

#### 4.2 Elbow Extension Test for Elbow Fracture (with thanks to Matt Hickey, Epi 204 2013)

Appelboam et al.[1] studied the elbow extension test (inability fully to extend the elbow) as a predictor of elbow fracture in 960 adult emergency department patients. All 647 patients who had a positive test (were unable to extend fully) received an x-ray (gold standard #1), but only 58 of the 313 patients with a negative test received an x-ray, of whom  $2/58 = 3.5\%$  showed fractures. The remaining 255 received clinical follow up for subsequent elbow problems (gold standard #2); only  $3/255 = 1.2\%$  had problems on follow up.

- a) Of the 647 patients with inability to fully extend the elbow (a positive test), 311 (48.1%) showed an elbow fracture. This 48.1% represents which index (Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value, etc.) of test accuracy?
- b) As above, of the 313 patients who had a negative elbow extension test, 2 had a positive x-ray, and 3 had problems on clinical follow-up and should be interpreted as false negatives. Assuming that x-rays and clinical follow-up always give the same answer, what was the negative predictive value (NPV) of the elbow extension test?
- c) Again, assuming that x-rays and clinical follow-up always give the same answer, create a 2x2 table using the numbers from part b above, and calculate sensitivity and specificity.
- d) Now re-create the 2x2 table in (c) above but assume that the rate of x-ray positivity among those with normal elbow extension who did not receive x-rays would have been the same as among those who did. Under this assumption, 9 of the 255 patients receiving clinical follow-up would have had positive x-rays had all patients in the study received an x-ray as a single gold standard. Combined with the 2 patients with positive x-rays from among the 58 who actually received an x-ray, there would be a total of 11 patients with normal elbow extension and a positive x-ray. Calculate sensitivity, specificity, PPV, and NPV.
- e) Under the assumption of Part (d) that 6 patients with negative index tests and negative clinical follow-up would have had a positive x-ray, how did using a differential gold standard in the actual study affect sensitivity and specificity relative to a study in which all patients received x-rays?
- f) If you were willing to do up to 20 X-rays to find one elbow fracture, would the possibility of differential verification bias significantly affect your decision to trust the elbow extension test based on this study (assuming the observed prior probability is similar to yours)?
- g) Repeat part F, but this time assume you are willing to do 50 X-rays to find one elbow fracture.

1. Appelboam A, Reuben AD, Bengner JR, Beech F, Dutson J, Haig S, et al. Elbow extension test to rule out elbow fracture: multicentre, prospective validation and observational study of diagnostic accuracy in adults and children. *BMJ*. 2008;337:a2428.