Supplementary material of the manuscript:

Asymptotic confidence intervals for the difference and the ratio of the weighted kappa coefficients of two diagnostic tests subject to a paired design

REVSTAT Statistical Journal

José Antonio Roldán-Nofuentes, Saad Bouh Sidaty-Regad

Appendix *A*

From now onwards, we are going to suppose that $0 < Se_h < 1$, $0 < Sp_h < 1$, $0 < p < 1$ and $q = 1 - p$. Performing algebraic operations it is verified that

$$\kappa_1(c) - \kappa_2(c) = \frac{pq}{D_1 D_2}\nu \tag{1}$$

where $D_h = p(1-Q_h)c + qQ_h(1-c)$ is the denominator of $\kappa_h(c)$, with $h = 1,2$, and

$$\nu = q\Delta_1 - c(\Delta_1 - p\Delta_2) \tag{2}$$

where $\Delta_1 = Se_1(1-Sp_2) - Se_2(1-Sp_1)$ and $\Delta_2 = Y_1 - Y_2 = Se_1 - Se_2 + Sp_1 - Sp_2$. Then $\kappa_1(c) > \kappa_2(c)$ if $\nu > 0$, since $D_h > 0$. Solving equation $\kappa_1(c) - \kappa_2(c) = 0$ in c it holds that

$$c' = c = \frac{q\Delta_1}{\Delta_1 - p\Delta_2}, \tag{3}$$

being c' a real value. From now onwards, the rules so that $\kappa_1(c) > \kappa_2(c)$, $\kappa_2(c) > \kappa_1(c)$ and $\kappa_1(c) = \kappa_2(c)$, considering that $i = 1$ and $j = 2$ (the demonstrations for $i = 2$ and $j = 1$ are analogous).

a) If $rTPF_{12} \geq 1$ and $rFPF_{12} < 1$, or $rTPF_{12} > 1$ and $rFPF_{12} \leq 1$, then $\kappa_1(c) > \kappa_2(c)$ for $0 \leq c \leq 1$.

Let us suppose in the first place that $rTPF_{12} = 1$ and that $rFPF_{12} < 1$, then $Se_1 = Se_2 = Se$ and $Sp_1 > Sp_2$. Substituting in equation (2) it holds that $\nu = (Sp_1 - Sp_2)[cp + (q-c)Se]$. Here $\nu > 0$ if $cp + (q-c)Se > 0$, since $Sp_1 > Sp_2$. If $c = 0$ or $c = 1$, then $cp + (q-c)Se > 0$ since $qSe > 0$, and $p(1-Se) > 0$ is verified; and as $Sp_1 > Sp_2$, then $\nu > 0$ and $\kappa_1(c) > \kappa_2(c)$. Let us suppose that $0 < c < 1$ and $p \geq Se$, then $cp + (q-c)Se = c(p - Se) + qSe > 0$, and it is verified that $\nu > 0$ and $\kappa_1(c) > \kappa_2(c)$. If $p < Se$, then $cp + (q-c)Se = (1-c)(Se - p) + (1-Se)p > 0$, since $(1-c)(Se-p) > 0$ and $(1-Se)p > 0$. Therefore, $\nu > 0$ and $\kappa_1(c) > \kappa_2(c)$.

Let us now suppose that $rTPF_{12} > 1$ and that $rFPF_{12} < 1$, then $Se_1 > Se_2$ and $Sp_1 > Sp_2$. It is easy to check that when $c = 0$ or $c = 1$ it is verified that $\nu > 0$ and, therefore, $\kappa_1(c) > \kappa_2(c)$. Moreover, as $rTPF_{12} > 1$ and $rFPF_{12} < 1$ then dividing both parameters $(rTPF_{12}/rFPF_{12} > 1)$ it holds that $\frac{rTPF_{12}}{rFPF_{12}} = \frac{Se_1(1-Sp_2)}{Se_2(1-Sp_1)} > 1$, verifying that $\Delta_1 = Se_1(1-Sp_2) - Se_2(1-Sp_1) > 0$. As $Se_1 > Se_2$ and $Sp_1 > Sp_2$ then $\Delta_2 = Se_1 - Se_2 + Sp_1 - Sp_2 > 0$. Furthermore, as it is verified that $Se_1 > Se_2$ then $1 - Se_1 < 1 - Se_2$, and $0 < \frac{1-Se_1}{1-Se_2} < 1$. Moreover, as $\frac{Sp_1}{Sp_2} > 1$ then $\frac{Sp_1}{Sp_2} - \frac{1-Se_1}{1-Se_2} = \frac{\Delta_3}{Sp_2(1-Se_2)} > 0$, when $\Delta_3 = (1-Se_2)Sp_1 - (1-Se_1)Sp_2 > 0$. It is easy to check that $\Delta_1 = \Delta_2 - \Delta_3$, so that $\Delta_2 > \Delta_1$. Equation (2) can be written as

$$\nu = (q-c)\Delta_1 + cp\Delta_2 . \tag{4}$$

Let us suppose that $0 < c < 1$, then if $q \geq c$ it is verified that $\nu > 0$ and $\kappa_1(c) > \kappa_2(c)$. Let us now suppose that $q < c$, then $q - c < 0$. Equation (4) can be written as

$$\nu = -(c-q)\Delta_1 + cp\Delta_2$$

being $c-q>0$. Let us suppose that

$$\nu < 0 \Rightarrow -(c-q)\Delta_1 + cp\Delta_2 < 0,$$

so that

$$-(c-q)\Delta_1 < -cp\Delta_2 \Rightarrow (c-q)\Delta_1 > cp\Delta_2 \Rightarrow c-q > cp\frac{\Delta_2}{\Delta_1}.$$

As $\Delta_2 > \Delta_1$ then $\frac{\Delta_2}{\Delta_1} > 1$, so that

$$c-q > cp\frac{\Delta_2}{\Delta_1} > cp > 0,$$

from where we obtain

$$c-q-cp > 0. \tag{5}$$

Performing algebraic operations

$$c-q-cp = q(c-1)$$

As $0<c<1$, $1-c>0$ and $c-1<0$, then $q(c-1)<0$, which is contradictory with expression (5). Therefore, if $q<c$ then $\nu>0$ and $\kappa_1(c)>\kappa_2(c)$.

The demonstrations for $rTPF_{12}>1$ and $rFPF_{12}\le 1$ are performed following a similar process to the previous one.

b). If $rTPF_{12}>1$ and $rFPF_{12}>1$, then:

b.1) $\kappa_1(c)>\kappa_2(c)$ if $0<c'<c\le 1$

b.2) $\kappa_1(c)<\kappa_2(c)$ if $0\le c<c'<1$

b.3) $\kappa_1(c)=\kappa_2(c)$ if $c=c'$, with $0<c'<1$

b.4) $\kappa_1(c)>\kappa_2(c)$ for $0\le c\le 1$ if $c'<0$ (or $c'>1$) and $rTPF_{12}>rFPF_{12}>1$

b.5) $\kappa_1(c) < \kappa_2(c)$ for $0 \le c \le 1$ if $c' < 0$ (or $c' > 1$) and $rFPF_{12} > rTPF_{12} > 1$

Firstly, we are going to demonstrate that c' cannot be equal to 0 or to 1. As $rTPF > 1$ and $rFPF > 1$, then it is verified that $Se_1 > Se_2$ and $Sp_1 < Sp_2$. If $c' = 0$ then $\Delta_1 = 0$, and it is verified that

$$\frac{Se_1}{Se_2} \times \frac{1 - Sp_2}{1 - Sp_1} = 1,$$

which is incompatible with $rTPF > 1$ and $rFPF > 1$, since as $\frac{Se_1}{Se_2} > 1$ and $0 < \frac{1 - Sp_2}{1 - Sp_1} < 1$ then it is verified that $\frac{Se_1}{Se_2} \times \frac{1 - Sp_2}{1 - Sp_1} \neq 1$. Therefore c' cannot be equal to 0 if $rTPF > 1$ and $rFPF > 1$. If $c' = 1$ then $\Delta_1 - \Delta_2 = Sp_2(1 - Se_1) - Sp_1(1 - Se_2) = 0$, and it is verified that

$$\frac{Sp_2}{Sp_1} \times \frac{1 - Se_1}{1 - Se_2} = 1,$$

which is incompatible with $rTPF > 1$ and $rFPF > 1$, since as $\frac{Sp_2}{Sp_1} > 1$ and $0 < \frac{1 - Se_1}{1 - Se_2} < 1$ then it is verified that $\frac{Sp_2}{Sp_1} \times \frac{1 - Se_1}{1 - Se_2} \neq 1$. Therefore, c' cannot be equal to 1 if $rTPF > 1$ and $rFPF > 1$.

Let us consider that $0 < c' < 1$, then we must verify one of the two following: 1) $0 < q\Delta_1 < \Delta_1 - p\Delta_2$, or 2) $\Delta_1 - p\Delta_2 < q\Delta_1 < 0$. Condition 1 implies that $\Delta_1 > 0$ and $\Delta_1 > p\Delta_2$, and Condition 2 implies that $\Delta_1 < 0$ and $\Delta_1 < p\Delta_2$.

Moreover, as $Se_1 > Se_2$ and $Sp_1 < Sp_2$ (which implies that $1 - Sp_1 > 1 - Sp_2$) then $Q_1 > Q_2$. Furthermore, if $c = c'$ then performing algebraic operations, each weighted kappa coefficient is expressed as

$$\kappa_h(c') = \frac{Y_h}{\tau_h},$$

when $\tau_h = \frac{\Delta_1 - Q_h \Delta_2}{\Delta_1 - p\Delta_2}$, with $h = 1,2$. As $Q_1 > Q_2$, then $\tau_2 - \tau_1 > 0$ if $\Delta_2 > 0$, and $\tau_2 - \tau_1 < 0$ if $\Delta_2 < 0$. If $\Delta_2 > 0$, then

$$\tau_2 - \tau_1 = \frac{\Delta_2 (Q_1 - Q_2)}{\Delta_1 - p\Delta_2} > 0 \Rightarrow \Delta_1 - p\Delta_2 > 0 \Rightarrow \Delta_1 > p\Delta_2 > 0.$$

If $\Delta_2 < 0$, then

$$\tau_2 - \tau_1 = \frac{\Delta_2 (Q_1 - Q_2)}{\Delta_1 - p\Delta_2} < 0 \Rightarrow \Delta_1 - p\Delta_2 > 0 \Rightarrow \Delta_1 > p\Delta_2 .$$

Therefore, whether $\Delta_2 > 0$ or $\Delta_2 < 0$, it is always verified that $\Delta_1 > p\Delta_2$. This condition is only compatible with Condition 1 obtained by the fact that $0 < c' < 1$, i.e. $0 < q\Delta_1 < \Delta_1 - p\Delta_2$. Therefore, it is always verified that $\Delta_1 > 0$ and $\Delta_1 > p\Delta_2$.

Moreover, from equation (3) it holds that $q\Delta_1 = c'(\Delta_1 - p\Delta_2)$, so that substituting this expression in equation (2) it holds that

$$\nu = (\Delta_1 - p\Delta_2)(c' - c). \tag{6}$$

As $\Delta_1 > p\Delta_2$ then $\Delta_1 - p\Delta_2 > 0$. Based on equation (6), if $0 \le c < c' < 1$ then $\nu > 0$ and $\kappa_1(c) > \kappa_2(c)$. If $0 < c' < c \le 1$ then $\nu < 0$ and $\kappa_1(c) < \kappa_2(c)$, and if $c = c'$ (with $0 < c' < 1$) then $\nu = 0$ and $\kappa_1(c) = \kappa_2(c)$.

If $c' < 0$ then one of the following two conditions must be verified: 1) $0 < q\Delta_1 < \Delta_1 < p\Delta_2 < \Delta_2$, or 2) $\Delta_2 < p\Delta_2 < \Delta_1 < q\Delta_1 < 0$. Condition 1 implies that $\Delta_1 > 0$ and therefore $Se_1\left(1-Sp_2\right) > Se_2\left(1-Sp_1\right)$, and from this inequality it holds that

$$\frac{Se_1}{Se_2} > \frac{1-Sp_1}{1-Sp_2} > 1 \Rightarrow rTPF_{12} > rFPF_{12} > 1.$$

As $q\Delta_1 > 0$ and $\Delta_1 - p\Delta_2 < 0$, then applying equation (2) it holds that $\nu > 0$ and therefore $\kappa_1\left(c\right) > \kappa_2\left(c\right)$. Condition 2 implies that $\Delta_1 < 0$ and therefore $Se_1\left(1-Sp_2\right) < Se_2\left(1-Sp_1\right)$, and it holds that

$$\frac{1-Sp_1}{1-Sp_2} > \frac{Se_1}{Se_2} > 1 \Rightarrow rFPF_{12} > rTPF_{12} > 1.$$

As $q\Delta_1 < 0$ and $\Delta_1 - p\Delta_2 > 0$, applying equation (2) again it holds that $\nu < 0$ and therefore $\kappa_1\left(c\right) < \kappa_2\left(c\right)$. If $c' > 1$, the demonstrations are similar to those of $c' < 0$.

c) If $rTPF_{12} < 1$ and $rFPF_{12} < 1$, then $rTPF_{21} > 1$ and $rFPF_{21} > 1$, and the demonstrations are analogous to case b).

Appendix *B*

Bloch (1997) has deduced the expressions of the variances of $\hat{\kappa}_1\left(c\right)$ and $\hat{\kappa}_2\left(c\right)$ and of the covariance between them. We then obtain equivalent expressions and we also deduce the variance of the ratio of the two weighted kappa coefficients, an expression which is necessary to apply the method to calculate the sample size explained in Section 5. Let $\boldsymbol{\omega} = \left(Se_1, Sp_1, Se_2, Sp_2, p\right)^T$ be the vector of parameters, where $Se_1 = \frac{p_{10} + p_{11}}{p}$,

$Sp_1 = \frac{q_{00}+q_{01}}{q}$, $Se_2 = \frac{p_{01}+p_{11}}{p}$ and $Sp_2 = \frac{q_{00}+q_{10}}{q}$, with $q = 1-p$. Applying the delta method, the matrix of the asymptotic variances-covariances of $\hat{\boldsymbol{\omega}}$ is

$$\Sigma_{\hat{\boldsymbol{\omega}}} = \left(\frac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{\pi}}\right) \Sigma_{\hat{\boldsymbol{\pi}}} \left(\frac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{\pi}}\right)^T .$$

Performing the algebraic operations it is obtained that

$$Var\left(\hat{S}e_1\right) = \frac{\left(p_{11}+p_{10}\right)\left(p_{01}+p_{00}\right)}{np^3} = \frac{Se_1\left(1-Se_1\right)}{np},$$

$$Var\left(\hat{S}e_2\right) = \frac{\left(p_{11}+p_{01}\right)\left(p_{10}+p_{00}\right)}{np^3} = \frac{Se_2\left(1-Se_2\right)}{np},$$

$$Var\left(\hat{S}p_1\right) = \frac{\left(q_{11}+q_{10}\right)\left(q_{01}+q_{00}\right)}{nq^3} = \frac{Sp_1\left(1-Sp_1\right)}{nq},$$

$$Var\left(\hat{S}p_2\right) = \frac{\left(q_{11}+q_{01}\right)\left(q_{10}+q_{00}\right)}{nq^3} = \frac{Sp_2\left(1-Sp_2\right)}{nq},\ Var\left(\hat{p}\right) = \frac{pq}{n},$$

$$Cov\left[\hat{S}e_1, \hat{S}e_2\right] = \frac{p_{11}p_{00}-p_{10}p_{01}}{np^3} = \frac{\varepsilon_1}{np},\ Cov\left[\hat{S}p_1, \hat{S}p_2\right] = \frac{q_{11}q_{00}-q_{10}q_{01}}{nq^3} = \frac{\varepsilon_0}{nq}$$

and

$$Cov\left(\hat{S}e_h, \hat{S}p_h\right) = Cov\left(\hat{S}e_h, \hat{p}\right) = Cov\left(\hat{S}p_h, \hat{p}\right) = 0 \text{, with } h = 1,2 .$$

The estimators of the variances-covariances are obtained substituting each parameter with its corresponding estimator, where $\hat{S}e_1 = \frac{s_{11}+s_{10}}{s}$, $\hat{S}e_2 = \frac{s_{11}+s_{01}}{s}$, $\hat{S}p_1 = \frac{r_{01}+r_{00}}{r}$, $\hat{S}p_2 = \frac{r_{10}+r_{00}}{r}$, $\hat{p} = \frac{s}{n}$, $\hat{q} = \frac{r}{n}$, $\hat{\varepsilon}_1 = \frac{\hat{p}_{11}}{\hat{p}} - \hat{S}e_1\hat{S}e_2 = \frac{s_{11}s_{00}-s_{10}s_{01}}{s^2}$ and $\hat{\varepsilon}_0 = \frac{\hat{q}_{00}}{\hat{q}} - \hat{S}p_1\hat{S}p_2 = \frac{r_{11}r_{00}-r_{10}r_{01}}{r^2}$. Applying the delta method, the variance of $\hat{\kappa}_h\left(c\right)$ is

$$Var\left[\hat{\kappa}_h\left(c\right)\right] \approx \left[\frac{\partial \kappa_h\left(c\right)}{\partial Se_h}\right]^2 Var\left(\hat{S}e_h\right) + \left[\frac{\partial \kappa_h\left(c\right)}{\partial Sp_h}\right]^2 Var\left(\hat{S}p_h\right) + \left[\frac{\partial \kappa_h\left(c\right)}{\partial p}\right]^2 Var\left(\hat{p}\right).$$

In this expression the covariances are zero. Performing the algebraic operations, it is obtained that

$$Var\left[\hat{\kappa}_h(c)\right] \approx \left[\frac{\kappa_h(c)}{pqY_h}\right]^2 \times \left[\left\{a_{h1}^2 Var\left(\hat{S}e_h\right) + a_{h2}^2 Var\left(\hat{S}p_h\right) + a_{h3}^2 Var\left(\hat{p}\right)\right\}\right]$$

with $h = 1,2$, and where

$$a_{h1} = pq - p(q-c)\kappa_h(c),$$

$$a_{h2} = a_{h1} + (q-c)\kappa_h(c)$$

and

$$a_{h3} = (1-2p)Y_h - \left[(1-c-2p)Y_h + Sp_h + c - 1\right]\kappa_h(c).$$

The expression of $\hat{V}ar\left[\hat{\kappa}_h(c)\right]$ is obtained substituting in the previous expressions each parameter with its estimator. Regarding the covariance between $\hat{\kappa}_1(c)$ and $\hat{\kappa}_2(c)$, applying the delta method again it is obtained that

$$Cov\left[\hat{\kappa}_1(c), \hat{\kappa}_2(c)\right] \approx \frac{\partial\kappa_1(c)}{\partial Se_1}\frac{\partial\kappa_2(c)}{\partial Se_2} Cov\left[\hat{S}e_1, \hat{S}e_2\right] + \frac{\partial\kappa_1(c)}{\partial Sp_1}\frac{\partial\kappa_2(c)}{\partial Sp_2} Cov\left[\hat{S}p_1, \hat{S}p_2\right] + \frac{\partial\kappa_1(c)}{\partial p}\frac{\partial\kappa_2(c)}{\partial p} Var\left(\hat{p}\right).$$

In this expression, the rest of the covariances are equal to zero. Performing the algebraic operations it is obtained that

$$Cov\left[\hat{\kappa}_1(c), \hat{\kappa}_2(c)\right] \approx \frac{\kappa_1(c)\kappa_2(c)}{p^2q^2Y_1Y_2} \times \left[a_{11}a_{21} Cov\left(\hat{S}e_1, \hat{S}e_2\right) + a_{12}a_{22} Cov\left(\hat{S}p_1, \hat{S}p_2\right) + a_{13}a_{23} \hat{V}ar\left(\hat{p}\right)\right].$$

The expression of $\hat{C}ov\left[\hat{\kappa}_1(c), \hat{\kappa}_2(c)\right]$ is obtained substituting in this equation each parameter with its estimator.

Regarding the ration of the two weighted kappa coefficients, the variance of θ is easily calculated applying the delta method again, i.e.

$$Var\left(\hat{\theta}\right) \approx \sum_{h=1}^{2}\left(\frac{\partial\theta}{\partial\kappa_h(c)}\right)^2 Var\left[\hat{\kappa}_h(c)\right] + 2\frac{\partial\theta}{\partial\kappa_1(c)}\frac{\partial\theta}{\partial\kappa_2(c)}Cov\left[\hat{\kappa}_1(c),\hat{\kappa}_2(c)\right].$$

Performing the algebraic operations,

$$Var\left(\hat{\theta}\right) \approx \frac{\kappa_2^2(c)Var\left[\hat{\kappa}_1(c)\right] + \kappa_1^2(c)Var\left[\hat{\kappa}_2(c)\right] - 2\kappa_1(c)\kappa_2(c)Cov\left[\hat{\kappa}_1(c),\hat{\kappa}_2(c)\right]}{\kappa_2^4(c)}, \quad (7)$$

and substituting in this equation each parameter with its estimator, we obtain the expression of $\hat{V}ar\left(\hat{\theta}\right)$. The expression of variance of $\hat{V}ar\left[\ln\left(\hat{\theta}\right)\right]$ is calculated in a similar way to in the previous case, but considering $\ln(\theta)$ instead of θ.

Appendix *C*

The selection of the *CI* with the best asymptotic behaviour, both for the difference δ and for the ratio θ, was made taking the following steps: 1) Choose the *CIs* with the least failures ($CP > 93\%$), 2) Choose the *CIs* that are the most accurate i.e. those with the lowest *AL*. The first step in this method establishes that the *CI* does not fail when $CP > 93\%$. In the simulation experiments the *CIs* were calculated to a 95% confidence i.e. $\gamma = 1 - \alpha = 0.95$ is the nominal confidence and $\alpha = 5\%$ is the nominal error. Then $\Delta\alpha = \alpha - \alpha^* = \gamma^* - \gamma$, where γ^* is the *CP* calculated and α^* is the type I error.

Moreover, the hypothesis test to check the equality of the two weighted kappa coefficients is $H_0 : \kappa_1(c) = \kappa_2(c)$ vs $H_1 : \kappa_1(c) \neq \kappa_2(c)$. Based on the difference of both parameters, this hypothesis test is equivalent to test $H_0 : \delta = 0$ vs $H_1 : \delta \neq 0$. This test can be solved through different methods. Applying Bloch's method (1997), the test statistic is given by equation (equation (10) of the manuscript). The statistics for the bootstrap method and for the Bayesian method are obtained computationally.

In step 1 of the method, a *CI* has a failure if $CP \leq 93\%$, i.e. if $\Delta\alpha \leq -2$. In this situation, the type I error of the corresponding hypothesis test is $\geq 7\%$, and therefore it is a very liberal hypothesis test and it can give false significances. The criteria of 93% has been used by other authors (Price and Bonett, 2004; Martín-Andrés and Álvarez-Hernández, 2014a, 2014b; Montero-Alonso and Roldán-Nofuentes, 2018). If $\Delta\alpha > 2\%$, i.e. $CP > 97\%$, then the hypothesis test is very conservative (its type I error is very small, $< 3\%$), but it does not give false significances. Consequently, the choose of the optimal *CI* is linked to the decisions of the hypothesis test, and it is preferable to choose a conservative test rather than a very liberal one (as there will be no false significances because its type I error is lower than the nominal one). The method for the *CIs* for the ratio θ is justified in a similar way.