

This is the author's version of the following supplementary, as accepted for publication: -

Lord, S., St John, A., Bossuyt, P.M.M., Sandberg, S., Monaghan, P.J., O'Kane, M., Cobbaert, C.M., Roddiger, R., Lennartz, L., Gelfi, C., Horvath, A.R., and the Test Evaluation Working Group of the European Federation of Clinical Chemistry and Laboratory Medicine. (2019) Setting clinical performance specifications to develop and evaluate biomarkers for clinical use. *Annals of Clinical Biochemistry: International Journal of Laboratory Medicine, Early View Online First*. doi: 10.1177/0004563219842265

<https://journals.sagepub.com/doi/abs/10.1177/0004563219842265>

Supplementary Figure: Two-by-two table for estimating clinical test accuracy and the trade-off between patient health benefits versus harms when acting on test results with worked example for calculation of minimum acceptable clinical performance.

Two-by-two table for test accuracy measures				Benefit:harm trade-off
Test	Target present	condition absent		
positive	true positive (TP)	false positive (FP)	PPV $TP/TP+FP$	TP:FP threshold Better disease outcomes versus more iatrogenic harm
negative	false negative (FN)	true negative (TN)	NPV $TN/TN+FN$	TN:FN threshold Less iatrogenic harm versus more worse disease outcomes
	Sensitivity $TP/TP+FN$	Specificity $TN/TN+FP$		

Worked example

Serum biomarker as a screening test for ovarian cancer in asymptomatic women aged 50 years and older. All positives tests (TP and FP) will be referred for further investigation by intravaginal ultrasound and potentially laparotomy.

1. Set FP:TP threshold

Trigger question: What is the highest number of individuals having a FP that you would be prepared to accept for one additional individual to have the benefit of a TP finding?

For the purpose of this example, if you would tolerate up to 49 women being referred for unnecessary further testing for one additional individual to have the benefit of one TP finding, the FP:TP threshold = 49:1

This threshold indicates you value the potential benefit of identifying one TP 49 times higher than avoiding the potential harmful consequences of a FP, i.e. harm:benefit ratio = 1:49

2. Calculate the minimum acceptable PPV

$$\begin{aligned}
\text{Minimum PPV} &= TP/TP+FP \\
&= 1/(1+49) \\
&= 1/50 \\
&= 2\%
\end{aligned}$$

A biomarker that achieves $PPV \geq 2\%$ warrants further evaluation.

3. Calculate the minimum acceptable sensitivity and specificity

Apply the following formula (see Pepe et al. 2016):

$$\frac{\text{Sensitivity}}{(1 - \text{Specificity})} \geq \frac{(1 - \text{prevalence})}{\text{prevalence}} * \text{harm: benefit ratio}$$

Estimate prevalence of ovarian cancer in women ≥ 50 years at first screen

= 5 per 10,000 = 0.0005 (estimated from Buys et al. 2011)

$$\begin{aligned}
\frac{\text{Sensitivity}}{(1 - \text{Specificity})} &\geq \frac{(0.9995)}{0.0005} * 1:49 \\
&\geq 1999 * 0.02 \\
&\geq 40
\end{aligned}$$

Any combination of sensitivity and specificity meeting this value (sensitivity/(1-specificity ≥ 40) warrants further evaluation.

Note, one can calculate the lowest minimum acceptable specificity across the full range of sensitivity values by assuming 100% sensitivity can be achieved as follows:

$$\begin{aligned}
(1 - \text{Specificity}) &\leq \text{Sensitivity}/40 \\
&\leq 1.00/40 \\
&\leq 0.025 (2.5\%)
\end{aligned}$$

Thus minimum acceptable specificity is $(1 - 0.025) = 0.975$, i.e. a biomarker must achieve a specificity of 97.5% to warrant further evaluation.

Specificity must be even higher if biomarker sensitivity is < 100%

Use the formula to check the minimum acceptable sensitivity at different specificity levels. For example, if a specificity of 98% can be achieved:

Minimum acceptable sensitivity $\geq 40^* (1 - \text{Specificity})$

$\geq 40^* (1 - 0.98)$

$\geq 40^* (0.02)$

$\geq 0.8 (80\%)$

References

Buyss SS, Partridge E, Black A, et al. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. JAMA. 2011; 305: 2295-303.

Pepe MS, Janes H, Li CI, Bossuyt PM, et al. Early-Phase Studies of Biomarkers: What Target Sensitivity and Specificity Values Might Confer Clinical Utility? Clin Chem. 2016; 62: 737-42.