Supplemental Material:

1. From biomarker levels in cases and controls to risk stratification: The algebra

The basic algebra of screening is simple. Understanding the terminology is the only pre-
requisite for understanding the math, except for the delta formula.

The 2 x 2 table from a cross-sectional study is the simplest setting that can demonstrate the
complexity. In a 2 x 2 table disease status is in columns and biomarker status – whether the
biomarker is above a fixed threshold, called positive at the level that is believed to be positively
associated with disease – is in rows.

<table>
<thead>
<tr>
<th></th>
<th>Case (Disease)</th>
<th>Control (Non-disease)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomarker above</td>
<td>A = (1 - β)π</td>
<td>C = α (1 - π)</td>
<td>A + C = (1 - β)π + α (1 - π)</td>
</tr>
<tr>
<td>threshold</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomarker below</td>
<td>B = βπ</td>
<td>D = (1 - α)(1 - π)</td>
<td>B + D = βπ + (1 - α)(1 - π)</td>
</tr>
<tr>
<td>threshold</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>A + B = π</td>
<td>C + D = 1 - π</td>
<td>A + B + C + D = 1</td>
</tr>
</tbody>
</table>

Think of screening for a rare disease from the perspective of a clinician. Now consider the 2 x 2
table of disease status and biomarker level. Of the proportion π with prevalent disease,
A = (1 - β)π are biomarker positive, B = βπ are biomarker negative. Of the cPrevalence, or
proportion 1 - π, without prevalent disease, C = α (1 - π) are biomarker positive and the others,
D = (1 - α)(1 - π), are biomarker negative.

The clinical value of the test must be expressed in terms of predictive values, as noted by Pepe
(1). The positive predictive value, PPV, which is the risk of disease given a positive test from
the clinical perspective, is simply \( PPV = \frac{A}{A + C} \). The essential comparator of PPV is the complement of the negative predictive value, which is the risk of disease after a negative test \( cNPV = \frac{B}{B + D} \).

The Youden score is usually defined as sum of sensitivity and specificity less one of \( Y = (1 - \beta) + (1 - \alpha) - 1 \); we prefer to think of \( Y = (1 - \beta) - \alpha \), or the difference between sensitivity and cSpecificity. When Youden is 0, sensitivity equals cSpecificity and, equivalently, cSensitivity equals Specificity, so the biomarker is not predictive; the obvious corollary that Youden of 0 implies PPV\( = cNPV = \pi \) follows because the cofactors of \( \pi \) and \( 1 - \pi \) are equal in A and in C, and also equal in B and D.

The Youden can be seen on the ROC curve for any possible threshold: the Youden is simply the vertical distance between the sensitivity, \( 1 - \beta \), of the threshold and the cSpecificity, \( \alpha \), which is the sensitivity for a worthless biomarker with the same cSpecificity \( \alpha \), as shown on the ROC diagonal. The ROC curve intersects the diagonal at thresholds where Youden=0.

The difference PPV-cNPV is a simple measure of the clinical value of the test, or the difference between risks; if PPV is close to cNPV, the screening test will not be very helpful, even if the sensitivity and specificity are high. But even predictive values do not account for the frequency of a positive test. When the frequency of positivity is very high or very low, the estimated risk for a few patients’ risk will change substantially after the test, but the risk estimated risk among the others will change only slightly. For example, the PPV for breast cancer before age 70 from carrying a BRCA mutation is around 50%, and the cNPV is about 10%, but the low frequency (\(~0.2\%) means that the clinical value of BRCA testing is less than for testing with another
biomarker with the same predictive values but a higher frequency of positivity. Thus, simply maximizing PPV by increasing threshold to extreme does not yield a useful clinical test.

Working with the representations in terms of odds, such as $\frac{\pi}{1-\pi} = \frac{\text{Prev}}{c\text{Prev}}$ instead of probabilities like $\pi$ or $\text{Prev}$, the following equations follow from the table:

$$\frac{PPV}{cPPV} = \frac{\text{Prev} \cdot \text{Sensitivity}}{c\text{Prev} \cdot c\text{Specificity}}$$

or the odds of disease after a positive test $\frac{A}{C}$ is the product of the prevalence odds and the ratio $\frac{1-\beta}{\alpha}$, called the likelihood ratio positive or the ratio of the likelihood or chance that the biomarker is positive in cases (sensitivity) the chance that the biomarker is false (cSpecificity) in cases.

Similarly,

$$\frac{cNPV}{NPV} = \frac{\text{Prev} \cdot c\text{Sensitivity}}{c\text{Prev} \cdot \text{Specificity}}$$

where the odds of disease after a negative test $\frac{B}{D}$ is the ratio of the prevalence odds and the likelihood ratio negative, or the ratio of the chance that the biomarker is negative in cases and the chance that the biomarker is negative in controls, $\frac{\beta}{1-\alpha}$. Thus the likelihood ratios are factors that update the prior odds to obtain conditional odds of disease after a positive and negative disease. Clearly, if LR positive is above 1, and the chance of disease increase after a positive test than LR negative is below 1, and the chance of disease decreases after a negative test. By
convention, positivity of a biomarker is defined so that the sum $\alpha + \beta < 1$, to ensure that a positive test increases the conditional probability of disease.

We define $\Delta$ as a summary measure of the distance between the means of the distributions of biomarker levels in cases and controls, standardized by a function of the variation within cases and controls. $\Delta$ does not depend on sample size or significance of tests. Consider a continuous biomarker with equal variance $\sigma^2$ in cases and controls, and mean difference $\Delta$ expressed in units of the common standard deviation $\sigma$. For a given $\Delta$, the sensitivity can be calculated from the specificity by the formula $1 - \beta = 1 - F[F^{-1}(1 - \alpha) - \Delta]$, where $F$ is the distribution of the biomarker; when the biomarker is assumed to be normally or log-normally distributed, $F$ and $F^{-1}$ can be evaluated using standard tables. The AUC is the integral of the sensitivity over the range of specificity from 0 to 1. Figure 5 shows the relationship between $\Delta$ and AUC for a normally distributed biomarker with equal variance in cases and controls.

2. **From difference in Means to Risk Stratification: A spreadsheet**

The supplementary materials include a spreadsheet with four sheets.

The first sheet includes the table of contents and links to the individual sheets.

The second sheet “Delta and Sens, Spec” calculates the Delta (cell F13) from the means, standard errors, and sample sizes (cells D6-E8) of the biomarker levels in cases and controls. The spreadsheet calculates the PPV, cNPV, difference between PPV and cNPV, as well as case numbers detected per 1000 individuals screened for pairs of sensitivity and specificity consistent with the observed delta (F13) if the biomarker is normally distributed, for a given disease prevalence.

The third sheet “Specificity and PPV” calculates the specificity (E17:L26) required to achieve a desired PPV (E13-L13) given a prior unconditional probability of disease (E5) and test sensitivity (D17-D26).

The fourth sheet “Sensitivity and cNPV” shows the sensitivity (E21:L33) required to achieve a desired cNPV (E13-L13) given a prior unconditional probability of disease (E5) and test specificity (D21-D33).

Users can customize the spreadsheet for a given application.