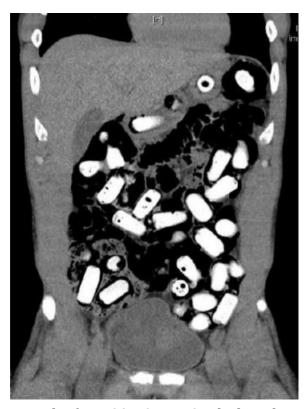5.3. **ED interpretation of CT scans for body packing** (with thanks to Dr. Kimberly Kallianos, Epi 204, 2015)

Individuals suspected of drug smuggling by ingestion of drug packages (known as body packers) may be brought to emergency departments for abdominal computed tomography (CT) scanning.

Sometimes the diagnosis is obvious (Figure), but in other cases emergency department clinicians may sometimes find it challenging to interpret these CT scans if formal radiology interpretation is not available overnight. Missing concealed drug packages has important clinical implications, as the packages may rupture leading to fatal overdose.

Asha et al [1] investigated (among other things) the inter-rater reliability of the ED physicians for whether the CT scan was or was not positive for packing.

The authors reported Kappa = 0.46 (95% CI 0.30−0.62, P<0.001). For parts a to c, which of the following statements about that Kappa are true? Explain your answers.

a. The Kappa of 0.46 indicates agreement was worse than would be expected by chance alone, since by chance we would expect ~50% agreement.

Example of a positive CT scan in a body packer

False. Whether or not we would expect 50% agreement by chance depends on whether we are willing to assume the marginals are fixed, but either way a Kappa >0 indicates better agreement than expected.

b. If ED raters agreed that the approximate prevalence of packing on CT scans was only about 25%, then we would expect them to agree > 50% of the time, even if they did not know anything about how to read CT scans.

True. As an example, if they read 100 CT scans, the marginals of the 2×2 table would be as shown below and expected agreement would be $(25 \times 25/100 + 75 \times 75/100)/100 = 0.625$.

Observer 2	Observer 1		
	+	-	Total

+			**25**
-			**75**
Total	**25**	**75**	**100**

c. The authors of this study could have obtained a higher Kappa value (without at all changing their study or their data) simply by calculating a quadratic-weighted Kappa.

False. Using quadratic-weighting will generally inflate Kappa, but that option is only available when there are > 2 ordered categories.

d. If you look at the figure, it's hard to believe Kappa was only 0.46. Why do you think K was not higher?

Presumably, most of the time the packing is much subtler. This dramatically illustrates that the results of a study of Kappa will depend on the *spectrum* of abnormality in the sample of patients evaluated.

1.   Asha SE, Cooke A. Sensitivity and Specificity of Emergency Physicians and Trainees for Identifying Internally Concealed Drug Packages on Abdominal Computed Tomography Scan: Do Lung Windows Improve Accuracy? J Emerg Med. 2015;49(3):268-73.