Modeling the Accuracy of Three Detection Methods of Grapevine leafroll-associated virus 3 During the Dormant Period Using a Bayesian Approach

Antonio Olmos, Edson Bertolini, Ana B. Ruiz-García, Carmen Martínez, Rosa Peiró, and Eduardo Vidal

First, third, fourth, and sixth authors: Centro de Protección Vegetal y Biotecnología, Instituto Valenciano de Investigaciones Agrarias (IVIA), Ctra. Moncada-Náquera km 4.5, 46113 Moncada, Valencia, Spain; second author: Departamento de Fitossanidade, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul (UFRGS), Avenida Bento Gonçalves 7712, 91450-000 Porto Alegre, RS, Brazil; and fifth author: Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de Valencia (COMAV-UPV), 46022 Valencia, Spain. Current address of sixth author: Technical and Regulatory Affairs Department at SAPEC AGRO ESPAÑA. Accepted for publication 8 January 2016.

ABSTRACT

Olmos, A., Bertolini, E., Ruiz-García, A. B., Martínez, C., Peiró, R., and Vidal, E. 2016. Modeling the accuracy of three detection methods of *Grapevine leafroll-associated virus 3* during the dormant period using a Bayesian approach. *Phytopathology* 106:510-518.

Grapevine leafroll-associated virus 3 (GLRaV-3) has a worldwide distribution and is the most economically important virus that causes grapevine leafroll disease. Reliable, sensitive, and specific methods are required for the detection of the pathogen in order to assure the production of healthy plant material and control of the disease. Although different serological and nucleic acid-based methods have been developed for the detection of GLRaV-3, diagnostic parameters have not been established, and there is no gold standard method. Therefore, the main aim of this work was to determine the sensitivity, specificity, and likelihood ratios of three commonly used methods, including one serological test (double-antibody sandwich enzyme-linked immunosorbent assay [DAS-ELISA]) and two nucleic acid-based techniques (spot and conventional real-time reverse transcription-polymerase chain reaction [RT-PCR]). Latent class models using a Bayesian

approach have been applied to determine diagnostic test parameters and to facilitate decision-making regarding diagnostic test selection. Statistical analysis has been based on the results of a total of 281 samples, which were collected during the dormant period from three different populations. The best-fit model out of the 49 implemented models revealed that DAS-ELISA was the most specific method (value = 0.99) and provided the highest degree of confidence in positive results. Conversely, conventional real-time RT-PCR was the most sensitive method (value = 0.98) and produced the highest degree of confidence in negative results. Furthermore, the estimation of likelihood ratios showed that in populations with low GLRaV-3 prevalence the most appropriate method could be DAS-ELISA, while conventional real-time RT-PCR could be the most appropriate method in medium or high prevalence populations. Combining both techniques significantly increases detection accuracy. The flexibility and power of Bayesian latent class models open new possibilities for the evaluation of diagnostic tests for plant viruses.

Additional keywords: OpenBUGS, post-test probability.

Grapevine leafroll disease (GLD) is one of the most important viral diseases affecting grapevine (*Vitis vinifera*). Currently, there are five recognized *Grapevine leafroll-associated virus* (GLRaV) species, including one member of the genus *Closterovirus* (GLRaV-2), three members of the genus *Ampelovirus* (GLRaV-1, GLRaV-3, and GLRaV-4 and its strains, GLRaV-5, -6, -9, GLRaV-Pr, GLRaV-De, and GLRaV-Car) and one member of the genus *Velarivirus* (GLRaV-7) (Martelli et al. 2012). *Grapevine leafroll-associated virus 3* (GLRaV-3) has the most prevalent worldwide distribution of the GLD complex (Maree et al. 2013). It has a positive-sense, single-stranded RNA genome of 18,498 nucleotides in length that comprises 12 open reading frames (Martelli et al. 2012). GLRaV-3 has been introduced to most grape growing regions via the exchange and propagation of infected plant material. Afterward, local spread occurs due to vegetative propagation and semipersistent transmission by several insect species belonging to the family *Pseudococcidae* (Homoptera order), such as *Pseudococcus maritimus* and *Planococcus citri* (Maliogka et al. 2015). GLD symptoms show high variability in some grape varieties, such as some rootstocks and certain white

V. vinifera cultivars, where infection remains completely symptomless (Maree et al. 2013). Therefore, the use of visual inspection as a diagnostic method has important drawbacks in the control of the disease, as asymptomatic infected plants can only be detected using serological (e.g., enzyme-linked immunosorbent assay [ELISA]) or nucleic acid-based methods (e.g., polymerase chain reaction [PCR]) (Walsh and Pietersen 2013). For instance, biological indexing using indicator plants detects GLD on Cabernet Franc or Pinot Noir vines, while molecular methods identify GLRaV-3 (López-Fabuel et al. 2013; Maliogka et al. 2015; Maree et al. 2013). However, the diagnostic parameters of these detection methods have not been estimated. Indeed, the scarcity of such estimators is a common problem in plant pathology (Olmos et al. 2008; Turechek et al. 2013; Vidal et al. 2012b). Therefore, the main purpose of this study was to estimate and compare the diagnostic parameters of three methods, including one serological test (double-antibody sandwich [DAS]-ELISA using polyclonal antibodies) and two nucleic acid-based techniques (spot and conventional real-time reverse transcription [RT]-PCR). Analyses were performed using data collected from three populations tested during the dormant period. The dormant period is the recommended time for testing for GLRaV-3 because plant material for propagation is traded during this period of the year and because of the high accumulation of virions in the phloem of lignified canes (Osman et al. 2012). In fact, ELISA- and PCR-based methods were able to detect more infected plants during the dormant period versus the growing vegetative period (Bertolini et al. 2010).

Corresponding author: E. Vidal; E-mail address: evidal@damastats.com

*The e-Xtra logo stands for “electronic extra” and indicates that one supplementary appendix is published online.

<http://dx.doi.org/10.1094/PHYTO-10-15-0246-R>
© 2016 The American Phytopathological Society

The most common diagnostic parameters used to evaluate detection tests are the sensitivity, specificity, accuracy, and positive and negative likelihood ratios (Table 1). Sensitivity is the capacity of a method to detect an infected plant within a population of infected plants, that is, the proportion of true positives. Specificity is the capacity of a method to produce as negative results a noninfected plant in the population of noninfected plants. It is the proportion of true negatives detected by the techniques out of the noninfected plants. Accuracy is the capacity of a method to correctly diagnose infected and noninfected plants, estimated by the proportion of true results (both true positives and true negatives) among the total number of plants analyzed. Finally, likelihood ratios have been reported to be the best measure of the usefulness of a diagnostic test. Positive likelihood ratio (LR+) expresses how much more probable it is to find a positive result in an infected plant than in a healthy one, while negative likelihood ratio (LR-) evaluates how much more probable it is to obtain a negative result in a healthy plant than in an infected one (Altman and Bland 1994; Deeks and Altman 2004; McNeil and Adelstein 1976; Olmos et al. 2008). Bayesian logic uses the positive and negative likelihood ratios to convert a pre-test probability or prevalence of infection into a post-test probability, according to Bayes' theorem (Bianchi and Alexander 2006; Deeks and Altman 2004) and has been applied to plant virology (Massart et al. 2008, 2009; Olmos et al. 2008).

The diagnostic parameters of a detection method can be calculated when the real sanitary status of the population of study is known. However, without a real knowledge of the sanitary status of the population, the suitability of a detection method is usually estimated by comparison with a gold standard technique (Enøe et al. 2000). Thus, the absence of a perfect gold standard method may lead to a misclassification of the infection status (van Smeden et al. 2014). To overcome this problem, statistical latent class models (LCMs) have been employed (Conraths and Schares 2006; Enøe et al. 2000; Reitsma et al. 2009; van Smeden et al. 2014). LCMs combine the results of multiple detection tests to estimate the prevalence, the sensitivity and the specificity of a method when a gold standard assay is not available (Reitsma et al. 2009; van Smeden et al. 2014). Hui and Walter (1980) provided formulae for estimating the sensitivity and specificity of two different methods by using two populations with different prevalences and maximum likelihood functions. In those cases in which there are no exact solutions to Hui-Walter formulae, iterative/numerical methods such as Newton-Raphson (Tanner 1996) or expectation-maximization algorithms (Dempster et al. 1977) are applied. This is also the case

when the model is expanded to allow estimation of sensitivity and specificity for more than two tests or populations (Enøe et al. 2000). It is important to note that, due to its simplicity and numerical stability, the expectation-maximization algorithm is well suited for solving LCMs (Becker et al. 1997). A Bayesian approach can be coupled with the previous methodology (Joseph et al. 1995, 2001) by combining maximum likelihood functions with beta distributions obtained from prior information using the Gibbs sampling algorithm, an iterative process based on Markov chain Monte Carlo methods (Gelfand and Smith 1990).

LCMs have been used as a statistical tool to evaluate plant virus detection tests in the absence of a gold standard (Turechek et al. 2013; Vidal et al. 2012a). In this study, several diagnostic parameters, including the sensitivity, specificity and likelihood ratios were estimated by fitting LCMs using a Bayesian approach to select the most suitable and reliable GLRaV-3 detection method for different infection prevalences. The use of Bayesian LCMs opens new possibilities in plant pathology diagnostic studies.

MATERIALS AND METHODS

Plant material. A total of 281 grapevine plants were sampled in winter (dormant period), which were collected from three different populations. Population 1 (P1) included 95 table grapevine plants from Designation of Origin (D.O.) "Vinalopó bagged table grape" (Spain). Population 2 (P2) included 155 wine grapevine plants from "D.O. Alicante" (Spain). Population 3 (P3) included 31 wine grapevine plants from "D.O. Rioja" (Spain).

GLRaV-3 monitoring. GLRaV-3 detection was performed by DAS-ELISA, using polyclonal antibodies, spot real-time RT-PCR, and conventional real-time RT-PCR, according to Bertolini et al. (2010). Briefly, at least five lignified canes were collected per grapevine plant. Extracts were prepared by grinding cambial scraping cuttings at a ratio of approximately 1/20 (wt/vol) in phosphate buffered saline, pH 7.2, supplemented with 0.2% (wt/vol) DIECA and 2% (wt/vol) polyvinyl-pyrrolidone (PVP-10). Grinding took place in individual plastic bags with a heavy net (Plant Print Diagnostics, Valencia, Spain) to avoid contamination among samples. The same crude extracts were used for DAS-ELISA, the spotting procedure, and total RNA purification.

DAS-ELISA using polyclonal antibodies against the viral coat protein was performed according to the manufacturer's instructions (AgriTest, Valenzano, Italy). The DAS-ELISA results were considered positive when the absorbance of a sample was equal to or greater than two times the absorbance of a healthy control (extract of GLRaV-3-free grapevine plant).

The spot procedure was performed by loading 5 µl of plant crude extract onto a 0.45-µm positively charged nylon membrane (Olmos et al. 1996), and viral targets were extracted with glycine buffer (0.1 M glycine, 0.05 M NaCl, 1 mM EDTA, pH 8.0) (Osman and Rowhani 2006).

Total RNA for conventional real-time RT-PCR, was extracted from 200 µl of crude extract using an Ultraclean Plant RNA isolation kit (Mobio, Carlsbad, CA) following the manufacturer's instructions.

The GLRaV3-56f and GLRaV3-285r primers and the TaqMan probe GLRaV3-181p, based on the Hsp70h gene (Osman and

TABLE 1. Contingency table for assessment of parameters of a diagnostic method^a

Test result	Disease status		Total
	Infected	Healthy	
+	a	b	a + b
-	c	d	c + d
Total	a + c	b + d	

^a Diagnostic parameters: sensitivity (Se) = $a/(a + c)$, specificity (Sp) = $d/(b + d)$, accuracy (A) = $[(a + d)/(a + b + c + d)]$, positive likelihood ratio (LR+) = $[Se/(1 - Sp)] = [a(b + d)]/[b(a + c)]$ and negative likelihood ratio (LR-) = $[1 - Se]/Sp = [c(b + d)]/[d(a + c)]$, being a = true positives, b = false positives, c = false negatives, and d = true negatives.

TABLE 2. Informative priors of sensitivity and specificity of double-antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA), spot real-time reverse transcription-polymerase chain reaction (RT-PCR), and conventional real-time RT-PCR for the analysis of the initial Bayesian models

Diagnostic test	Sensitivity	Mode ^a	Beta ^b	Specificity	Mode ^a	Beta ^b
DAS-ELISA	>0.80	0.86	(108.41, 18.49)	>0.90	0.99	(34.17, 1.34)
Spot real-time RT-PCR	>0.90	0.93	(262.96, 20.17)	>0.90	0.98	(42.11, 1.84)
Conventional real-time RT-PCR	>0.90	0.97	(53.58, 2.63)	>0.80	0.89	(51.41, 7.23)

^a Mode of sensitivity and specificity at 95% certainty.

^b Beta prior distribution obtained with the R platform package *prevalence*.

Rowhani 2006), were used for the spotting procedure and conventional real-time RT-PCR according to Bertolini et al. (2010). Real-time RT-PCR was performed in a StepOne Plus thermocycler (Applied Biosystems). The final volume of each reaction mixture was 12 µl, containing 1× AgPath-ID One-step RT-PCR master mix (Ambion),

TABLE 3. Results of the analysis of 281 samples by double-antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA), spot real-time reverse transcription-polymerase chain reaction (RT-PCR), and conventional real-time RT-PCR from three populations and their estimated prevalence^a

ELISA ^b	Spot ^c	qPCR ^d	P1 ^e	P2 ^f	P3 ^g	Total
(–)	(–)	(–)	5	53	29	87
(–)	(+)	(–)	0	4	0	4
(–)	(–)	(+)	22	57	1	80
(–)	(+)	(+)	7	6	0	13
(+)	(–)	(–)	0	0	0	0
(+)	(+)	(–)	0	0	0	0
(+)	(–)	(+)	2	2	0	4
(+)	(+)	(+)	59	33	1	93
Total			95	155	31	281
Estimated prevalence ^h			71.58%	26.45%	3.23%	39.15%

^a Obtained Fleiss' kappa index (95% confidence interval): 0.52 (0.45:0.60).

^b ELISA: DAS-ELISA.

^c Spot: spot real-time RT-PCR.

^d qPCR: conventional real-time RT-PCR.

^e P1: population 1, 95 table grapevine samples from Designation of Origin (D.O.) "Vinalopó bagged table grape" (Spain).

^f P2: population 2, 155 wine grapevine samples from "D.O. Alicante" (Spain).

^g P3: population 3, 31 wine grapevine samples from "D.O. Rioja" (Spain).

^h Number of samples positive by at least two techniques divided per total number of analyzed samples.

1× RT-PCR enzyme mix (Ambion), 0.5 µM each primer, 150 nM TaqMan probe, and 3 µl of template. The real-time RT-PCR protocol consisted of one step at 45°C for 10 min and 95°C for 10 min followed by 45 cycles of amplification (95°C for 15 s and 60°C for 1 min). Data acquisition and analysis were performed with StepOne Plus 2.0 software. The default threshold set by the machine was adjusted slightly above the noise at the narrowest point of the linear part of the amplification curve according to the StepOne Plus manufacturers. The spotting procedure and conventional real-time RT-PCR results were considered positive when the threshold cycle (Ct), which is the cycle at which a significant increase in fluorescence occurs that passes the threshold line, was reached. A Ct value below 45 was considered as a positive result according to Bertolini et al. (2010) in order to minimize false negative results, as multiplication of GLRaV-3 free plant material is a requirement under the European Directive 2005/43/EC.

Agreement between techniques. Global agreement between the three tested detection methods was evaluated using Fleiss' kappa index (Fleiss 1971), while agreement between pairs of detection techniques was evaluated using Cohen's kappa index (Cohen 1960). Both kappa indexes indicate the degree of agreement beyond what is expected by chance. The benchmarks suggested by Landis and Koch (1977) were used to categorize both kappa indexes, where <0.00 is poor agreement, 0.00 to 0.20 is slight agreement, 0.21 to 0.40 is fair agreement, 0.41 to 0.60 is moderate agreement, 0.61 to 0.80 is substantial agreement, and 0.81 to 1.00 is almost perfect agreement.

The confidence interval for Fleiss' kappa was estimated using a percentile bootstrap procedure (Falotico and Quatto 2010) to

TABLE 4. Comparison and agreement measures between the three possible pairwise combinations of the three detection techniques (double-antibody sandwich enzyme-linked immunosorbent assay [DAS-ELISA], spot real-time reverse transcription-polymerase chain reaction [RT-PCR], and/or conventional real-time RT-PCR) for the detection of *Grapevine leafroll-associated virus 3* in 281 grapevine samples

Comparison 1				Comparison 2				Comparison 3			
ELISA ^a				ELISA ^a				Spot ^b			

avoid bias. Fleiss' kappa and its confidence interval were estimated using the R platform package *raters* (<https://cran.r-project.org/>).

McNemar's χ^2 test (McNemar 1947) was used to detect the bias effect due to a high difference on the proportion of positive results between both tests, producing higher Cohen's kappa index values than expected (Feinstein and Cicchetti 1990). If McNemar's χ^2 test was significant, the bias-adjusted kappa (BAK) index (Byrt et al. 1993) was applied to correct the possible bias effect. Cohen's kappa and BAK indexes were estimated using the R platform package *epi-R* (<https://cran.r-project.org/>).

Estimation of diagnostic parameters using LCMs. The sensitivity and specificity were estimated using the *OpenBUGS* program (Lunn et al. 2009), which implements the Gibbs sampling algorithm. The Bayesian models that were applied assumed one of the following: (i) two tests were dependent, and the third test was independent of the first two (Branscum et al. 2005); and (ii) all three tests were independent, deleting the covariance term (Toft et al. 2005). Tests measuring different biological phenomena would probably be conditionally independent, while tests measuring similar biological processes could be classified as dependent (Gardner et al. 2000; Georgiadis et al. 2003). Thus, the models applied were classified

TABLE 6. Deviance information criterion (DIC) values for the Bayesian models using informative and noninformative priors

Priors		Model			
Informative	Noninformative ^a	A ^b	B ^c	C ^d	D ^e
Se: ELISA ^f , spot ^g , qPCR ^h Sp: ELISA, spot, qPCR Prevalence: P1, P2, P3	–	104.7 ⁱ (A1)	104.1 (B1)	103.7 (C1)	100.5 (D1)
Se: ELISA, spot Sp: ELISA, spot Prevalence: P1, P2, P3	Se: qPCR Sp: qPCR	97.9 (A2)	97.9 (B2)	97.7 (C2)	94.6 (D2)
Se: ELISA, qPCR Sp: ELISA, qPCR Prevalence: P1, P2, P3	Se: spot Sp: spot	105.8 (A3)	105.4 (B3)	104.7 (C3)	99.4 (D3)
Se: spot, qPCR Sp: spot, qPCR Prevalence: P1, P2, P3	Se: ELISA Sp: ELISA	105.4 (A4)	104.9 (B4)	104.5 (C4)	101.2 (D4)
Se: ELISA Sp: ELISA Prevalence: P1, P2, P3	Se: spot, qPCR Sp: spot, qPCR	99.0 (A5)	NC ^j (B5)	98.7 (C5)	100.9 (D5)
Se: spot Sp: spot Prevalence: P1, P2, P3	Se: ELISA, qPCR Sp: ELISA, qPCR	98.6 (A6)	98.8 (B6)	NC (C6)	95.5 (D6)
Se: qPCR Sp: qPCR Prevalence: P1, P2, P3	Se: ELISA, spot Sp: ELISA, spot	106.5 (A7)	106.3 (B7)	105.5 (C7)	NC (D7)
Se: ELISA, spot, qPCR Sp: ELISA, spot, qPCR Prevalence: P1, P2, P3	Se: ELISA, spot, qPCR Sp: ELISA, spot, qPCR Prevalence: P1, P2, P3	99.8 (A8)	–	–	–
Se: ELISA, spot, qPCR Sp: ELISA, spot, qPCR Prevalence: P1, P2, P3	Se: qPCR Sp: qPCR Prevalence: P1, P2, P3	108.9 (A9)	108.2 (B8)	107.9 (C8)	101.2 (D8)
Se: ELISA, qPCR Sp: ELISA, qPCR Prevalence: P1, P2, P3	Se: spot Sp: spot Prevalence: P1, P2, P3	102.1 (A10)	102.0 (B9)	101.1 (C9)	96.6 (D9)
Se: spot, qPCR Sp: spot, qPCR Prevalence: P1, P2, P3	Se: ELISA Sp: ELISA Prevalence: P1, P2, P3	110.1 (A11)	NC (B10)	109.0 (C10)	91.2 (D10)
Se: ELISA Sp: ELISA Prevalence: P1, P2, P3	Se: spot, qPCR Sp: spot, qPCR Prevalence: P1, P2, P3	109.7 (A12)	109.1 (B11)	NC (C11)	100.9 (D11)
Se: spot Sp: spot Prevalence: P1, P2, P3	Se: ELISA, qPCR Sp: ELISA, qPCR Prevalence: P1, P2, P3	103.2 (A13)	–	–	–
Se: qPCR Sp: qPCR Prevalence: P1, P2, P3	Se: ELISA, spot Sp: ELISA, spot Prevalence: P1, P2, P3	102.8 (A14)	–	–	–
–	Se: ELISA, spot, qPCR Sp: ELISA, spot, qPCR Prevalence: P1, P2, P3	110.7 (A15)	–	–	–
–	Se: ELISA, spot, qPCR Sp: ELISA, spot, qPCR Prevalence: P1, P2, P3	NC (A16)	–	–	–

^a Noninformative: uniform distribution (β (1, 1)).

^b Models A: Bayesian models which assume that three tests are conditional independent.

^c Models B: Bayesian models which assume that spot and conventional real-time reverse transcription-polymerase chain reaction (RT-PCR) are conditional dependent tests while enzyme-linked immunosorbent assay (ELISA) test is independent of other two tests.

^d Models C: Bayesian models which assume that ELISA and conventional real-time RT-PCR are conditional dependent tests while spot real-time RT-PCR test is independent of other two tests.

^e Models D: Bayesian model which assumes that ELISA and spot real-time RT-PCR are conditional dependent tests while conventional real-time RT-PCR test is independent of other two tests.

^f ELISA: double-antibody sandwich (DAS)-ELISA.

^g Spot: spot real-time RT-PCR.

^h qPCR: conventional real-time RT-PCR.

ⁱ DIC.

^j NC: not converge.

TABLE 7. Sensitivity, specificity, prevalences, positive and negative conditional correlations, and deviance information criterion (DIC) for the three diagnostic methods (double-antibody sandwich enzyme-linked immunosorbent assay [DAS-ELISA], spot real-time reverse transcription-polymerase chain reaction [RT-PCR], and conventional real-time RT-PCR), estimated by a Bayesian approach^a

Model	DAS-ELISA		Spot real-time RT-PCR		Conventional real-time RT-PCR	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
A2	0.870 (0.821:0.911) ^d	0.994 (0.974:1.000)	0.937 (0.908:0.959)	0.965 (0.931:0.987)	0.993 (0.963:1.000)	0.528 (0.453:0.603)
B2	0.868 (0.818:0.909)	0.993 (0.973:1.000)	0.936 (0.908:0.958)	0.969 (0.935:0.988)	0.992 (0.962:1.000)	0.529 (0.454:0.603)
C2	0.867 (0.816:0.909)	0.993 (0.972:1.000)	0.937 (0.909:0.959)	0.968 (0.933:0.991)	0.989 (0.947:1.000)	0.527 (0.451:0.602)
D10	0.762 (0.688:0.827)	0.993 (0.969:0.999)	0.787 (0.705:0.860)	0.934 (0.898:0.981)	0.986 (0.962:0.997)	0.723 (0.638:0.820)

(continued on next page)

^a Results of the best-fit implemented model (A2, B2, C2, and D10) are shown.

^b rhoD: conditional correlation for positive results.

^c rhoDc: conditional correlation for negative results.

^d Bayesian estimates (median and 95% probability intervals).

as (i) type A, which assumed that the three diagnostic tests were independent; (ii) type B, which assumed that the spot and conventional real-time RT-PCR tests were dependent, while DAS-ELISA was independent; (iii) type C, which assumed that the DAS-ELISA and conventional real-time RT-PCR tests were dependent, while the spot real-time RT-PCR was independent; and (iv) type D, which assumed that the DAS-ELISA and spot real-time RT-PCR tests were dependent, while conventional real-time RT-PCR was independent.

The R platform package *R2OpenBUGS* (<https://cran.r-project.org/>) was used to implement the *OpenBUGS* program. The *OpenBUGS* code for the applied models included two modifications: (i) the analysis was performed simultaneously for three populations and three methods and (ii) LRs and their combinations ($LR_1 = LR_1 \times LR_2 \times LR_3$) were calculated at each iteration based on the sensitivity and specificity, according to Joseph and Gyorkos (1996). Modifications of the *OpenBUGS* code are shown in Supplementary Appendix S1.

The conditional correlations between tests for sensitivity (rhoD) and specificity (rhoDc) were estimated using conditional dependent models. Convergence was determined using the Gelman-Rubin method (Brooks and Gelman 1998), which was implemented in *OpenBUGS* software using three chains per model. The first 10,000 iterations were discarded as the burn-in phase, and inferences were based on the subsequent 100,000 iterations.

Median values and percentiles were respectively used as point estimates and probability intervals of the different parameters (Enøe et al. 2000). Because the diagnostic parameters of the assessed GLRaV-3 detection techniques had not been previously published, estimated values for the detection of *Plum pox virus* on stone fruit trees (Capote et al. 2009) by DAS-ELISA (5B-IVIA monoclonal antibody), spot real-time RT-PCR, and conventional real-time RT-PCR during the dormant period were used to model the prior beta distribution (beta (a, b)) using the R platform package *prevalence* (<https://cran.r-project.org/>). The uniformity of diagnostic parameter values across populations was checked by analyzing and comparing them in each of the three populations using a conditional independent Bayesian model (Georgiadis et al. 2003). Informative priors for the analysis of the initial models are shown in Table 2. Similarly, the prevalence of the GLRaV-3 virus in each population was modeled, using the same package, considering as a positive result those which tested positive by at least two of the three methods. Thus, prevalence for GLRaV-3 in each population was represented by the following: $Prevalence_1 = \text{beta}(999.99, 389.5)$, mean 0.72; $Prevalence_2 = \text{beta}(31.25, 87.11)$, mean 0.26; and $Prevalence_3 = \text{beta}(10.75, 291.27)$, mean 0.04. The influence of the prior distributions was checked by comparing the initial models to new models with noninformative priors that had uniform distributions (beta (1, 1)) for sensitivity, specificity and prevalence values. This comparison is a key step to obtain distributions, which are mainly driven by the data, and not influenced by priors.

A total of 49 models were evaluated. The deviance information criterion (DIC) was applied to determine the fit of each model. The

models with better goodness-of-fit were those with the lowest DIC values (Spiegelhalter et al. 2002). The median value of the posterior distribution was used to estimate LRs via a Bayesian approach (Fosgate et al. 2006).

LRs were interpreted according to Jaeschke et al. (2002). $LR+ > 10$ and $LR- < 0.1$ will result in large conclusive changes from pre-test to post-test probabilities. $LR+$ between 5 and 10 or $LR-$ from 0.1 to 0.2, indicate moderate shifts in pre-test to post-test probabilities. $LR+$ values from 2 to 5 or $LR-$ between 0.5 and 0.2 indicate small (but sometimes important) changes from pre-test to post-test probabilities. $LR+$ between 1 and 2 or $LR-$ from 0.5 to 1 indicate small (and rarely important) shifts from pre-test to post-test probabilities. Finally $LRs = 1$ indicate no change from pre-test to post-test probabilities. Bayes' theorem was used to estimate the post-test probability of disease for an individual plant based on detection test results (Deeks and Altman 2004) using the following formulae: pre-test probability = prevalence, pre-test odds = prevalence/(1 - prevalence), post-test odds = pre-test odds $\times$ LR, and post-test probability = post-test odds/(1 + post-test odds).

RESULTS

Agreement between techniques. The agreement among DAS-ELISA, spot real-time RT-PCR, and conventional real-time RT-PCR by Fleiss' kappa index was 0.52 (Table 3), showing moderate agreement (Landis and Koch 1977).

Table 4 shows the comparison by pairs between DAS-ELISA, spot real-time RT-PCR, and conventional real-time RT-PCR for the detection of GLRaV-3 and the agreement based on Cohen's kappa index (Landis and Koch 1977). It is important to note that a bias effect was detected by McNemar's χ^2 test in all comparisons. BAK indexes were applied to correct the bias. The total percentage of coincidental results between DAS-ELISA and spot real-time RT-PCR was 92.53% with total Cohen's kappa and BAK indexes of 0.84. These results showed a high agreement. The total percentage of coincidental results between DAS-ELISA and conventional real-time RT-PCR was 66.90% with a total Cohen's kappa index of 0.40 and a BAK index of 0.34. These results indicated moderate and fair agreements, respectively. The total percentage of coincidental results between spot real-time RT-PCR and conventional real-time RT-PCR was 68.68% with a total Cohen's kappa index of 0.42 and a BAK index of 0.35. These results showed moderate and fair agreements, respectively.

Calculation of diagnostic parameters using LCMs. Sensitivity and specificity of the three diagnostic methods were estimated by 49 variants of a Bayesian model that assumed constant values for the same technique across the three populations of study. This requirement was checked by estimating these parameters using a conditional independent Bayesian model (Table 5). The estimated median values for the sensitivity were similar across all populations for DAS-ELISA, spot real-time RT-PCR, and conventional real-time RT-PCR. Median values for specificity of DAS-ELISA and

TABLE 7. (continued from preceding page)

Prevalence			rhoD ^b	rhoDc ^c	DIC
Population 1	Population 2	Population 3			
0.720 (0.697:0.742)	0.258 (0.207:0.313)	0.034 (0.018:0.057)	–	–	97.9
0.720 (0.697:0.742)	0.260 (0.209:0.315)	0.034 (0.018:0.057)	0.138 (–0.011:0.541)	–0.067 (–0.176:0.093)	97.9
0.720 (0.696:0.742)	0.260 (0.209:0.316)	0.034 (0.018:0.058)	0.139 (–0.020:0.478)	0.017 (–0.081:0.130)	97.7
0.912 (0.813:0.969)	0.354 (0.250:0.492)	0.052 (0.006:0.165)	0.727 (0.553:0.846)	0.156 (–0.008:0.533)	91.2

spot real-time RT-PCR were also similar for all populations. However, the specificity of conventional real-time RT-PCR in population 3 (0.910) was different from those obtained for population 1 (0.671) and population 2 (0.623). As conventional real-time RT-PCR had almost perfect sensitivity in all populations, the specificity value obtained in our study can be a good estimate of the weighted mean of this parameter in the three populations (Johnson et al. 2009; Toft et al. 2005).

A total of 49 different models were fit to the data (Table 6). The estimated DIC values for each model ranged from (i) 97.9 to 110.7 in type A models, (ii) 97.9 to 109.1 in type B models, (iii) 97.7 to 107.9 in type C models, and (iv) 91.2 to 101.2 in type D models. This fact indicates that the best-fit models were those that assumed DAS-ELISA and spot real-time RT-PCR to be dependent tests, while conventional real-time RT-PCR being independent of the other two (type D models). This finding was confirmed as a 95% probability interval of conditional correlation for positive results (rhoD) for model D10, with the lowest DIC value (91.2), gave a value of 0.727 (0.553:0.846), which did not include zero (Table 7). Therefore, the hypothesis suggesting the conditional dependence of DAS-ELISA and spot real-time RT-PCR could be accepted.

The sensitivities and specificities for the best fitting models of each group are shown in Table 7. Conventional real-time RT-PCR had the highest sensitivity, while DAS-ELISA was the most specific for the detection of GLRaV-3 in all models. Spot real-time RT-PCR always presented intermediate sensitivity and specificity values compared with the other two techniques.

The estimated likelihood ratios of the detection methods for the best fitting models are shown in Table 8. The positive likelihood ratios (LR+) of DAS-ELISA and spot real-time RT-PCR were >10 in all models. The negative likelihood ratios (LR–) of the conventional real-time RT-PCR were <0.1 in all models. The LRs obtained combining two or three methods using model D10 are shown in Table 9.

The post-test probabilities of GLRaV-3 infection of an individual plant after a diagnosis by one test or the combination of two or three tests, based on the prevalence of the disease, are shown in Table 10. Only the results obtained using the best-fit model (D10) are shown. Model D10 showed that for populations with a very low prevalence (0.1%) of infection the estimated post-test probability of GLRaV-3 infection after a positive detection result by DAS-ELISA (9.31%) was higher than after a positive diagnostic result by spot or conventional real-time RT-PCR (1.52 and 0.35%, respectively). This difference increases with infection prevalence, peaking at a prevalence of 10%, and decreasing among tests from this point. The estimated post-test probability of GLRaV-3 infection after a negative result was practically equal for all techniques. However, for prevalences higher than 0.5%, the post-test probability of GLRaV-3 infection after a negative result by conventional real-time RT-PCR was lower than after a negative result by DAS-ELISA or spot real-time RT-PCR (0.01% versus 0.12 and 0.11%, respectively). This difference increased with higher prevalences, post-test probabilities of infection after positive results were similar for all techniques at these prevalences.

DISCUSSION

Accurate detection of GLRaV-3 during the dormant period is required because plant material for multiplication purposes is traded

in this period. Consequently, it is necessary to improve detection methods and to assess the risks associated with each test in terms of false positive and false negative rates. Determining the best diagnostic technique to apply under different levels of disease prevalence requires a proper evaluation of test performance. The issues that are relevant for judging detection test performance are not widely disseminated in plant pathology. It is important to note that, in general, there are few diagnostic scenarios where a gold standard test exists (López et al. 2009). Thus, the lack of awareness on these issues may lead to misinterpretation of results.

An essential step in the evaluation of a detection test is to obtain valid estimates of its diagnostic parameters (Lewis and Torgerson 2012). In this study, three detection methods routinely used for diagnosis of GLRaV-3 were compared, and according to Fleiss' kappa index there was only moderate agreement among them. However, when comparing the methods in pairs, Cohen's kappa and BAK indexes showed a high agreement between DAS-ELISA and spot real-time RT-PCR, whereas conventional real-time RT-PCR only showed fair or moderate agreement with the other tests. The bias effect determined by McNemar's χ^2 test did not affect Cohen's kappa when DAS-ELISA was compared with spot real-time RT-PCR because the BAK index value was similar to Cohen's kappa index. The opposite was found when comparing conventional real-time RT-PCR to the other two techniques. This finding was expected because bias effect is only reflected in Cohen's kappa index when low agreement exists between methods (Byrt et al. 1993). Ninety-three and 84 more plants (out of 281) tested positive by conventional real-time RT-PCR and spot real-time RT-PCR, respectively, than by DAS-ELISA. However, in our conditions, the high sensitivity of conventional real-time RT-PCR could lead to a misclassification of the plant's sanitary status, due to a high rate of false positives.

In order to decide which is the best detection method or combination of methods, diagnostic parameters for each test were estimated by LCMs. This approach provides a solution in cases where no detection method is considered a gold standard (Enøe et al. 2000; Turechek et al. 2013; Vidal et al. 2012a). Combining a Bayesian approach and the use of LCMs allows parameters to be estimated using prior information (Lewis and Torgerson 2012).

In the current study, different Bayesian models were applied to estimate the sensitivity, specificity and likelihood ratios of DAS-ELISA, spot real-time RT-PCR, and conventional real-time RT-PCR for GLRaV-3 detection and were used to evaluate the dependence or independence of the different assessed techniques. A new code was used to perform this analysis using the *OpenBUGS* program, in order to apply the models to three different populations and to calculate the likelihood ratios at each iteration. In addition, models with noninformative priors distributions were included to assess the influence of prior distributions on parameter estimates (Georgiadis et al. 2003).

The best-fit models, which showed the lowest DIC values, were those that assumed DAS-ELISA and spot real-time RT-PCR techniques to be dependent and conventional real-time RT-PCR to be independent (type D models). This finding was confirmed in model D10, which showed a 95% probability interval of conditional correlation for sensitivity (rhoD) that did not include zero. It would be expected that the three detection methods had a degree of

TABLE 8. Likelihood ratios of double-antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA), spot real-time reverse transcription-polymerase chain reaction (RT-PCR), and conventional real-time RT-PCR estimated by a Bayesian approach^a

Model	DAS-ELISA		Spot real-time RT-PCR		Conventional real-time RT-PCR	
	LR+ ^b	LR- ^c	LR+	LR-	LR+	LR-
A2	133.1 (32.89:1,777) ^d	0.131 (0.090:0.181)	27.06 (13.62:70.77)	0.066 (0.043:0.095)	2.100 (1.810:2.490)	0.013 (0.000:0.070)
B2	131.1 (33.26:1,741)	0.134 (0.092:0.183)	30.20 (14.46:81.03)	0.066 (0.043:0.095)	2.100 (1.810:2.500)	0.014 (0.001:0.072)
C2	127.7 (31.30:1,698)	0.134 (0.092:0.185)	29.51 (14.00:105.5)	0.065 (0.013:0.094)	2.080 (1.790:2.480)	0.022 (0.001:0.102)
D10	102.6 (24.34:1,380)	0.241 (0.174:0.316)	15.39 (7.580:40.91)	0.225 (0.147:0.313)	3.550 (2.710:5.470)	0.019 (0.004:0.054)

^a Results of the best-fit implemented models (A2, B2, C2, and D10) are shown.

^b LR+: positive likelihood ratio.

^c LR-: negative likelihood ratio.

^d Bayesian estimates (median and 95% probability intervals).

TABLE 9. Combined likelihood ratios of two and three detection methods (double-antibody sandwich enzyme-linked immunosorbent assay [DAS-ELISA], spot real-time reverse transcription-polymerase chain reaction [RT-PCR], and conventional real-time RT-PCR) estimated by the Bayesian model D10

DAS-ELISA	Diagnostic method		Parameter
	Spot real-time RT-PCR	Conventional real-time RT-PCR	
(+) ^b	(+)	* ^d	1,675 (279.2:27,740) ^e
(+)	(-) ^c	*	22.86 (5.220:310.7)
(-)	(+)	*	3.670 (1.740:10.09)
(-)	(-)	*	0.054 (0.027:0.096)
(+)	*	(+)	372 (84.71:5056)
(+)	*	(-)	1.920 (0.226:30.48)
(-)	*	(+)	0.852 (0.518:1.600)
(-)	*	(-)	0.004 (0.001:0.013)
*	(+)	(+)	55.82 (25.85:158.1)
*	(+)	(-)	0.289 (0.050:1.240)
*	(-)	(+)	0.789 (0.441:1.560)
*	(-)	(-)	0.004 (0.001:0.013)
(+)	(+)	(+)	6,071 (966.2:101,300)
(+)	(+)	(-)	30.72 (2.820:613.8)
(+)	(-)	(+)	83.48 (17.50:1165)
(+)	(-)	(-)	0.427 (0.050:6.880)
(-)	(+)	(+)	13.38 (5.630:40.84)
(-)	(+)	(-)	0.069 (0.012:0.299)
(-)	(-)	(+)	0.192 (0.082:0.471)
(-)	(-)	(-)	0.001 (0.000:0.003)

^a Combined likelihood ratio: combined likelihood ratios obtained by multiplying positive likelihood ratio (for a positive result) or negative likelihood ratio (for a negative result) of each test.

^b (+): positive result.

^c (-): negative result.

^d *: not tested.

^e Bayesian estimates (median and 95% probability intervals).

conditional dependence because they can detect the presence of GLRaV-3 in plants with different sanitary statuses. Thus, the dependence between DAS-ELISA and spot real-time RT-PCR is justified. However, the high sensitivity of conventional real-time RT-PCR could explain its independence from the other two tests, considering that when a method has a sensitivity (or specificity) equal to 1, it is conditionally independent (Branscum et al. 2005).

The sensitivity and specificity of DAS-ELISA, spot real-time RT-PCR, and conventional real-time RT-PCR were estimated using the best Bayesian LCMs. In all models, conventional real-time RT-PCR turned out to be the most sensitive method for the detection of GLRaV-3 and DAS-ELISA was the most specific technique. Spot real-time RT-PCR presented intermediate sensitivity and specificity values compared with the other two techniques.

Because sensitivity does not consider false positive results and specificity does not consider false negative results, likelihood ratios for the three techniques were calculated as alternative statistics for summarizing diagnostic accuracy. Thus, according to Jaeschke et al. (2002), the LR+ values calculated for model D10 showed that the use of DAS-ELISA and spot real-time RT-PCR

gave high confidence in providing positive results. On the other hand, LR- values calculated for this model indicated a high degree of confidence in negative results for conventional real-time RT-PCR.

In addition, LRs were used to estimate post-test probability at different prevalences. At low prevalences, LR+ values obtained for DAS-ELISA and spot real-time RT-PCR provided large conclusive changes from pre-test to post-test probability, while conventional real-time RT-PCR provided lower changes. Conversely, at high prevalences, LR- values obtained for conventional real-time RT-PCR provided conclusive changes from pre-test to post-test probability, while DAS-ELISA and spot real-time RT-PCR provided lower changes. Therefore, at low prevalences, methods with the highest LR+ give more confidence in the positive results than the other methods, while all methods give similar confidence for negative results. However, at high prevalences, methods with the lowest LR- values provide more confidence in negative results, while the confidence in positive results is similar for all methods. For example (Table 10), in a population with a 1% prevalence of infection, a positive result by DAS-ELISA means that a plant has 50.89% probability of being infected, whereas conventional real-time RT-PCR suggests this probability is 3.46%. A negative result by DAS-ELISA has 0.24% probability of being infected while conventional real-time RT-PCR gives 0.02%. However, in populations with higher prevalences of infection (e.g., 60%), a negative result by DAS-ELISA means that the probability of infection in a plant is 26.55%; as such, infected plants might escape the detection method. At this prevalence, a negative result by conventional real-time RT-PCR represents 2.77% probability of infection, a positive result by DAS-ELISA gives 99.35% probability of infection, and a positive result by conventional real-time RT-PCR 84.19% probability of infection. In general, DAS-ELISA could be the selected method for low prevalences and conventional real-time RT-PCR for those populations with high prevalence of infection. Spot real-time RT-PCR has diagnostic parameters that lie between those of DAS-ELISA and conventional real-time RT-PCR.

Combining LRs of methods increases the confidence of the results. For example, the confidence of a negative diagnosis for a plant that tested negative by DAS-ELISA and conventional real-time RT-PCR is almost 100% at any prevalence (Table 10). Consequently, it is recommended that DAS-ELISA and conventional real-time RT-PCR be used to evaluate critical materials, such as mother plants and plant material for propagation and/or exportation. In the case of discordant results between DAS-ELISA and conventional real-time RT-PCR, combining LRs facilitates the diagnostic of the sanitary status of a plant. Thus, the application of a third method, spot real-time RT-PCR, significantly improves the confidence of the result. For example, if DAS-ELISA tests positive and conventional real-time RT-PCR tests negative, the post-test probability of infection at 5% of prevalence is 9.18%. When the third method test is included, a negative result decreases the post-test probability of infection from 9.18 to 2.20%, giving more confidence in the negative result. If spot real-time RT-PCR gives a positive result, the post-test probability of infection

TABLE 10. post-test probability of *Grapevine leafroll-associated virus 3* infection based on double-antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA), spot real-time reverse transcription-polymerase chain reaction (RT-PCR), and conventional real-time RT-PCR results estimated by the Bayesian model D10 at different prevalence conditions

Pre-test probability	0.01%	0.10%	0.50%	1.00%	5.00%	10.00%	20.00%	30.00%	60.00%	90.00%
post-test probability of infection (%)										
Diagnostic method										
ELISA ^a (+) ^b	0.10	9.31	34.02	50.89	84.38	91.94	96.25	97.78	99.35	99.89
ELISA (-) ^c	0.00	0.02	0.12	0.24	1.25	2.61	5.68	9.36	26.55	68.44
Spot ^d (+)	0.02	1.52	7.18	13.45	44.75	63.10	79.37	86.83	95.85	99.28
Spot (-)	0.00	0.02	0.11	0.23	1.17	2.44	5.33	8.79	25.23	66.94
qPCR ^e (+)	0.00	0.35	1.75	3.46	15.74	28.29	47.02	60.34	84.19	96.97
qPCR (-)	0.00	0.00	0.01	0.02	0.10	0.21	0.47	0.81	2.77	14.60
Diagnostic methods										
ELISA (+); spot (+)	1.65	62.64	89.38	94.42	98.88	99.47	99.76	99.86	99.96	99.99
ELISA (+); spot (-)	0.02	2.24	10.30	18.76	54.61	71.75	85.11	90.74	97.17	99.52
ELISA (-); spot (+)	0.00	0.37	1.81	3.57	16.19	28.97	47.85	61.13	84.63	97.06
ELISA (-); spot (-)	0.00	0.01	0.03	0.05	0.28	0.60	1.33	2.26	7.49	32.71
ELISA (+); qPCR (+)	0.37	27.13	65.15	78.98	95.14	97.64	98.94	99.38	99.82	99.97
ELISA (+); qPCR (-)	0.00	0.19	0.96	1.90	9.18	17.58	32.43	45.14	74.23	94.53
ELISA (-); qPCR (+)	0.00	0.09	0.43	0.85	4.29	8.65	17.56	26.75	56.10	88.46
ELISA (-); qPCR (-)	0.00	0.00	0.00	0.00	0.02	0.04	0.10	0.17	0.60	3.47
spot (+); qPCR (+)	0.06	5.29	21.91	36.05	74.61	86.12	93.31	95.99	98.82	99.80
spot (+); qPCR (-)	0.00	0.03	0.15	0.29	1.50	3.11	6.74	11.02	30.24	72.23
spot (-); qPCR (+)	0.00	0.08	0.39	0.79	3.99	8.06	16.48	25.27	54.20	87.66
spot (-); qPCR (-)	0.00	0.00	0.00	0.00	0.02	0.04	0.10	0.17	0.60	3.47
Diagnostic methods										
ELISA (+); spot (+); qPCR (+)	5.72	85.87	96.83	98.40	99.69	99.85	99.93	99.96	99.99	100.00
ELISA (+); spot (+); qPCR (-)	0.03	2.98	13.37	23.68	61.79	77.34	88.48	92.94	97.88	99.64
ELISA (+); spot (-); qPCR (+)	0.08	7.71	29.55	45.75	81.46	90.27	95.43	97.28	99.21	99.87
ELISA (+); spot (-); qPCR (-)	0.00	0.04	0.21	0.43	2.20	4.53	9.65	15.47	39.04	79.35
ELISA (-); spot (+); qPCR (+)	0.01	1.32	6.30	11.91	41.32	59.79	76.99	85.15	95.25	99.18
ELISA (-); spot (+); qPCR (-)	0.00	0.01	0.03	0.07	0.36	0.76	1.70	2.87	9.38	38.31
ELISA (-); spot (-); qPCR (+)	0.00	0.02	0.10	0.19	1.00	2.09	4.58	7.60	22.36	63.34
ELISA (-); spot (-); qPCR (-)	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.04	0.15	0.89

^a ELISA: DAS-ELISA.

^b (+): positive result.

^c (-): negative result.

^d Spot: spot real-time RT-PCR.

^e qPCR: conventional real-time RT-PCR.

increases from 9.18 to 61.7%, confirming the positive result obtained by DAS-ELISA (Table 10).

The LCMs statistical approach is well known in medical and veterinary disciplines but has only been recently implemented to evaluate detection methods in plant pathology (Turechek et al. 2013; Vidal et al. 2012a). This study reports the flexibility and power of using LCMs in combination with a Bayesian approach to obtain diagnostic parameters in the absence of a gold standard method. This methodology facilitates decision-making regarding detection method selection. The present study is the first to report sensitivities, specificities, and likelihood ratios for common GLRaV-3 diagnostic techniques and should lead to a better management and control of GLRaV-3. Furthermore, it opens new possibilities for the diagnosis of grapevine viruses.

ACKNOWLEDGMENTS

This work was supported by the INIA RTA2014-00061 and IVIA 5405 projects. We thank A. I. Mulet (MA in English Language and Literature, Timpanogos Intermediate School, Heber City, UT) for her revision of this manuscript.

LITERATURE CITED

- Altman, D. G., and Bland, J. M. 1994. Diagnostic tests 1: Sensitivity and specificity. *BMJ* 308:1552.
- Becker, M. P., Yang, I., and Lange, K. 1997. EM algorithms without missing data. *Stat. Methods Med. Res.* 6:38-54.
- Bertolini, E., García, J., Yuste, A., and Olmos, A. 2010. High prevalence of viruses in table grape from Spain detected by real-time RT-PCR. *Eur. J. Plant Pathol.* 128:283-287.
- Bianchi, M. T., and Alexander, B. M. 2006. Evidence based diagnosis: Does the language reflect the theory? *BMJ* 333:442-445.
- Branscum, A. J., Gardner, I. A., and Johnson, W. O. 2005. Estimation of diagnostic-test sensitivity and specificity through Bayesian modelling. *Prev. Vet. Med.* 68:145-163.
- Brooks, S. P., and Gelman, A. 1998. Alternative methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat.* 7:434-455.
- Byrt, T., Bishop, J., and Carlin, J. B. 1993. Bias, prevalence and kappa. *J. Clin. Epidemiol.* 46:423-429.
- Capote, N., Bertolini, E., Olmos, A., Vidal, E., Martínez, M. C., and Cambra, M. 2009. Direct simple preparation methods for the detection of *Plum pox virus* by real-time RT-PCR. *Int. Microbiol.* 12:1-6.
- Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20:37-46.
- Conraths, F. J., and Schares, G. 2006. Validation of molecular diagnostic techniques in the parasitological laboratory. *Vet. Parasitol.* 136:91-98.
- Deeks, J. J., and Altman, D. G. 2004. Diagnostic tests 4: Likelihood ratios. *BMJ* 329:168-169.
- Dempster, A., Laird, N., and Rubin, D. 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B* 39:1-38.
- Enøe, C., Georgiadis, M. P., and Johnson, W. O. 2000. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev. Vet. Med.* 45:61-81.
- Falotico, R., and Quatto, P. 2010. On avoiding paradoxes in assessing inter-rater agreement. *Ital. J. Appl. Stat.* 22:151-160.
- Feinstein, A. R., and Cicchetti, D. V. 1990. High agreement but low kappa: I. The problems with two paradoxes. *J. Clin. Epidemiol.* 43:543-549.
- Fleiss, J. L. 1971. Measuring nominal scale agreement among many raters. *Psychol. Bull.* 76:378-382.
- Fosgate, G. T., Adesiyun, A. A., Hird, D. W., and Hietala, S. K. 2006. Likelihood ratio estimation without a gold standard: A case study evaluating a brucellosis c-ELISA in cattle and water buffalo of Trinidad. *Prev. Vet. Med.* 75:189-205.
- Gardner, I. A., Stryhn, H., Lind, P., and Collins, M. T. 2000. Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. *Prev. Vet. Med.* 45:107-122.

- Gelfand, A., and Smith, A. F. M. 1990. Sampling-based approaches to calculating marginal densities. *J. Am. Stat. Assoc.* 85:398-409.
- Georgiadis, M. P., Johnson, W. O., Gardner, I. A., and Singh, R. 2003. Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. *J. R. Stat. Soc. Ser. C* 52:63-76.
- Hui, S. L., and Walter, S. D. 1980. Estimating the error rates of diagnostic tests. *Biometrics* 36:167-71.
- Jaeschke, R., Guyatt, G., and Lijmer, J. 2002. Diagnostic tests. Pages 121-140 in: *Users' Guides to the Medical Literature*. G. Guyatt and D. Rennie, eds. AMA Press, Chicago.
- Johnson, W. O., Gardner, I. A., Metoyer, C. N., and Brascum, A. J. 2009. On the interpretation of test sensitivity in the two-test two-population problem: Assumptions matter. *Prev. Vet. Med.* 91:116-121.
- Johnson, W. O., Gastwirth, J. L., and Pearson, L. M. 2001. Screening without a "Gold Standard": The Hui-Walter paradigm revisited. *Am. J. Epidemiol.* 153:921-924.
- Joseph, L., and Gyorkos, T. W. 1996. Inferences for likelihood ratios in the absence of a "gold standard". *Med. Decis. Making* 16:412-417.
- Joseph, L., Gyorkos, T. W., and Coupal, L. 1995. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *Am. J. Epidemiol.* 141:263-272.
- Landis, J. R., and Koch, G. G. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33:159-174.
- Lewis, F. I., and Torgerson, P. R. 2012. A tutorial in estimating the prevalence of disease in humans and animals in the absence of a gold standard diagnostic. *Emerg. Themes Epidemiol.* 9:9.
- López, M. M., Llop, P., Olmos, A., Marco-Noales, E., Cambra, M., and Bertolini, E. 2009. Are molecular tools solving the challenges posed by detection of plant pathogenic bacteria and viruses? *Curr. Issues Mol. Biol.* 11:13-46.
- López-Fabuel, I., Wetzels, T., Bertolini, E., Bassler, A., Vidal, E., Torres, L., Yuste, A., and Olmos, A. 2013. Real-time multiplex RT-PCR for the simultaneous detection of the five main grapevine viruses. *J. Virol. Methods* 188:21-24.
- Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. 2009. The BUGS project: Evolution, critique and future directions (with discussion). *Stat. Med.* 28:3049-3067.
- Maliogka, V. I., Martelli, G. P., Fuchs, M., and Katis, N. I. 2015. Control of viruses infecting grapevine. Pages 175-227 in: *Advances in Virus Research*, Vol. 91. G. Loebenstein and N. I. Katis, eds. Academic Press, Burlington.
- Maree, H. J., Almeida, R. P. P., Bester, R., Chooi, K. M., Cohen, D., Dolja, V. V., Fuchs, M., Golino, G. A., Jooste, A. E. C., Martelli, G. P., Naidu, R. A., Rowhani, A., Saldarelli, P., and Burger, J. T. 2013. Grapevine leafroll-associated virus 3. *Front. Microbiol.* 4:82.
- Martelli, G. P., Ghanem-Sabanadzovic, A. N., Agranovsky, A. A., Rwanhah, A. M., Dolja, V. V., Dovas, C. I., Fuchs, M., Gugerli, P., Hu, J. S., Jelkmann, W., Katis, N. I., Maliogka, V. I., Melzer, M. J., Menzel, W., Minafra, A., Rott, M. E., Rowhani, A., Sabanadzovic, S., and Saldarelli, P. 2012. Taxonomic revision of the family *Closteroviridae* with special reference to the Grapevine leafroll-associated members of the genus *Ampelovirus* and the putative species unassigned to the family. *J. Plant Pathol.* 94:7-10.
- Massart, S., Brostaux, Y., Barbarossa, L., Batlle, A., César, V., Dutrecq, O., Fonseca, F., Komorowska, B., Olmos, A., Steyer, S., Wetzels, T., Kummert, J., and Jijakli, M. H. 2009. Inter-laboratory evaluation of two RT-PCR based methods for detection of four fruit tree viruses. *Ann. Appl. Biol.* 154:133-141.
- Massart, S., Brostaux, Y., Barbarossa, L., César, V., Cieslinska, M., Dutrecq, O., Fonseca, F., Guillem, R., Laviña, A., Olmos, A., Steyer, S., Wetzels, T., Kummert, J., and Jijakli, M. H. 2008. Inter-laboratory evaluation of a duplex RT-PCR method using crude extracts for the simultaneous detection of *Prune dwarf virus* and *Prunus necrotic ringspot virus*. *Eur. J. Plant Pathol.* 122:539-547.
- McNeil, B. J., and Adelstein, S. J. 1976. Determining the value of diagnostic and screening tests. *J. Nucl. Med.* 17:439-448.
- McNemar, Q. 1947. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 12:153-157.
- Olmos, A., Bertolini, E., Capote, N., and Cambra, M. 2008. An evidence-based approach to *Plum pox virus* detection by DASi-ELISA and RT-PCR in dormant period. *Virology (Auckl.)* 1:1-8.
- Olmos, A., Dasi, M. A., Candresse, T., and Cambra, M. 1996. Print-capture PCR: A simple and highly sensitive method for the detection of *Plum pox virus* (PPV) in plant tissues. *Nucleic Acids Res.* 24:2192-2193.
- Osman, F., Golino, D., and Rowhani, A. 2012. The effect of seasonal changes on the titer of grapevine viruses in infected *Vitis vinifera* and their downstream detection using RT-PCR and qRT-PCR. Pages 218-219 in: *Proc. 17th Congr. Int. Council Study Virus Virus-Like Dis. Grapevine (ICVG)*, Davis, CA.
- Osman, F., and Rowhani, A. 2006. Application of a spotting sample preparation technique for the detection of pathogens in woody plants by RT-PCR and real-time PCR (TaqMan). *J. Virol. Methods* 133:130-136.
- Reitsma, J. B., Rutjes, A. W. S., Khan, K. S., Coomarasamy, A., and Bossuyt, P. M. 2009. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *J. Clin. Epidemiol.* 62:797-806.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. 2002. Bayesian measures of model complexity and fit (with discussion). *J. R. Stat. Soc. Ser. B* 64:583-639.
- Tanner, M. A. 1996. *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*, 3rd ed. Springer, New York.
- Toft, N., Jørgensen, E., and Søren, H. 2005. Diagnosing diagnostic tests: Evaluating the assumptions underlying the estimation of sensitivity and specificity in the absence of a gold standard. *Prev. Vet. Med.* 68:19-33.
- Turechek, W. W., Webster, C. G., Duan, J., Roberts, P. D., Kousik, C. S., and Adkins, S. 2013. The use of latent class analysis to estimate the sensitivities and specificities of diagnostic tests for *Squash vein yellowing virus* in cucurbit species when there is no gold standard. *Phytopathology* 103:1243-1251.
- van Smeden, M., Naaktgeboren, C. A., Reitsma, J. B., Moons, K. G. M., and de Groot, J. A. H. 2014. Latent class models in diagnostic studies when there is no reference standard—A systematic review. *Am. J. Epidemiol.* 179: 423-431.
- Vidal, E., Moreno, A., Bertolini, E., and Cambra, M. 2012a. Estimation of the accuracy of two diagnostic methods for the detection of *Plum pox virus* in nursery blocks by latent class models. *Plant Pathol.* 61:413-422.
- Vidal, E., Yokomi, R. K., Moreno, A., Bertolini, E., and Cambra, M. 2012b. Calculation of diagnostic parameters of advanced serological and molecular tissue-print methods for detection of *Citrus tristeza virus*: A model for other plant pathogens. *Phytopathology* 102:114-121.
- Walsh, H. A., and Pietersen, G. 2013. Rapid detection of *Grapevine leafroll-associated virus* type 3 using a reverse transcription loop-mediated amplification method. *J. Virol. Methods* 194:308-316.