: Agreement table for triage nurse and nurse administrator at the author's hospital, using the emergency severity index³ for nursing triage.

Table 1: Qualitative classification of kappa values*

| Kappa value | Degree of agreement |
|-------------|---------------------|
| ≤ 0 | None |
| 0.01–0.20 | Poor |
| 0.21–0.40 | Slight |
| 0.41–0.60 | Fair |
| 0.61–0.80 | Good |
| 0.81–0.92 | Very good |
| 0.93–1.00 | Excellent |

*Adapted, with permission of the publisher, from Byrt T. How good is that agreement? [letter]. *Epidemiology* 1996;7:561.