Laboratory-based evaluation of a simplified point-of-care test intended to support treatment decisions in non-severe bovine clinical mastitis

Francisco B Malcata, P Theo Pepler, Ruth N Zadoks, Lorenzo Viora

SUPPLEMENTARY FILE: Material and methods full detail

Reference test

Sheep blood agar (5% vol/vol; SBA) and MacConkey agar number 3 plates (E&O Laboratories Limited, Bonnybridge, Scotland) were inoculated with 0.01 ml of milk each using disposable sterile calibrated plastic loops. Plates were incubated aerobically at 37°C and examined after approximately 48 hours. Samples without growth of visible colonies were considered negative for mastitis-associated pathogens. Samples that yielded three or more colony types were considered contaminated and excluded from data analysis in accordance with NMC guidelines. For the remaining plates, each colony type was sub-cultured onto SBA for purification. From each pure culture, a colony was selected and grown aerobically in 2 ml of Brain Heart Infusion broth for 24 hours at 37°C without shaking. Isolates were preserved with 15% glycerol (v/v) in cryovials at -80°C and submitted to an external laboratory (Laboratoire de Microbiologie, Vétoquinol SA, Lure, France) for species identification by MALDI-ToF MS analysis, using Vitek-MS and the V3.1.0 database (bioMérieux, Marcy-l'Étoile, France).

Comparator test

The sectors of the comparator test contain selective indicator media for gram-negative organisms, staphylococci and gram-positive catalase negative cocci, respectively (Viora *et al.* 2014) (Supplemental Figure S1). Based on the manufacturer's guidelines, eight common mastitis-associated pathogen species or genera could be identified after 48 hr incubation: *E. coli*, *Klebsiella* spp., *Staphylococcus aureus*, non-aureus staphylococci (NAS), *Streptococcus uberis*, *Enterococcus* spp., *Streptococcus dysgalactiae*, and *Streptococcus agalactiae*.

Data analysis

For culture-positive samples with gram-positive or gram-negative species as identified by the reference test, matching results from the slide test or the comparator test were considered true positives (TP) and non-matching results were considered false negatives (FN) or false positives (FP). Samples that were negative for an outcome of interest on the reference test with matching results on slide test or comparator test were considered true negative (TN). For example, if a sample yielded *Staphylococcus haemolyticus* with the reference test, gram-negative growth other than *E. coli* on the slide test and Klebsiella spp. on the comparator test, it was considered a TP for growth, FN for gram-positive organisms, FP for gram-negative organisms and TN for *E. coli*. From those classifications, sensitivity (Se), specificity (Sp), accuracy (Acc), positive predictive value (PPV) and negative predictive value (NPV) were calculated as follow:

$$Se = TP/(TP+FN)$$

$$Sp = TN/(FP+TN)$$

$$Acc = (TP+TN)/n$$

$$PPV = TP/(TP+FP)$$

$$NPV = TN/(FN+TN)$$

Epidemiological parameters were expressed as percentages with 95% Wilson type confidence intervals (CI), calculated using the Hmisc package in R (Harrel Jr & Dupont, 2019). Wilson intervals are preferred over exact intervals and Wald (normal approximation) type intervals, as they have coverage probability closer to the nominal value (Agresti & Coull, 1998) and confidence limits that do not exceed the boundaries of the unit interval. The parameter estimates for the slide test and the comparator test are not independent because they are derived from the same sample. To account for this dependence when comparing Se, Sp, Acc, PPV and NPV for the two tests, Wald type confidence intervals for the differences between these measures were calculated using formulae derived from Kosinski (2013); see below for full detail. If the 95% confidence interval for the difference between tests exclude zero, test performance was considered significantly different.

The methods by which Wald confidence intervals were calculated for differences in sensitivity, specificity, overall accuracy, positive predictive value, and negative predictive value between the VétoSlide and VétoRapid tests are now described.

Sensitivity

Tabulate the number of positive samples (N), as per the reference test:

		VétoRapid		
		TP	FN	Total
VétoSlide	TP	n_{11}	n_{12}	$n_{1\bullet}$
	FN	n_{21}	n_{22}	$n_{2\bullet}$
	Total	$n_{\bullet 1}$	$n_{\bullet 2}$	N

Dividing every cell in the table by N gives the proportions:

		VétoRapid		
		TP	FN	Total
VétoSlide	TP	p_{11}	p_{12}	$p_{1\bullet}$
	FN	p_{21}	p_{22}	$p_{2\bullet}$
	Total	$p_{\bullet 1}$	$p_{\bullet 2}$	1

The sensitivities of VétoSlide and VétoRapid are $Se_{VS} = \frac{n_{1\bullet}}{N} = p_{1\bullet}$ and $Se_{VR} = \frac{n_{\bullet 1}}{N} = p_{\bullet 1}$, respectively. The difference between the two sensitivities are $D = Se_{VR} - Se_{VS}$, and the variance of this difference is (Agresti, 2012: p414),

$$\hat{\sigma}^2(D) = \frac{(p_{12}+p_{21})-(p_{12}-p_{21})^2}{N} \text{ (Equation 1)}$$

A 95% Wald confidence interval, under the hypothesis of no difference between the sensitivities, can be calculated with,

$$D \pm z_{1-\alpha/2}\hat{\sigma}(D) \text{(Equation 2)}$$

where $z_{1-\alpha/2} = 1.96$ is the appropriate quantile from a standard normal distribution.

Example. Suppose there are 100 positive (bacterial culture) samples that are tabulated as follows:

		VétoRapid		
		TP	FN	Total
VétoSlide	TP	80	2	82
	FN	10	8	18
	Total	90	10	N=100

The cell proportions are:

		VétoRapid		
		TP	FN	Total
VétoSlide	TP	0.80	0.02	0.82
	FN	0.10	0.08	0.18
	Total	0.90	0.10	1

The sensitivities are $Se_{VS} = 0.82$ and $Se_{VR} = 0.90$, and $D = Se_{VR} - Se_{VS} = 0.08$. The variance of this difference (D) is,

$$\hat{\sigma}^2(D) = \frac{(0.02 + 0.10) - (0.02 - 0.10)^2}{100} = 0.00114.$$

The 95% Wald confidence interval for the difference is,

$$0.08 \pm 1.96\left(\sqrt{0.00114}\right),$$

which is the interval, [0.014; 0.146]. By contrast, if the dependence between the sensitivities is ignored, the 95% confidence interval is [-0.016; 0.176], which includes the value 0.

Specificity

Tabulate the number of negative samples (N), as per the reference test:

		VétoRapid		
		TN	FP	Total
VétoSlide	TN	n_{11}	n_{12}	$n_{1\bullet}$
	FP	n_{21}	n_{22}	$n_{2\bullet}$
	Total	$n_{\bullet 1}$	$n_{\bullet 2}$	N

Divide all cells by N to give the proportions:

		VétoRapid		
		TN	FP	Total
VétoSlide	TN	p_{11}	p_{12}	$p_{1\bullet}$
	FP	p_{21}	p_{22}	$p_{2\bullet}$
	Total	$p_{\bullet 1}$	$p_{\bullet 2}$	1

The specificities of VétoRapid and VétoSlide are $Sp_{VR} = \frac{n_{\bullet 1}}{N} = p_{\bullet 1}$ and $Sp_{VS} = \frac{n_{1\bullet}}{N} = p_{1\bullet}$, respectively.

The difference between the two specificities are $D = Sp_{VR} - Sp_{VS}$, and the variance of this difference, $\hat{\sigma}^2(D)$, is calculated as per Equation 1. A 95% Wald confidence interval for D is calculated as per Equation 2.

Accuracy

Consider the classification of all samples (N) as either *Correct (= TP + TN)* or *Incorrect (= FP + FN)*, by each of the two diagnostic tests. Tabulate these cases as follows:

		VétoRapid		
		Correct	Incorrect	Total
VétoSlide	Correct	n_{11}	n_{12}	$n_{1\bullet}$
	Incorrect	n_{21}	n_{22}	$n_{2\bullet}$
	Total	$n_{\bullet 1}$	$n_{\bullet 2}$	N

Divide all cells by N to give the proportions:

		VétoRapid		
		Correct	Incorrect	Total
VétoSlide	Correct	p_{11}	p_{12}	$p_{1\bullet}$
	Incorrect	p_{21}	p_{22}	$p_{2\bullet}$
	Total	$p_{\bullet 1}$	$p_{\bullet 2}$	1

The overall accuracies of VétoRapid and VétoSlide are $Acc_{VR} = \frac{n_{\bullet 1}}{N} = p_{\bullet 1}$ and $Acc_{VS} = \frac{n_{1\bullet}}{N} = p_{1\bullet}$, respectively. The difference between the two accuracies is $D = Acc_{VR} - Acc_{VS}$.

The variance of D, $\hat{\sigma}^2(D)$, is calculated as per Equation 1, and a 95% Wald confidence interval for D is calculated as per Equation 2.

Positive predictive value

For Sensitivity, Specificity and Accuracy we condition on disease status; that is, the denominator (N) is the same for both tests (VétoSlide and VétoRapid). However, for positive predictive value (PPV) we condition on test outcome; that is, $N = TP + FP$ (the number of positives indicated by the specific test), which will differ for the two tests. Similarly for negative predictive value (NPV), $N = TN + FN$, which again differs for the two tests. Thus we cannot use the same methodology as for sensitivity, specificity and accuracy to compare the PPVs and NPVs of VétoSlide and VétoRapid.

Kosinski (2013) provided formulae for the variance of the contrast between two PPVs calculated from paired data. Tabulate the results (number of cases) for VétoRapid and VétoSlide as follows, using the letters *a* to *h* to indicate cells in the table:

		Bacteria+		Bacteria-	
		VétoRapid		VétoRapid	
		Positive	Negative	Positive	Negative
VétoSlide	Positive	*a*	*b*	*e*	*f*
	Negative	*c*	*d*	*g*	*h*

"Bacteria+" and "Bacteria-" indicate results from the reference test (bacterial culture).

Positive predictive values for the VétoRapid and VétoSlide, respectively, are:

$$PPV_{VR} = \frac{a + c}{a + c + e + g}$$

$$PPV_{VS} = \frac{a + b}{a + b + e + f}$$

The PPVs are calculated on partly dependent subsets of the total number of samples. The covariance of the PPVs is:

$$Cov(PPV_{VR}, PPV_{VS}) = \frac{a(1 - PPV_{VR})(1 - PPV_{VS}) + ePPV_{VR}PPV_{VS}}{2a + b + 2e + f + c + g}.$$

The difference between the PPVs is $D = PPV_{VR} - PPV_{VS}$. The variance of this difference is calculated as,

$$\hat{\sigma}^2(D) = \frac{PPV_{VR}(1 - PPV_{VR})}{a + c + e + g} + \frac{PPV_{VS}(1 - PPV_{VS})}{a + b + e + f}$$

$$- 2Cov(PPV_{VR}, PPV_{VS}) \left[\frac{1}{a + b + e + f} + \frac{1}{a + c + e + g}\right]$$

A 95% Wald confidence interval for the difference is calculated as per Equation 2.

Negative predictive value

With reference to the table in the PPV section above, the negative predictive values of VétoRapid and VétoSlide are, respectively:

$$NPV_{VR} = \frac{f + h}{b + d + f + h}$$

$$NPV_{VS} = \frac{g + h}{c + d + g + h}$$

The covariance of the NPVs is,

$$Cov(NPV_{VR}, NPV_{VS}) = \frac{dNPV_{VR}NPV_{VS} + h(1 - NPV_{VR})(1 - NPV_{VS})}{b + 2d + f + 2h + c + g}.$$

The difference between the NPVs is calculated as $D = NPV_{VR} - NPV_{VS}$, and the variance of this difference is,

$$\hat{\sigma}^2(D) = \frac{NPV_{VR}(1 - NPV_{VR})}{b + d + f + h} + \frac{NPV_{VS}(1 - NPV_{VS})}{c + d + g + h}$$

$$- 2Cov(NPV_{VR}, NPV_{VS}) \left[\frac{1}{b + d + f + h} + \frac{1}{c + d + g + h}\right].$$

A 95% Wald confidence interval for the difference is calculated as per Equation 2.

References

Agresti A 2012 Categorical Data Analysis. 3rd Edition. *Wiley*. New York, USA

Harrell Jr, FE & C Dupont 2019 Hmisc: Harrell miscellaneous. R package version 4.2-0, Available at: https://CRAN.R-project.org/package=Hmisc (accessed 5 May 2019)

Viora L, EM Graham, DJ Mellor, K Reynolds, PBA Simoes & TE Geraghty 2014 Evaluation of a culture-based pathogen identification kit for bacterial causes of bovine mastitis. *Veterinary Record* **175** 89

Kosinski AS 2013 A weighted generalized score statistic for comparison of predictive values of diagnostic tests. *Statistics in Medicine* **32** 964–977.