

Measures for concordance and discordance with applications in disease control and prevention

Peer-reviewed author version

AERTS, Marc; JUGA, Adelino & HENS, Niel (2019) Measures for concordance and discordance with applications in disease control and prevention. In: STATISTICAL METHODS IN MEDICAL RESEARCH, 28(10-11), p. 3086-3099.

DOI: 10.1177/0962280218796252

Handle: <http://hdl.handle.net/1942/29855>

# Measures for concordance and discordance with applications in disease control and prevention

Marc Aerts<sup>1</sup>, Adelino J.C. Juga<sup>1,2</sup> and Niel Hens<sup>1,3</sup>

## Abstract

Bivariate binary response data appear in many applications. Interest goes most often to a parameterization of the joint probabilities in terms of the marginal success probabilities in combination with a measure for association, most often being the odds ratio. Using e.g. the bivariate Dale model, these parameters can be modelled as function of covariates. But the odds ratio and other measures for association are not always measuring the (joint) characteristic of interest. Agreement, concordance, synchrony are in general facets of the joint distribution distinct from association and the odds ratio as in the bivariate Dale model can be replaced by such an alternative measure. Here we focus on the so-called conditional synchrony measure. But, as indicated by several authors, such a switch of parameter might lead to a parameterization that does not always lead to a permissible joint bivariate distribution. In this contribution, we propose a new parameterization in which the marginal success probabilities are replaced by other conditional probabilities as well. The new parameters, one homogeneity parameter and two synchrony/discordance parameters, guarantee that the joint distribution is always permissible. Moreover, having a very natural interpretation, they are of interest on their own. The applicability and interpretation of the new parameterization is shown for three interesting settings: quantifying HIV serodiscordance among couples in Mozambique, concordance in the infection status of two related viruses, and the diagnostic performance of an index test in the field of major depression disorders.

---

## Keywords

Association, asynchrony, concordance, discordance, marginal homogeneity, maximum likelihood, McNemar's test, synchrony

## 1 Introduction

In medical applications as well as in other fields, it is often of interest to examine the 'resemblance' between two or more observations from paired or matched outcomes. Here focus is on two binary paired or matched outcomes. Examples considered in this paper are the HIV status among couples; the infection statuses for the same individual for both the Varicella-Zoster Virus and the Parvo B19-virus, viruses that are similar in their transmission being close contact; and the diagnostic performance of the Whooley questions as a screening tool for depression amongst older adults in primary care. Resemblance can be measured in different ways, depending on the characteristic of interest. It could be represented by an association parameter, such as the Pearson product-moment correlation or, for binary data, by the cross-product ratio or odds ratio. However, often association is not of interest but rather agreement. Agreement and association are in general distinct facets of the joint distribution. Strong agreement requires strong association, but strong association can exist without strong agreement.<sup>1</sup> A well-known measure for agreement is Cohen's kappa, see Agresti<sup>1</sup> for extensions and ways to model agreement.

Measures for association and agreement are typically symmetric and can be misleading if one of the agreeing outcomes is very dominant, such as the (negative, negative) combination in our first example of the HIV status among couples. Indeed, as the majority of pairs agree in being negative, symmetric measures of association or agreement might be high even if there is only a small number of agreeing positive pairs. In the context of measuring synchrony in neuronal firing, Faes *et al.*<sup>2</sup> proposed a new measure of synchrony, the conditional synchrony measure (CSM), which is the probability of two neurons firing together, given that at least one of the two is active. Faes *et al.* state that, although the odds ratio is an attractive association measure with nice mathematical properties (such as the absence of range restrictions, regardless of the marginal probabilities), it is less suitable for quantifying synchrony, due to

---

<sup>1</sup>Interuniversity Institute for Biostatistics and statistical Bioinformatics, Hasselt University, BE3590 Diepenbeek, Belgium

<sup>2</sup>Department of Mathematics and Informatics, Faculty of Sciences, Eduardo Mondlane University, Maputo, Mozambique

<sup>3</sup>Centre for Health Economics Research and Modeling Infectious Diseases, Centre for the Evaluation of Vaccination, Vaccine & Infectious Disease Institute (WHO Collaborating Centre), University of Antwerp, Antwerp, Belgium

### Corresponding author:

Marc Aerts, Interuniversity Institute for Biostatistics and statistical Bioinformatics, Hasselt University, BE3590 Diepenbeek, Belgium

Email: marc.aerts@uhasselt.be

its symmetry treating 0-0 matches of equal importance as 1-1 matches. Similar to the CSM but being rather interested in discordance, Juga *et al.*<sup>3</sup> defined the HIV conditional (sero)discordance measure (CDM) as the conditional probability that the couple is HIV discordant, given that at least one of them, man or woman, is HIV positive.

As noted by Faes *et al.* and Juga *et al.*, a reparameterization of the joint bivariate binary distribution in terms of the marginal “success” probabilities and with the OR parameter replaced by the CSM (or the CDM) does not lead to a permissible joint distribution for the full ranges of all parameters, as the Fréchet bounds can be violated.<sup>4</sup> This puts constraints on the parameters of the joint bivariate distribution which are difficult to translate to the regression parameters when introducing dependency of the parameters on risk factors and other covariates. Moreover the constraints hinder fitting the models, leading to computational issues such as convergence problems. The objective of this paper is to solve this non-permissibility problem, and to introduce an alternative parameterization guaranteeing a permissible distribution for all combination of values. The alternative parameterization, no longer including the marginal success probabilities, is shown to be of interest on its own, and offers additional insights for particular applications.

In the next section, three settings with illustrative datasets are introduced. Then the new measures and the new parameterization are presented and covariate models for the different parameters and maximum likelihood inference is briefly described. The illustrative datasets are analyzed using the new parameterization and the paper ends with final conclusions, considerations and ideas about further research.

## 2 Applications and datasets

In the following subsections three different settings and specific data examples are introduced in the field of disease control and prevention.

### 2.1 HIV serodiscordance among couples in Mozambique

We consider the same setting as in Juga *et al.*<sup>3</sup>, based on data from the 2009 National Survey of Prevalence, Risk Behavioural and Information about HIV and AIDS (INSIDA<sup>5</sup>). This survey was a cross-sectional two-stage survey, carried out by the National Institute of Health in collaboration with the National Bureau of Statistics of Mozambique. The objective is to model HIV serodiscordance among couples as a function of different risk factors and other covariates.

Let  $y_{ij} = (y_{ij1}, y_{ij2})$  denote the HIV status (1 if positive, 0 if negative) of a (female,male)-couple  $j = 1, \dots, n_i$  in enumeration area (EA)  $i$  with  $n_i$  sampled couples,  $i = 1, \dots, N$ . For the INSIDA data, the total number of EAs is  $N = 270$  and  $\sum_{i=1}^N n_i = 2159$  is the total number of couples. Expressing

that the covariates  $x_{1,ij}$ ,  $x_{2,ij}$  and  $x_{3,ij}$  can be possibly different subvectors of the full covariate/factor vector  $x_{ij}$ , Juga *et al.*<sup>3</sup> fitted several joint models for both marginal probabilities to be HIV positive, complemented with a new conditional (sero)discordance measure CDM (defined in Section 3):

$$\begin{aligned}\text{logit}(P\{y_{ij1} = 1\}) &= \beta_1^T x_{1,ij} + b_{1,i}, \\ \text{logit}(P\{y_{ij2} = 1\}) &= \beta_2^T x_{2,ij} + b_{2,i}, \\ \text{logit}(\text{CDM}) &= \beta_3^T x_{3,ij} + b_{3,i},\end{aligned}\tag{1}$$

where

$$(b_{1,i}, b_{2,i}, b_{3,i}) \sim N_3(0, \Sigma),\tag{2}$$

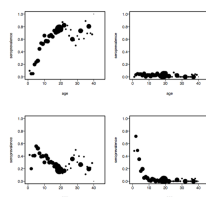
are distributed as a trivariate normal distribution with mean zero-vector and covariance matrix  $\Sigma$ . Note that  $x_{\ell,ij}$  ( $\ell = 1, 2, 3$ ) are vectors of covariates, some of these covariates are specific for the male/female individual, some specific for the couple, and some are at the level of the province. They considered different choices for the covariance matrix  $\Sigma = V R V$  (including full/partial correlated, full/partial shared, full/partial equal, independent). From their final model, Juga *et al.*<sup>3</sup> concluded that the HIV prevalence for the province where a couple was located as well as the union number for the woman within a couple are factors associated with HIV serodiscordance.

As will be discussed in Section 3, the parameterization with both HIV marginal probabilities and the conditional discordance measure CDM does not satisfy the Fréchet inequalities<sup>4</sup>, causing computational difficulties with some of the models. In Section 5.1 we will reanalyze these data with the same type of models, but based on our new parameterization as introduced in Section 3.2.

## 2.2 Varicella Zoster Virus and Parvo B19 concordance

The Varicella-Zoster Virus (VZV) and the Parvo B19-virus (B19) are similar in that transmission occurs during close contacts. The contact rate and the infectiousness of the pathogen determine the spread of the infection in a population. It has been shown that the contact rate depends on age through heterogeneity in mixing of individuals from different age-classes. Several approaches have been proposed to model multi-sera data. Hens *et al.*<sup>6</sup> used a marginal model (bivariate Dale model<sup>7</sup> with odds ratio as association parameter) and conditional models (modelling one infection status conditional on the other) to model the multi-sera VZV-B19 data from Belgium. Hens *et al.*<sup>8</sup> studied the behaviour of the bivariate-correlated gamma frailty model for cross-sectionally collected serological data on hepatitis A and B.

Here we reanalyze the Belgian VZV-B19 serological data. In a period from November 2001 until March 2003, 2381 serum samples in Belgium were collected and consecutively tested for VZV and B19.<sup>9</sup> Together with the test result for VZV and B19, gender and age of the individuals were recorded.



**Figure 1.** VZV and B19 data, as function of age. Proportion of samples that tested positive on both VZV and B19 (top left panel), that tested positive on B19 only (top right panel), that tested positive on VZV only (lower left panel), and that tested negative on both viruses (lower right panel), based on a cross-sectional survey in Belgium anno 2001-2003. The size of the dots is proportional to the number of serum samples collected in the corresponding age category.

Samples from children under 6 months were omitted, as test results are driven by maternal antibodies in this early stage of life. The maximum age of 40 years was fixed by design; it was considered not important to test for older ages given that it concerns childhood infections.

Figure 1 depicts the bivariate distribution of VZV and B19 as a function of age. In Section 5.2 we will propose and discuss new measures to provide other and new insights in the joint occurrence of both infections, as function of age and gender.

### 2.3 Diagnostic performance and concordance of the Whooley questions

Based on a cross-sectional validation study, conducted with 766 patients aged  $\geq 75$  from UK primary care and recruited via 17 general practices based in the North of England during the pilot phase of a randomized controlled trial, Bosanquet *et al.*<sup>10</sup> assessed the diagnostic performance of the Whooley questions (Whooley *et al.*<sup>11</sup>) as a screening tool for major depression disorder (MDD) amongst older adults in UK primary care. Sensitivity, specificity and likelihood ratios comparing the index test (two Whooley questions) for an MDD-diagnosis were ascertained by the reference standard Mini International Neuropsychiatric Interview (MINI<sup>12</sup>). Participants completed a self-reported, written version of the index test, the Whooley questions: (WQ1) During the past month, have you often been bothered by feeling down, depressed, or hopeless? (yes=1/no=0); (WQ2) During the past month, have you often been bothered by little interest or pleasure in doing things? (yes=1/no=0). In the standard method of scoring the Whooley questions, participants who respond yes to at least one of the two questions were classified as screening positive for depression.

Table 1 shows a  $2 \times 2$  table cross-classifying the index test (positive if being positive for at least one Whooley question) with MINI as the golden standard reference (GSR), as well as tables cross-classifying both Whooley questions against each other, unconditionally and conditional on the GSR

**Table 1.** Whooley questions data. Left table: index test versus golden standard reference. Right table: Whooley Question 2 versus Whooley Question 1, unconditionally and conditional on the golden standard reference.

Index Test			WQ2						
GSR	0	1	WQ2		GSR=0				GSR=1
			WQ1	0	1	0	1	0	1
0	458	273	0	460	41	458	40	2	1
1	2	33	1	95	170	91	142	4	28

status. Concordance between index and golden standard reference, and concordance and discordance between both Whooley questions is of interest and will be discussed in Section 5.3.

### 3 Measuring con(dis)cordance and (a)synchrony

First we briefly review existing measures, including the conditional synchrony measure. Next the new parameterization is introduced, discussed and relations with other parameters are examined. A final subsection focuses on the estimation of the new parameters by maximum likelihood.

#### 3.1 Existing measures

Consider a bivariate binary outcome  $y = (y_1, y_2)$  and a (possibly multivariate) covariate  $x$  and let

$$\pi_{k\ell}(x) = P(y_1 = k, y_2 = \ell | x), \quad k, \ell = 0, 1, \quad (3)$$

with  $0 < \pi_{k\ell}(x) < 1$  and  $\sum_{k,\ell=0}^1 \pi_{k\ell}(x) = 1$ , denote the conditional joint distribution of  $y$  given  $x$  (dependency of  $x$  suppressed from notation, if not relevant), with marginal conditional (success) probabilities  $\pi_{1+}(x) = \pi_{10}(x) + \pi_{11}(x)$  and  $\pi_{+1}(x) = \pi_{01}(x) + \pi_{11}(x)$ . Many parameterizations are theoretically possible, but in many cases interest goes in the effect of  $x$  on the marginal probabilities (most often with a logit link allowing an odds ratio interpretation), complemented with an association parameter, such as the correlation  $\rho = (\pi_{11} - \pi_{1+}\pi_{+1}) / \sqrt{\pi_{1+}(1 - \pi_{1+})\pi_{+1}(1 - \pi_{+1})}$ , (with a Fisher-z link) or the odds ratio  $\phi = (\pi_{00}\pi_{11}) / (\pi_{01}\pi_{10})$  (with a log link).

But such association measures are not the target parameter of interest in case interest goes to con(dis)cordance or (a)synchrony. In the context of measuring synchrony in neuronal firing, Faes *et al.*<sup>2</sup> stated that the odds ratio is less suitable to quantify synchrony due to its symmetry, treating 0-0 matches of equal importance as 1-1 matches, and proposed a new measure of synchrony, the conditional

synchrony measure CSM, defined as

$$\text{CSM}(x) = \frac{\pi_{11}(x)}{\pi_{10}(x) + \pi_{01}(x) + \pi_{11}(x)}, \quad (4)$$

being the probability of two neurons firing together, given that at least one of the two is active. In order to model HIV serodiscordance among couples in Mozambique, Juga *et al.*<sup>3</sup> introduced the conditional (sero)discordance measure  $\text{CDM}(x) = 1 - \text{CSM}(x)$  and showed that the CDM measure is a more direct and relevant measure to study the effects of risk factors.

A limitation of this parameterization, with the marginal probabilities combined with  $\text{CSM}(x)$  or  $\text{CDM}(x)$ , is that not all values of  $\pi_{1+}(x)$ ,  $\pi_{+1}(x)$  and of  $\text{CSM}(x)$  (or  $\text{CDM}(x)$ ) result in a permissible joint distribution  $\pi_{k\ell}(x)$ . Indeed, the Fréchet bounds<sup>4</sup> need to hold: given values  $0 < \pi_{1+}(x) < 1$  and  $0 < \pi_{+1}(x) < 1$ , a permissible joint distribution is only obtained if

$$\max\left\{1 - \frac{\pi_{1+}(x)}{\pi_{+1}(x)}, 1 - \frac{\pi_{+1}(x)}{\pi_{1+}(x)}\right\} \leq \text{CDM}(x) \leq \min\{\pi_{1+}(x) + \pi_{+1}(x), 1\}. \quad (5)$$

When modelling the dependency on  $x$  and possibly including additional random effect structures on all three parameters ( $\pi_{+1}$ ,  $\pi_{1+}$  and CDM or CSM) to account for additional heterogeneity, the constraints (5) may cause problems when fitting some models (computational issues, non-convergence, ...).

### 3.2 New measures and new parameterization

While it is common to include the two marginal probabilities as part of a model parameterization, particular alternative parameters might be of interest too and might shed more light on the research questions at hand. In the three applications of interest, focus is not on the marginal probabilities. The new parameterization proposed here abandons the common starting point of adding the parameter of interest to the marginal probabilities, but takes an opposite approach: next to the CSM or CDM measure, which other parameters of interest can be introduced to obtain a complete parameterization of the joint distribution?

As a first parameter define the conditional probability that  $y_1$  is positive, given that both disagree:

$$\pi = P(y_1 = 1 | y_1 \neq y_2) = \frac{\pi_{10}}{\pi_{10} + \pi_{01}}, \quad (6)$$

or, alternatively,  $1 - \pi$  being the probability that  $y_2$  is positive in a disagreeing pair  $(y_1, y_2)$ . This parameter focuses on disagreeing pairs and the probability that one is more dominant than the other.

Note that  $\pi = 0.5$  implies symmetry ( $\pi_{k\ell} = \pi_{\ell k}$ ) and hence marginal homogeneity ( $\pi_{1+} = \pi_{+1}$ ), and  $\pi > 0.5$  implies that  $\pi_{1+} > \pi_{+1}$ . Actually  $\pi$  is the central parameter in McNemar's (exact) test



for matched pairs with, under the null hypothesis of marginal homogeneity,  $\hat{\pi} \sim \text{Bin}(n^*, 0.5)$  with  $n^*$  the total number of disagreeing pairs (see e.g.<sup>1</sup>). This implies that, although the two marginal success probabilities  $\pi_{1+}, \pi_{+1}$  are no longer model parameters, their equality  $\pi_{1+} = \pi_{+1}$  can still be directly tested, while accounting for covariate effects. Actually, instead of the parameter  $\pi$ , one could use the relative difference

$$2\pi - 1 = \frac{\pi_{10} - \pi_{01}}{\pi_{10} + \pi_{01}} = \frac{\pi_{1+} - \pi_{+1}}{\pi_{10} + \pi_{01}}.$$

In the sequel we will use definition (6) and refer to it as the (marginal) homogeneity parameter.

As a second parameter, define (as before) the “positive” conditional synchrony measure CSM, being the probability that both are agreeing (both positive) given that at least one is positive:

$$\sigma_+ = P(y_1 = y_2 | y_1 + y_2 \geq 1) = \frac{\pi_{11}}{\pi_{10} + \pi_{01} + \pi_{11}}, \quad (7)$$

or, alternatively, the (positive) CDM, now denoted as  $\delta_+ = 1 - \sigma_+$ .

Finally, define the third parameter as the “negative” conditional synchrony measure, being the probability that both are agreeing (both negative), given that at most one is positive:

$$\sigma_- = P(y_1 = y_2 | y_1 + y_2 \leq 1) = \frac{\pi_{00}}{\pi_{00} + \pi_{10} + \pi_{01}}. \quad (8)$$

Again, the third parameter can also be defined as (negative) CDM, being  $\delta_- = 1 - \sigma_-$ .

The homogeneity parameter  $\pi$  determines the relative ratio of the off-diagonal probabilities of disagreement, independently of the values of  $\sigma_+$  and  $\sigma_-$ , whereas the parameters  $\sigma_+, \sigma_-$  (or alternatively  $\delta_+, \delta_-$ ) focus on the diagonal probabilities of agreement, and

$$\pi_{11}/\pi_{00} = \left(\frac{\sigma_+}{1 - \sigma_+}\right) / \left(\frac{\sigma_-}{1 - \sigma_-}\right).$$

The measure  $\sigma_+$  tends to 1 if and only if  $\sigma_-$  does so. The parameters  $\delta_+, \delta_-, \sigma_+, \sigma_-$  are invariant for switching  $y_1$  with  $y_2$ , but  $\pi$  will switch to  $1 - \pi$ . Depending on the application and the particular parameters of interest, one can opt for a particular combination, e.g.  $\pi, \delta_+$  and  $\sigma_-$ . The joint probabilities  $\pi_{k\ell}$  can easily be expressed in terms of the new parameters, e.g. in terms of  $\pi, \sigma_+$  and  $\sigma_-$

$$\begin{aligned} \pi_{00} &= \frac{\sigma_-(1 - \sigma_+)}{1 - \sigma_- \sigma_+}, & \pi_{10} &= \frac{(1 - \sigma_-)(1 - \sigma_+)}{1 - \sigma_- \sigma_+} \pi, \\ \pi_{01} &= \frac{(1 - \sigma_-)(1 - \sigma_+)}{1 - \sigma_- \sigma_+} (1 - \pi), & \pi_{11} &= \frac{\sigma_+(1 - \sigma_-)}{1 - \sigma_- \sigma_+}. \end{aligned}$$

It can also readily be shown that, for any combination of values for the three conditional probabilities  $0 < \pi, \sigma_-, \sigma_+ < 1$ , the Fréchet bounds are satisfied, and we obtain a permissible joint distribution  $\pi_{k\ell}$ .

The marginal success probabilities can be written as

$$\pi_{1+} = \frac{(1 - \sigma_-)(1 - (1 - \sigma_+)(1 - \pi))}{1 - \sigma_- \sigma_+}, \quad \pi_{+1} = \frac{(1 - \sigma_-)(1 - (1 - \sigma_+)\pi)}{1 - \sigma_- \sigma_+},$$

and the odds ratio as

$$\phi = \left(\frac{\sigma_-}{1 - \sigma_-}\right)\left(\frac{\sigma_+}{1 - \sigma_+}\right)\frac{1}{\pi(1 - \pi)}. \quad (9)$$

Identity (9) shows that the odds ratio  $\phi$  decomposes in three factors, each related to one of the three new parameters. The association in terms of the odds ratio increases multiplicatively with the odds of both synchrony measures  $\sigma_-$  and  $\sigma_+$ , and converges to infinity as the homogeneity parameter  $\pi$  tends to 0 or 1. The minimal value of  $\phi$ , for fixed values of  $\sigma_-$  and  $\sigma_+$ , is obtained for  $\pi = 0.5$ , corresponding to marginal homogeneity. Of course, this factorization is not helping in characterizing independence, but independence is not of interest in our settings of interest.

The relation with Cohen's kappa measure of agreement takes the form

$$\kappa = \frac{\sigma_-(1 - \sigma_+) + \sigma_+(1 - \sigma_-) - K(\pi, \sigma_-, \sigma_+)}{(1 - \sigma_- \sigma_+) - K(\pi, \sigma_-, \sigma_+)},$$

for a rather complicated function  $K(\pi, \sigma_-, \sigma_+)$ . But an immediate consequence is that there is perfect agreement according to Cohen's kappa,  $\kappa = 1$ , if and only if there is perfect negative ( $\sigma_- = 1$ ) or perfect positive ( $\sigma_+ = 1$ ) conditional synchrony.

A bit different and more “asymmetric setting” is that of measuring the accuracy of diagnostic tests. Assume  $y_1$  represents the true disease status, and  $y_2$  another alternative test. Sensitivity  $\text{Se} = P(y_2 = 1|y_1 = 1)$  and specificity  $\text{Sp} = P(y_2 = 0|y_1 = 0)$  relate to the new parameters as

$$\text{Se} = \frac{\sigma_+}{\sigma_+ + (1 - \sigma_+)\pi}, \quad \text{Sp} = \frac{\sigma_-}{\sigma_- + (1 - \sigma_-)(1 - \pi)},$$

and for the positive predictive value  $\text{PPV} = P(y_1 = 1|y_2 = 1)$  and the negative predictive value  $\text{NPV} = P(y_1 = 0|y_2 = 0)$  it holds that

$$\text{PPV} = \frac{\sigma_+}{\sigma_+ + (1 - \sigma_+)\pi}, \quad \text{NPV} = \frac{\sigma_-}{\sigma_- + (1 - \sigma_-)\pi}.$$

First of all, note that  $\text{Se}$  and  $\text{PPV}$  do not depend on  $\sigma_-$  and  $\text{Sp}$  and  $\text{NPV}$  do not on  $\sigma_+$ . As to be expected,  $\text{Se}$  and  $\text{NPV}$  decrease whereas  $\text{Sp}$  and  $\text{PPV}$  increase with the homogeneity parameter  $\pi$ . If  $\pi$  is very close

to 1,  $Se \approx \sigma_+$  and  $Sp \approx 1$ . Furthermore, in that case,  $PPV \approx 1$ , and  $NPV \approx \sigma_-$ . Similarly, if  $\pi$  is very close to 0,  $Se \approx NPV \approx 1$ ,  $Sp \approx \sigma_-$  and  $PPV \approx \sigma_+$ .  $Se$  and  $PPV$  increase with  $\sigma_+$  and  $Sp$  and  $NPV$  increase with  $\sigma_-$ . All of them tend to 1 whenever  $\sigma_+$  tends to 1 (and thus  $\sigma_-$  tends to 1).

The diagnostic odds ratio

$$DOR = \left( \frac{Se}{1 - Se} \right) \left( \frac{Sp}{1 - Sp} \right) = \left( \frac{PPV}{1 - PPV} \right) \left( \frac{NPV}{1 - NPV} \right) = \left( \frac{\sigma_+}{1 - \sigma_+} \right) \left( \frac{\sigma_-}{1 - \sigma_-} \right) \frac{1}{\pi(1 - \pi)} \quad (10)$$

is used as a measure of the effectiveness of a diagnostic or screening test. It is independent of prevalence, it is a single indicator of test performance, ranging from zero to infinity and higher values (above 1) are indicative of better test performance<sup>13</sup>. Equation (10) decomposes the DOR in three factors: given the value of the marginal homogeneity parameter  $\pi$ , higher values of the DOR correspond to higher values of one or both synchrony measures  $\sigma_+, \sigma_-$ .

Although our application as introduced in Section 2.3 is based on a cross-sectional study, allowing to estimate the disease prevalence by the case study prevalence, this might be not the case for other study designs. Consider for instance the case-control design, with data about a screening test result for the diseased and non-diseased subpopulation (as defined by a golden standard or reference test). Such a design allows the estimation of the sensitivity and specificity, but not the prevalence of the disease. The formulas

$$\begin{aligned} \pi &= \frac{(1 - Se)Prev}{(1 - Se)Prev + (1 - Sp)(1 - Prev)}, \\ \sigma_+ &= \frac{Se \times Prev}{1 - Sp(1 - Prev)}, \\ \sigma_- &= \frac{Sp(1 - Prev)}{1 - Se \times Prev}, \end{aligned}$$

with  $Prev = \pi_{1+}$  the prevalence of the disease, show the dependency of the homogeneity parameter and both conditional synchrony measures on the disease prevalence. These formulas allow us to combine data with knowledge about the disease prevalence (from other data or literature) to estimate or to model the parameters  $\pi, \sigma_+$  and  $\sigma_-$  for a case-control study (typically in the Bayesian paradigm).

## 4 Estimation and inference

Consider quadrinomial observations  $y_{jk\ell} = I(y_{j1} = k)I(y_{j2} = \ell)$  for  $k = 0, 1$  and  $\ell = 0, 1$ , for  $j = 1, \dots, n$ . Then, by the orthogonality  $\pi \perp (\sigma_-, \sigma_+)$ , the quadrinomial likelihood, with  $\pi_{00} + \pi_{01} + \pi_{10} +$

$$\pi_{11} = 1,$$

$$\mathcal{L}_{(4)}(\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11}) = \prod_{j=1}^n \pi_{00}^{y_{j00}} \pi_{01}^{y_{j01}} \pi_{10}^{y_{j10}} \pi_{11}^{y_{j11}},$$

factorizes into a trinomial and binomial likelihood  $\mathcal{L}_{(4)} = \mathcal{L}_{(3)} \times \mathcal{L}_{(2)}$  with

$$\begin{aligned} \mathcal{L}_{(3)}(\sigma_-, \sigma_+) &= \prod_{j=1}^n \pi_{00}^{y_{j00}} \pi_{11}^{y_{j11}} (\pi_{01} + \pi_{10})^{y_{j01} + y_{j10}}, \\ &= \frac{(1 - \sigma_-)(1 - \sigma_+)}{(1 - \sigma_- \sigma_+)} \prod_{j=1}^n \left( \frac{\sigma_-}{1 - \sigma_-} \right)^{y_{j00}} \left( \frac{\sigma_+}{1 - \sigma_+} \right)^{y_{j11}}, \end{aligned}$$

with

$$\mathcal{L}_{(2)}(\pi) = \prod_{j=1}^n \pi^{y_{j10}} (1 - \pi)^{y_{j01}}.$$

In case no parameters are common to the models for  $(\sigma_-, \sigma_+)$  and  $\pi$ , both likelihoods can be maximized separately, and, in case interest only goes to the conditional synchrony measures, the disagreeing observations can be collapsed and it suffices to only maximize the trinomial likelihood.

The dependency of the three conditional probabilities  $\pi$ ,  $\sigma_+$  and  $\sigma_-$  (or any other eligible combination of interest from the sets  $\{\sigma_+, \delta_+\}$  and  $\{\sigma_-, \delta_-\}$ ) on covariates can be modelled with three components:

$$h_1(\pi_j) = \beta_1^T x_{1j}, \quad h_2(\sigma_{+j}) = \beta_2^T x_{2j}, \quad h_3(\sigma_{-j}) = \beta_3^T x_{3j}, \quad (11)$$

where  $h_1, h_2$  and  $h_3$  are link functions (logit, probit, cloglog, ...). We will focus on the logit link as it allows a more appealing interpretation of covariate effects in terms of odds ratios. The model components (11) can be embedded in different frameworks of estimation and inference; we will opt for full maximum likelihood.

Note that, when using the relative difference parameter  $2\pi - 1$ , ranging from -1 to 1, the logit link is not appropriate but rather a Fisher-z link would be in order. The identity, with  $\Delta_r = 2\pi - 1$ ,

$$\log\left(\frac{\pi}{1 - \pi}\right) = -\log\left(\frac{1 - \pi}{\pi}\right) = \log\left(\frac{1 + \Delta_r}{1 - \Delta_r}\right),$$

implies that models for  $\pi$ ,  $1 - \pi$  and  $\Delta_r$  with respective links are identical (only opposite slopes for the models for  $1 - \pi$ ). Depending on the application at hand, one might be more interested in interpreting the estimates in terms of  $\pi$ ,  $1 - \pi$  or  $\Delta_r$ . For the latter choice, a zero intercept would reflect marginal homogeneity for all covariate values equal to 0, and the effect of a covariate as represented by the estimated slope would reflect non-homogeneity in one or the other direction: a positive slope would

indicate a higher marginal probability of success in the first variable, a negative slope would indicate a higher marginal probability of success in the second variable.

## 5 Applications

In this section we revisit the three applications introduced in Section 2 and show how in each example model (11) can be formulated and we illustrate the use and interpretation of the three conditional probabilities  $\pi$ ,  $\sigma_+$  and  $\sigma_-$  (or variations thereof). Data analyses were performed in SAS version 9.4 using PROC NLMIXED (exemplifying code in the supplemental material).

### 5.1 Modelling HIV serodiscordance among couples in Mozambique

We reanalyse the HIV data introduced in Section 2.1 based on model components for i) the (homogeneity) probability  $\pi_{ij}$  that the female partner of couple  $j$  in EA  $i$  is HIV positive, given that both partners differ in their HIV status; ii) the probability  $\delta_{+ij}$  that only one is HIV positive, given that at least one of the two partners is positive (positive serodiscordance); and iii) the probability  $\sigma_{-ij}$  that both are negative given that at most one of the two partners is positive (negative seroconcordance):

$$\begin{aligned}\text{logit}(\pi_{ij}) &= \beta_1^T x_{1,ij} + b_{\pi,i} \\ \text{logit}(\delta_{+ij}) &= \beta_2^T x_{2,ij} + b_{\delta+,i} , \\ \text{logit}(\sigma_{-ij}) &= \beta_3^T x_{3,ij} + b_{\sigma-,i}\end{aligned}$$

with

$$(b_{\pi,i}, b_{\delta+,i}, b_{\sigma-,i}) \sim N_3(0, \Sigma), \quad (12)$$

trivariate normally distributed random EA-effects, with mean zero-vector and covariance matrix  $\Sigma$ .

Most of the sample designs for household surveys such as INSIDA are complex and involve stratification, multistage sampling, and unequal sampling rates, and it is necessary to account for the particular survey design in the statistical analyses using appropriate weights. We followed the same approach as Juga *et al.*<sup>3</sup>. For more details on the calculation of the weights as used in our analyses, we refer to Juga *et al.*

After following the same model building procedure as in Juga *et al.*, the best fitting final model had no random EA-effect  $b_{\pi,i}$  on parameter  $\pi_{ij}$  and correlated random EA-effects  $(b_{\delta+,i}, b_{\sigma-,i})$  on  $(\delta_{+ij}, \sigma_{-ij})$ . This model has an AIC value of 2509.2, considerably improving the fit of the best model in Juga *et al.*<sup>3</sup> (AIC=2571.8 for a partial-equal random effects type of model). All models converged, in contrast to the

**Table 2.** HIV serodiscordance example. Estimates (standard errors) of the final model with no random EA-effect for parameter  $\pi_{ij}$  and correlated random EA-effects ( $b_{\delta_+,i}, b_{\sigma_-,i}$ ) on  $(\delta_{+ij}, \sigma_{-ij})$ , with variance components  $\sigma_{\delta_+}^2$ ,  $\sigma_{\sigma_-}^2$  and correlation  $\rho_{\delta_+\sigma_-}$ .

Effect	$\pi$	$\delta_+$	$\sigma_-$
Intercept	-0.82(0.40)*	2.48(0.69)*	3.65(0.37)*
HIV prevalence			
5 – 15%	-	-0.93(0.53)	-0.88(0.28)*
> 15%	-	-1.09(0.56)	-1.64(0.32)*
Union number woman			
More than once	0.57(0.28)*	-0.74(0.26)*	-0.49(0.17)*
STI man			
Yes	-	-1.41(0.42)*	-
Condom use woman			
Not used	1.87(0.79)*	-	-
Wealth index			
Poorer	-	1.13(0.37)*	0.42(0.22)
Middle	-	0.20(0.37)	0.52(0.24)*
Variance components			
$\sigma_{\delta_+}^2$	-	1.29(0.43)†	
$\sigma_{\sigma_-}^2$	-	0.72(0.20)†	
$\rho_{\delta_+\sigma_-}$	-	-0.62(0.14)*	

\* Significant at 5% level based on a likelihood ratio test, † Significant at 5% level, using  $\chi_{0,1}^2$ -mixture  
a - sign in a column refers to a non-significant effect at 5% after which the covariate was deleted (stepwise)

experiences of Juga *et al.*, who reported that the models with independent random effects and full or partial random effects did not convergence.

Table 2 shows the estimates of the final model. The following covariates appear in the final model with fixed effects: ‘HIV prevalence’ is the prevalence of HIV at the couple’s residence at the level of the province, categorized into three categories using cutpoints of 5% and 15% and with 0-5% as reference category; ‘Union number woman’ refers to whether the woman has been married or lived with a man once (reference category) or more than once; ‘STI man’ is yes if the man answered yes to any of 3 questions about symptoms of sexually transmitted infections STIs (no is reference); ‘Condom use woman’ refers to whether the female respondent of the couple used a condom the last time she had sexual intercourse with the other partner (yes is reference); ‘Wealth index’ refers to the economic status of the couple with three categories Poorer, Middle and Richer (reference). For more details, see<sup>14</sup>.

Comparing our results with those of Juga *et al.*<sup>3</sup> and focusing on the common conditional serodiscordance parameter  $CDM = \delta_+$  in both models, similar effects were obtained for the effect of

‘HIV prevalence’ and ‘Union number woman’. Additional to those effects, our new model identified a significant effect of ‘STI man’ and ‘Wealth index’: the probability for both partners of a couple to differ in HIV status, given that at least one of both is positive, decreases in case the man has reported STIs and increases in case the couple’s wealth index is poorer rather than richer.

The negative synchrony depends on the ‘HIV prevalence’ (the higher the prevalence the lower the synchrony), the ‘Union number woman’ (lower synchrony in case the woman has been married or lived with a man more than once) and the ‘Wealth index’ (higher synchrony for the middle category).

Finally for the homogeneity parameter  $\pi$  it can be observed that marginal homogeneity (both partners having the same probability to be HIV positive) does not hold in case the woman has been married or lived with a man only once and the man has indicated to have no STI symptoms (95% CI [0.166,0.491] for  $\pi$ , implying the probability to be HIV positive is lower for the woman) and in case the woman has been married or lived with a man more than once and the man has indicated to have STI symptoms (95% CI [0.517,0.960] for  $\pi$ , implying the probability to be HIV positive is higher for the woman).

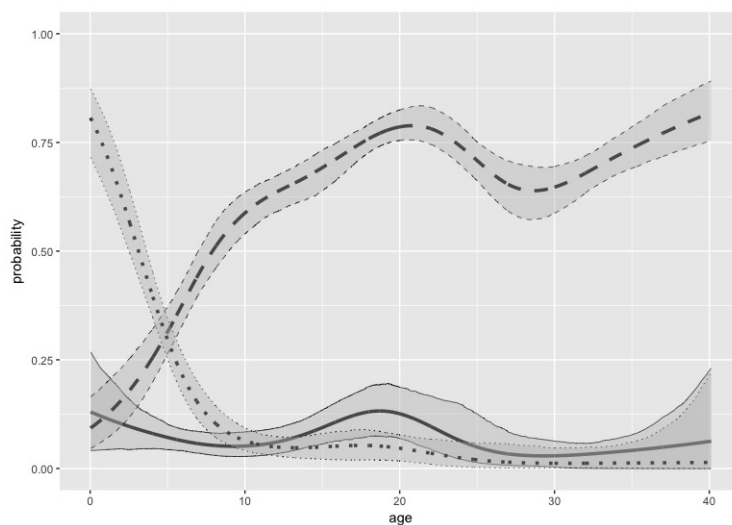
## 5.2 Varicella Zoster Virus and Parvo B19 concordance

**Table 3.** VZV and B19 data. B19 versus VZV, unconditionally and conditional on gender.

	VZV		VZV			
			Female		Male	
B19	0	1	0	1	0	1
0	174	725	91	357	83	368
1	57	1425	28	733	29	692

Let  $y_1$  refer to the B19 infection status and  $y_2$  to the VZV infection status of the same individual. We are interested in the dependency of  $\pi$ ,  $\sigma_+$  and  $\sigma_-$  on age and gender. Table 3 shows the observed bivariate frequencies of  $(y_1, y_2)$ , unconditionally and conditionally on gender. Model (11) is applied with an indicator for gender and a cubic spline for age restricted to be linear before the first knot and after the last knot, with knots located at the 9 deciles of the observed age distribution. Gender was not significant for any parameter (p-values 0.8762, 0.2002, 0.8641 for  $\pi$ ,  $\sigma_+$  and  $\sigma_-$  respectively). Age has a significant effect on all three parameters  $\pi$ ,  $\sigma_+$  and  $\sigma_-$  (p-values 0.004,  $<0.0001$ ,  $<0.0001$  respectively).

Figure 2 visualizes the dependencies on age. For an individual who is positive for one and negative for the other virus, the probability is lowest that he is positive for B19 (about 0.10) across all ages. Marginal homogeneity clearly does not hold, for any age (p-value  $< 0.00001$ ). If an individual is positive for at least one virus, the probability  $\sigma_+$  that he is positive for both increases from 0.10 to about 0.80 with a strange bump at the age of 20. The negative conditional synchrony  $\sigma_-$ , being the probability he



**Figure 2.** VZV and B19 example. Plot of the fitted  $\pi$  (solid line),  $\sigma_+$  (dashed line) and  $\sigma_-$  (dotted line) as function of age and based on the final model, together with 95% bootstrap percentile intervals using 1000 nonparametric bootstrap samples.

is negative for both given that he is at most positive for one of the virus, decreases rapidly during the first 10 years of life, from about 0.80 to about 0.05. The bump around the age of 20, visible more or less for all curves, is caused by an “artefact” in the data, in the sense that the prevalence to be positive for one only, or for both is expected to be monotone as a function of age (the older you are the higher the probability of ever been infected by one of the diseases). Already Figure 1 shows that this expected monotonicity constraint is violated by the patterns shown in the data. This phenomenon and possible model modifications and extensions (covering e.g. waning immunity) have been presented and discussed in e.g. Abrams and Hens<sup>15</sup>.

Using a bivariate Dale model and splines for the effect of age, Hens *et al.*<sup>6</sup> could not reject the null hypothesis of a constant OR (p-value 0.37). The estimated age and gender independent OR equaled 2.11 with 95% confidence interval (1.45, 3.23). This is another example showing that measures for agreement and for dependency can behave quite differently.

### 5.3 Diagnostic performance and concordance of the Whooley questions

Let  $y_1$  refer to the GSR and  $y_2$  to the index test of the same individual. Bosanquet *et al.*<sup>10</sup> report a high sensitivity of 0.943 (95% CI [0.808, 0.993]) and a modest specificity of 0.627 (95% CI [0.590, 0.662]) for the index test (at least one Whooley question positive). Table 4 shows estimates for all parameters of



interest. Note that the homogeneity parameter  $\hat{\pi} = 0.007$  is very small, implying marginal heterogeneity and that, as noticed in general,  $Sp \approx \sigma_-$ ,  $PPV \approx \sigma_+$  and  $Se \approx NPV \approx 1$ . So, if index and reference test disagree, it is highly unlikely that the reference test is the positive one. If at least one of the index or reference test is positive, is unlikely that both are positive (probability of 0.107), so low positive synchrony or concordance. If at least one is negative, they are both negative with probability 0.625.

Next it is interesting to get more insight in the (con/dis)cordance between both Whooley questions,

**Table 4.** Whooley questions example. Table of Index test  $\times$  GSR: point estimates, standard error estimates and 95% confidence intervals for sensitivity, specificity, positive predictive value, negative predictive value, probability  $\pi$ , positive and negative conditional synchrony.

	Est	se	95% CI
Se	0.943	0.0392	(0.866, 1.000)
Sp	0.627	0.0179	(0.591, 0.662)
PPV	0.108	0.0177	(0.073, 0.143)
NPV	0.996	0.0031	(0.990, 1.002)
$\pi$	0.007	0.0051	(0.000, 0.017)
$\sigma_+$	0.107	0.0176	(0.073, 0.142)
$\sigma_-$	0.625	0.0179	(0.591, 0.660)

defining the index test, and how this depends on the GSR status. So, let  $y_1$  now refer to WQ1 and  $y_2$  to WQ2. Table 5 shows the estimated parameters, unconditionally and conditionally on the GSR status. SAS code for the model with parameters modelled as a function of the GSR status (MINI) is included in the supplemental material.

**Table 5.** Whooley questions example. Table of WQ1  $\times$  WQ2, unconditionally and conditionally on GSR status: point estimates, standard error estimates and 95% confidence intervals for probability  $\pi$ , positive and negative conditional synchrony.

	Est	se	95% CI
$\pi$	0.699	0.0394	(0.616, 0.770)
$\sigma_+$	0.556	0.0284	(0.499, 0.610)
$\sigma_-$	0.772	0.0172	(0.736, 0.804)
$\pi$   GSR=0	0.695	0.0402	(0.611, 0.768)
$\sigma_+$   GSR=0	0.520	0.0302	(0.461, 0.580)
$\sigma_-$   GSR=0	0.778	0.0171	(0.742, 0.809)
$\pi$   GSR=1	0.800	0.1789	(0.308, 0.973)
$\sigma_+$   GSR=1	0.849	0.0624	(0.683, 0.936)
$\sigma_-$   GSR=1	0.286	0.1707	(0.072, 0.674)

The homogeneity probability  $\pi$  does not depend on the GSR status (p-val 0.6189). As  $\pi \perp (\sigma_-, \sigma_+)$ , it is not necessary to refit a simplified model with no effect of the GSR status on  $\pi$ . Indeed, fitting such model would lead to exactly the same results for  $\sigma_+ \mid \text{GSR}=0$  or 1 and  $\sigma_- \mid \text{GSR}=0$  or 1, and the GSR independent estimate for  $\pi$  would equal the estimate 0.699 in the collapsed table, being the first line of results in Table 5. So, whether or not the individual is depressed, if the answers on the two WQ's disagree, the probability is about 70% that WQ1 is answered positively (95% CI being [0.616, 0.770], marginal homogeneity does not hold). So, individuals with disagreeing answers on both Whooley questions tend to have been more often bothered by feeling down, depressed, or hopeless, than bothered by little interest or pleasure in doing things, regardless of their GSR status.

But  $\sigma_+$  and  $\sigma_-$  significantly depend on the GSR status, with p-values 0.0011 and 0.0103 respectively. The positive synchrony measure estimate  $\hat{\sigma}_+$  increases from 0.520 to 0.849 implying that, giving that at least one of the WQ's is answered positively, the probability that both are positively answered increases substantially for individuals suffering from a major depression disorder (according to GSR). On the other hand, the negative synchrony measure estimate  $\hat{\sigma}_-$  decreases from 0.778 to 0.286 implying that, giving that at most one of the WQ's is answered positively, the probability that both are negatively answered decreases substantially for individuals suffering from a major depression disorder (according to GSR).

Bosanquet et al.<sup>10</sup> mention in the discussion that the use of the two-item version of the Whooley questions, rather than a three-item version, in which the respondent is asked to state whether they would like help for any difficulties reported, is a potential limitation of their study. However, evidence from recent studies using a third help question, does not provide a conclusive answer whether to include or not the third question. It would be interesting to study in more depth the con- and discordance between all three questions in order to come up with an improved index test.

## 6 Conclusions and discussion

As interest goes to modelling a genuine synchrony/concordance measure rather than a typical association measure such as the odds ratio, the joint distribution of matched pairs of binary data need to be reparametrized accordingly. In this contribution, a new parameterization solved the existing permissibility issue with the conditional synchrony measure and related limitations of fitting appropriate models. This new parameterization is based on two synchrony measures, a positive and negative synchrony (or alternatively discordance) parameter, combined with a marginal homogeneity parameter, leading without any restrictions on any of these parameters to a permissible joint distribution for the matched binary pairs, thus facilitating the fitting of more flexible and appropriate models.

The usefulness of the new approach has been illustrated in three different areas of application in disease control and prevention. In the first application, the positive serodiscordance was the main parameter

of interest, but also the additional negative seroconcordance provides alternative new insights in this field. Whereas the same characteristic (HIV status) is measured for both partners of a couple in the first application, two different characteristics (VZV and B19 infection status) on one and the same individual are available in the second application. The negative and positive synchrony measures provide new information and insights in the joint process of acquiring both diseases having similar transmission routes. In a third, more distinct application, the accuracy of a screening test is to be assessed in relation to the true disease status (or a gold standard). The new synchrony measures allow to investigate the performance of the diagnostic test from another angle, different from but closely related to well-known accuracy measures such as sensitivity, specificity, predictive values, DOR, etc. and future use of these new measures will shed more light on their ultimate value in this particular field of application.

An advantage of the new approach is that, in case interest only goes to both synchrony measures and their models do not share any parameter in common with the model for the marginal homogeneity parameter, the disagreeing observations can be collapsed and the synchrony measures can be modelled by means of a simplified trinomial likelihood. It may be perceived as a disadvantage that the marginal success probabilities (such as the prevalence of one or both diseases) are not directly estimated or modelled as a function of covariates, as in the currently applied parameterizations. On the other hand, the homogeneity parameter still allows to investigate structural differences in both marginal parameters. Moreover, models for the new parameters  $\pi, \sigma_+, \sigma_-$  imply indirect models for the marginal success probabilities using their relation equations.

A first interesting methodological topic for further research is the modification and application of the HIV model of the first example to same-sex couples, as already mentioned by Juga *et al.*<sup>3</sup>. They suggested two approaches to deal with the exchangeability of both partners of a couple. But the parameterization proposed here offers an interesting third option. Indeed only the marginal homogeneity parameter depends on the order of the partners in a couple and would not be interpretable when using a random order (actually one would expect homogeneity in case of a random order). Being orthogonal to the other parameters, it would not affect the estimation of the synchrony measures.

Another interesting extension is to examine synchrony between three or more outcomes, as for instance three related infections in our second example. As the number of parameters grows exponentially with the outcome-dimension, defining a full set of appropriate homogeneity and synchrony parameters needs careful considerations in view of the application of interest. For the last example one often includes an inconclusive category, introducing one or both as a trinomial outcome. An extension in that direction poses interesting challenges. Also in the latter setting, one might like to account for an imperfect GSR by correcting for misclassifications.

## Acknowledgements

The authors would also like to acknowledge the support given by the Mozambican Health Ministry (MISAU) and Demography Health Survey (DHS) Program for providing the INSIDA survey data. The VZV-B19 example was based on a serum sample collected for the European Commission's ESEN2-project.

## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

This study was only possible thanks to the financial support of the Flemish Interuniversity Council (VLIR-UOS) in collaboration with Eduardo Mondlane University (UEM) through the DESAFIO Program. NH acknowledges funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement 682540 TransMID) and the Special Research Fund of Hasselt University.

## References

1. Agresti A. *Categorical Data Analysis*. Third Edition. New York: John Wiley and Sons Inc; 2013.
2. Faes C, Geys H, Molenberghs G, Aerts M, Cadarso-Suárez C, Acuña C, Cano M. A Flexible Method to Measure Synchrony in Neuronal Firing. *Journal of the American Statistical Association*. 2008.103:149-161.
3. Juga AJC, Hens N, Osman N, Aerts M. A flexible method to model HIV serodiscordance among couples in Mozambique. *PLoS ONE* 12(3): e0172959. doi:10.1371/journal.pone.0172959
4. Fréchet M. Sur les tableaux de corrélation dont les marges sont données. *Annals Université Lyon, Section A, Series 3*. 1951.14:53–77.
5. Instituto Nacional de Saúde (INS), Instituto Nacional de Estatística (INE). Inquérito Nacional de Prevalência, Riscos Comportamentais e Informação sobre o HIV e SIDA em Moçambique de 2009. 2010. 310pp (<https://dhsprogram.com/pubs/pdf/ais8/ais8.pdf>).
6. Hens N, Aerts M, Shkedy Z, Theeten H, Van Damme P, Beutels P. Modelling multisera data: the estimation of new joint and conditional epidemiological parameters. *Statistics in Medicine*. 2008.27:2651-2664.
7. Dale JR. Global cross-ratio models for bivariate, discrete, ordered responses. *Biometrics*. 1986.42:909-917.
8. Hens N, Wienke A, Aerts M, Molenberghs G. The correlated and shared gamma frailty model for bivariate current status data: An illustration for cross-sectional serological data. *Statistics in Medicine*. 2009.28:2785-2800.
9. Nardone A, Miller E. Serological surveillance of rubella in europe: European sero-epidemiology network (ESEN2). *Euro-surveillance*. 2004.9(4):5-7.

10. Bosanquet K, Mitchell N, Gabe R, Lewis H, McMillan D, Ekers D, Bailey D, Gilbody S. Diagnostic accuracy of the Whooley depression tool in older adults in UK primary care. *Journal of Affective Disorders*. 2015;182:39-43
11. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression: two questions are as good as many. *Journal of General Internal Medicine*. 1997;12:439-445.
12. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC. The mini-international neuropsychiatric interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*. 1998;59:22-33.
13. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PMM. The diagnostic odds ratio: a single indicator of test performance. *Journal of Clinical Epidemiology*. 2003;56:1129-1135.
14. Fishel JD, Bradley SEK, Young PW, Mbofana F, Botão C. HIV among Couples in Mozambique: HIV Status, Knowledge of Status, and Factors Associated with HIV Serodiscordance. Further Analysis of the 2009 Inquérito Nacional de Prevalência, Riscos Comportamentais e Informação sobre o HIV e SIDA em Moçambique 2009. Calverton, Maryland, USA. ICF International. 2011. 55pp (<https://dhsprogram.com/pubs/pdf/FA71/FA71.pdf>).
15. Abrams S, Hens, N. Modeling individual heterogeneity in the acquisition of recurrent infections: an application to parvovirus B19. *Biostatistics*. 2015;16:129-142.

## Supplemental material

```

/* (CON/DIS)CORDANCE BETWEEN BOTH WHOOLEY QUESTIONS AS A FUNCTION OF
THE GOLDEN STANDARD STATUS */

proc nlmixed data=WORK.DEPRESSION;

/* Starting values for intercepts and slopes */
parms b0_1=0 bA_1=0 b0_2=0 bA_2=0 b0_3=0 bA_3=0;

/* The dependency of the model parameters on the GSR status (MINI) */
eta1=b0_1+bA_1*MINI;
AP10=1/(1 + exp(-eta1));
eta2=b0_2+bA_2*MINI;
PCSM = 1/(1 + exp(-eta2));
eta3=b0_3+bA_3*MINI;
NCSM = 1/(1 + exp(-eta3));

/* The joint probabilities as function of the model parameters */
p11 = (PCSM-NCSM*PCSM)/(1-NCSM*PCSM);
p00 = (NCSM-NCSM*PCSM)/(1-NCSM*PCSM);
p10 = AP10*(1-NCSM)*(1-PCSM)/(1-NCSM*PCSM);
p01 = (1-AP10)*(1-NCSM)*(1-PCSM)/(1-NCSM*PCSM);

/* The model parameters for both statuses of the GSR*/
AP10GS0=1/(1 + exp(-b0_1));
PCSMGS0 = 1/(1 + exp(-b0_2));
NCSMGS0 = 1/(1 + exp(-b0_3));
AP10GS1=1/(1 + exp(-b0_1-bA_1));
PCSMGS1 = 1/(1 + exp(-b0_2-bA_2));
NCSMGS1 = 1/(1 + exp(-b0_3-bA_3));

/* Compute log likelihood for bivariate observations */
ll = depression*pleasure*log(p11) + depression*(1-pleasure)*log(p10)
      + (1-depression)*pleasure*log(p01) + (1-depression)*(1-pleasure)*log(p00);
model ll ~ general(ll);

```

```
estimate 'pi given GSR=0' AP10GS0;  
estimate 'sigma+ given GSR=0' PCSMG0;  
estimate 'sigma- given GSR=0' NCSMG0;  
estimate 'pi given GSR=1' AP10GS1;  
estimate 'sigma+ given GSR=1' PCSMG1;  
estimate 'sigma- given GSR=1' NCSMG1;  
run;
```