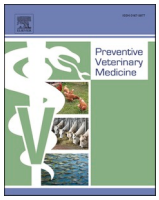
Preventive Veterinary Medicine 190 (2021) 105338

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed



Bayesian latent class model estimates of diagnostic accuracy for three test methods designed to detect spring viremia of carp virus

Sharon C. Clouthier [a,*], Carol McClure [b], Tamara Schroeder [a], Eric D. Anderson [c]

[a] Fisheries and Oceans Canada, Freshwater Institute, 501 University Crescent, Winnipeg, Manitoba, R3T 2N6, Canada
[b] AquaEpi Research, RR3, 300 Heartz Road, Charlottetown, Prince Edward Island, C1C 0H8, Canada
[c] Box 28, Group 30, RR2, Ste Anne, Manitoba, R5H 1R2, Canada

ARTICLE INFO

Keywords:
Spring viremia of carp virus
Diagnostic accuracy
Maximum likelihood and Bayesian latent class models
RT-qPCR
Virus isolation

ABSTRACT

Spring viremia of carp virus (SVCV) causes a systemic hemorrhagic disease that poses a significant risk to wild and cultured fish and is listed as notifiable by the World Organization for Animal Health. Validated molecular diagnostic tools for SVCV are required to accurately describe and analyze the ecology of the virus. Here, the diagnostic specificity (DSp) and sensitivity (DSe) (i.e. accuracy) of three SVCV diagnostic tests – 2 reverse transcription quantitative polymerase chain reaction (RT-qPCR) assays Q1G and Q2N and virus isolation by cell culture (VI) – were evaluated using 2-class latent class models run in maximum likelihood (ML) and Bayesian frameworks. Virus-free or experimentally-infected koi were sorted into three populations with low, moderate or high prevalence levels of SVCV (n = 269 fish in total). Koi kidney tissues were tested using Q2N and Q1G and for the VI assay, pools of kidney, spleen and gill tissues were used. All samples were blinded and analyzed in one laboratory. The ML and Bayesian approaches successfully estimated the diagnostic accuracy of the 3 tests with the exception of 1 ML model. The estimates were consistent across the two frameworks. The DSe estimates were higher for Q1G (>98 %) and Q2N (>96 %) compared to VI (>60 %). The DSp of all three tests varied by 12–15 % (79–91 % for Q1G, 79–94 % for Q2N and 81–97 % for VI) across same-fish samples revealing the potential range in test performance for one sample. The 3 fish populations had distinct SVCV prevalence levels estimated at 0–3 % (low), 70–73 % (moderate) and 95–96 % (high). The Bayesian covariance models revealed minor DSe dependence between Q1G and Q2N. The results suggested that SVCV diagnostic tests Q2N and Q1G are suitable for use as diagnostic assays and are fit for presumptive diagnosis, surveillance, and certification of populations or individuals as SVCV free.

1. Introduction

Spring viremia of carp virus (SVCV) causes a systemic hemorrhagic disease in cyprinid and ictalurid fish species with common carp *Cyprinus carpio* L. and koi considered to be the primary hosts (Fijan et al., 1971; Bachman and Ahne, 1973; Ahne, 1977; Wolf, 1988; Ahne et al., 2002; Dixon, 2008). The disease poses a significant risk to wild and cultured fish and possibly amphibians (Ip et al., 2016) and has been listed as notifiable by the World Organization for Animal Health (OIE) (OIE (World Organization for Animal Health), 2019a). Pathogen-specific diagnostic assays validated to internationally recognized standards such as those recommended by the OIE ensure the relevance of test results and provide essential information on virus eco-epidemiology to prevent the introduction of pathogens during the trade or movement of populations (OIE (World Organization for Animal Health), 2019b). Validation of a test involves a description of its analytical performance, diagnostic precision and diagnostic accuracy. The SVCV-specific reverse transcription quantitative polymerase chain reaction (RT-qPCR) tests Q1G and Q2N display analytical and diagnostic precision characteristics that indicate their fitness for detection of SVCV in asymptomatic carriers as well as animals exhibiting clinical signs of spring viremia of carp disease (Clouthier et al., 2021a, 2021b). Virus isolation by cell culture (VI) has been recommended by the OIE for targeted surveillance to demonstrate the absence of SVCV in susceptible fish populations (OIE (World Organization for Animal Health), 2019c). In this study, the diagnostic accuracy of SVCV tests Q1G, Q2N and VI (OIE (World

* Corresponding author.
E-mail addresses: sharon.clouthier@dfo-mpo.gc.ca (S.C. Clouthier), cmcclure@upei.ca (C. McClure), tamara.schroeder@dfo-mpo.gc.ca (T. Schroeder), mainebiotek@hotmail.com (E.D. Anderson).

https://doi.org/10.1016/j.prevetmed.2021.105338
Received 6 December 2020; Received in revised form 17 February 2021; Accepted 21 March 2021
Available online 23 March 2021

Organization for Animal Health), 2019c) was evaluated.

The diagnostic accuracy of a test is measured by its diagnostic sensitivity (DSe) and diagnostic specificity (DSp) (Dybkaer, 1995; OIE (World Organization for Animal Health), 2019b). DSe is the probability that a true positive sample will test positive and DSp is the probability that a true negative sample will test negative. Perfect reference or gold standard populations with known disease status as well as imperfect reference tests have been used to determine DSe and DSp (Dohoo et al., 2009). Gold standard reference tests do not exist for many infectious diseases including SVCV and imperfect reference tests such as the VI test for SVCV can lead to biased estimates of diagnostic accuracy. Latent class analysis is an alternative approach in which diagnostic test results are modelled as a mixture from a population of infected or naïve fish (Hui and Walter, 1980). In the absence of a perfect test, the unknown or latent variable is an animal's most likely infection status and in conventional latent class models, the two classes of latent variables are infected and non-infected (Enoe et al., 2000). Maximum likelihood (ML) estimation (MLE) and Bayesian inference statistical methods can be used to fit these models and generate estimates of population prevalence and test DSe and DSp (Hui and Walter, 1980; Enoe et al., 2000; Branscum et al., 2005).

Intrinsic to latent class models are 3 assumptions: 1) at least 2 tests are used and at least 2 sub-populations with different prevalence levels exist within a population, 2) tests are conditionally independent within each latent class and 3) test accuracy is consistent across populations (Hui and Walter, 1980). Although these assumptions enable model identifiability (i.e. sufficient data exists to estimate model parameters), biased estimates of test accuracy may result if one or more of these assumptions are unknowingly violated (Toft et al., 2005; Johnson et al., 2009). Given the infection status of a fish, conditional dependence of DSe or DSp is often observed between 2 assays that measure similar biological targets (Gardner et al., 2000). For example, the SVCV RT-qPCR tests Q1G and Q2N target viral nucleic acid encoding either the glycoprotein (Q1G) or nucleoprotein (Q2N) gene. The tests share a similar biological target and analytical sensitivity compared to the VI test and consequently conditional dependence may be observed in positive samples, particularly those with a low virus load. One advantage of the Bayesian approach for generating estimates is that the validity of the conditional independence assumption can be verified.

The goals of this study were to 1) generate DSe and DSp estimates for SVCV RT-qPCR tests Q1G, Q2N and the VI test using 2-class LCMs run in ML and Bayesian frameworks, 2) assess the effects of different prior distributions, population structures and datasets on accuracy estimates, 3) verify the validity of the conditional independence assumption using Bayesian inference and 4) use the information to determine if one or more of these tests is fit for the purpose of presumptive diagnosis, surveillance, and certification of populations or individuals as SVCV free.

2. Materials and methods

2.1. Reporting standards

The study is presented in accordance with the Standards for Reporting of Animal Diagnostic Accuracy (STRADAS) – Aquatic (Gardner et al., 2016) and the Standards for Reporting of Diagnostic accuracy studies that use Bayesian Latent Class Models (STARD-BLCM, Kostoulas et al., 2017). Details are provided in Table S1−1, Supplement 1.

2.2. Study design

An overview of the study is provided in Fig. 1. The study was designed prospectively, prior to collection and processing of test samples.

2.3. Viruses

SVCV NC2002, an isolate that originated from a disease outbreak in North Carolina (USA) in 2002 (Goodwin, 2002), was used in the virus exposure study as described in Clouthier et al. (2021b). SVCV isolate HHOcarp06 (Garver et al., 2006) from wild common carp in Ontario (Canada) was used to generate RNA for positive control reference material as described in Clouthier at al. (2021a, 2021b).

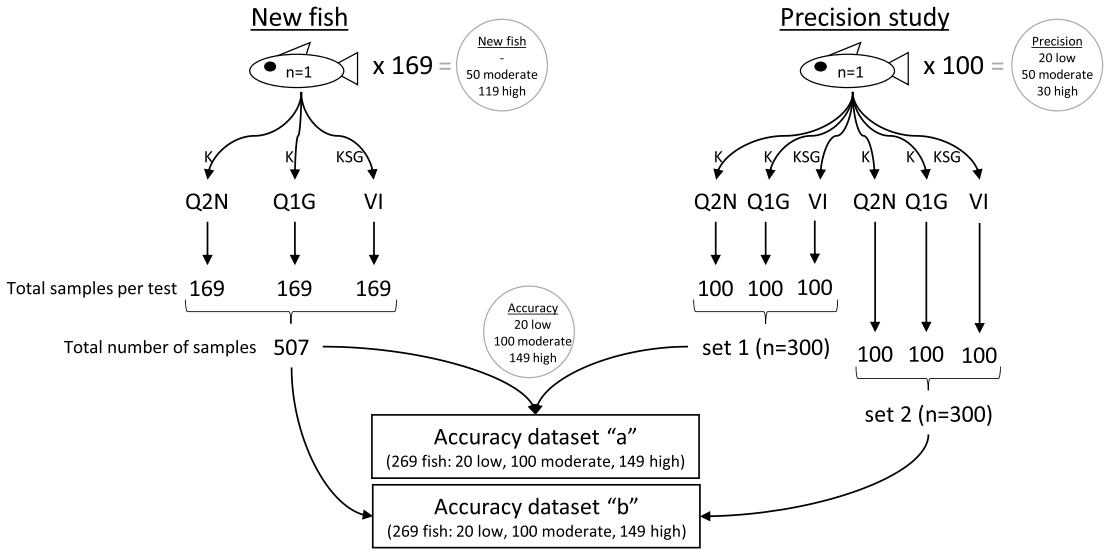


Fig. 1. Accuracy study plan. Schematic of the fish populations, tissue samples, and diagnostic tests used in the study. Koi *Cyprinus carpio* L. were sorted into 2 spring viremia of carp virus (SVCV) prevalence populations (new fish, n = 169), tissue samples were collected and a total of 507 samples were tested from the fish (left column – New fish). Test results from these fish were compiled with those of laboratory A from the precision study (Clouthier et al., 2021b) in which duplicate samples were tested from 100 fish sorted into 3 prevalence populations (right column – Precision study). The compilation resulted in a total of 269 fish from 3 prevalence populations for the study (central circle). The duplicate samples from 100 fish tested in the precision study resulted in two datasets per test for the accuracy study (right column – Precision study). Dataset "a" was used to generate accuracy estimates (top rectangle) and dataset "b" was used to assess the variability of estimates generated from same-fish samples (bottom rectangle) (see also Table 1). Q2N, reverse transcription quantitative PCR (RT-qPCR) assay targeting SVCV nucleoprotein nucleotides; Q1G, RT-qPCR assay targeting SVCV glycoprotein nucleotides; VI, virus isolation; K, kidney; KSG, pool of kidney, spleen and gill tissues.

2.4. Fish populations and study samples

Populations with different SVCV prevalence levels were used to ensure identifiability of latent class models and allowed for assessment of generalizability. Establishing that these fish could be used as a proxy for naturally infected populations was an important consideration because study samples were obtained from koi that had been exposed to SVCV under experimental conditions.

Three populations of koi *Cyprinus carpio* L. were used (n = 269): a negative reference population as well as moderate- or high-prevalence SVCV-positive experimental populations (Fig. 1). Of these 269 fish, 100 were analyzed in a precision study performed previously to evaluate the repeatability and reproducibility of Q1G, Q2N and VI (Clouthier et al., 2021b). The remaining 169 fish were analyzed here using the same 3 tests. Results from the 169 new fish were combined with those from the 100 precision study fish for a total population size of 269 fish (see Section 2.7.1 for more detail). The 169 new koi were derived from a prior SVCV exposure study (Clouthier et al., 2021b) and assigned to the moderate- or high-prevalence populations based on their SVCV status and virus load estimated from pre-testing of kidney tissue with Q1G (Fig. 2). Kidney tissue was collected for the RT-qPCR samples whereas kidney, gill and spleen tissues were collected and pooled for the virus isolation test. Tissue sample collection methods are described in Clouthier et al. (2021b).

The low-prevalence population (n = 20) consisted of naive koi from the previous precision study (Clouthier et al., 2021b) (Fig. 1). The moderate-prevalence population (n = 100) was comprised of 50 koi from the previous precision study (Clouthier et al., 2021b) plus an additional 50 koi from the SVCV-exposed group in this study (Fig. 1). The high-prevalence population (n = 149) consisted of 30 koi from the precision study (Clouthier et al., 2021b) plus an additional 119 koi from the SVCV-exposed group in this study (Fig. 1). Binary test results reported in Clouthier et al. (2021b) for the 100 precision study koi were used here in conjunction with the results for the 169 new koi from this study to generate estimates of test accuracy (Fig. 1) (further details are provided in Section 2.7.1).

2.5. Tests, sample analysis and test data structure

Results reported previously for the precision study samples (Clouthier et al., 2021b) and here for samples unique to this study were produced by one analyst in Laboratory A. The laboratory is located at the Freshwater Institute (Winnipeg, Manitoba, Canada) and is a member of the National Aquatic Animal Health Program. The analyst had extensive experience in each test and had successfully completed a Q1G proficiency panel prior to the diagnostic validation study (Clouthier et al., 2021b).

Samples were analyzed using 3 diagnostic tests – RT-qPCR assays Q1G and Q2N and virus isolation by cell culture on *epithelioma papulosum cyprini* (EPC) cells (Fijan et al., 1983; Winton et al., 2010). The Q1G and Q2N assays were performed as described in Clouthier et al. (2021a) and VI was performed as described by the OIE (OIE (World Organization for Animal Health), 2019c) with details provided in Clouthier et al. (2021b). Each sample was coded and labelled with a 5 digit random number to blind its identity from the analyst. With each batch of samples, positive and negative control samples for the Q2N, Q1G and VI tests were concurrently processed and evaluated as described in Clouthier et al. (2021a, 2021b). The precision study samples (duplicate sets of 100 samples per test) were completed by the end of March 2016 (Clouthier et al., 2021b) and another 507 samples (n = 169 per test) were tested in November 2016 through July 2017 for this study. The latter were used in combination with the previous precision study data to evaluate the DSe and DSp of the 3 assays.

RT-qPCR test results were reported by the analyst as quantification cycle (Cq) (Bustin et al., 2009) values and were then converted into binary data prior to statistical analysis in this study. The average Cq value for the sample run in replicates of 3 was scored as negative (i.e. 0) or positive (i.e. 1) according to the following thresholds: average Cq values <40 were scored as positive and those with average Cq values equal to 40 were recorded as negative.

The VI test results were reported by the analyst as binary data for each of the 2 replicates and then converted into an overall binary score of negative (i.e. 0) if both replicates were negative and positive (i.e. 1) if one or both replicates tested positive (Clouthier et al., 2021b).

2.6. Gold standard reference population method

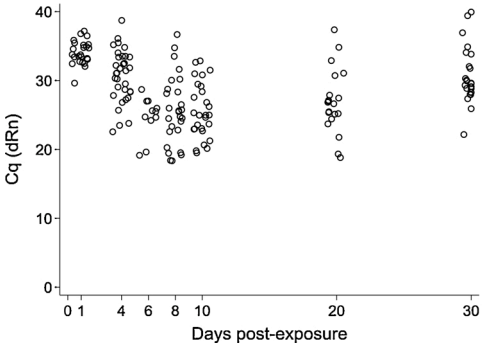
The proportion of negative fish testing negative (i.e. DSp) was determined using 2 × 2 tables as described by McClure et al. (2005) with the low prevalence population used as the gold standard negative reference population.

2.7. Latent class modelling methods - overview

In the absence of defined reference animals, estimates of DSe and DSp were generated with maximum-likelihood and Bayesian estimation procedures which were used to fit 2-class latent class models (LCM) and determine the disease status of each fish that would most likely yield the observed test results (Enoe et al., 2000). Bayesian latent class analysis has become the preferred approach for estimating the diagnostic accuracy of test methods (Johnson et al., 2019). The ML estimation procedures, estimates and code for running the software program "TAGS" ("Tests in the Absence of a Gold Standard", Pouillot et al., 2002) in R are provided in Supplements 2 and 3.

2.7.1. LCM datasets

Latent class analysis was conducted using binary test outcomes collated from the Q1G, Q2N and VI tests for each of the 169 sampled fish in this study in combination with each duplicate set of 100 sampled fish from the precision study (Clouthier et al., 2021b) (i.e. 2 sets of 269 fish). The 269 fish were grouped into 3 prevalence populations (Fig. 1, Table 1).



Time point	Number of fish per time point			
(days post-exposure)	Total	low	moderate	high
0 (naïve)	30	-	-	-
1	65	-	12	12
4	41	-	12	22
6	15	-	3	8
8	32	-	7	20
10	31	-	5	22
20	23	-	3	17
30	32	-	8	18
	269		**50**	**119**

Fig. 2. Q1G pre-testing for spring viremia of carp virus (SVCV) in experimentally infected koi *Cyprinus carpio* L. Quantification cycle (Cq) values from Q1G RT-qPCR pre-testing of kidney tissue from 169 koi collected at 1, 4, 6, 8, 10, 20 and 30 days following intraperitoneal injection with 1×10^6 $TCID_{50}$ of SVCV NC2002 per fish. A random jitter was added to expose the density of the data. The table inset presents the composition of each population relative to each targeted sampling time point. Results generated by laboratory A for the precision samples that are also part of this study are presented in Clouthier et al. (2021b).

Table 1
Composition of accuracy datasets "a" and "b". Estimates of test accuracy were generated with dataset "a". The intra-fish variability of accuracy estimates generated from same-fish samples was evaluated with dataset "b". Precision study set 1 and set 2 refer to duplicate sets of samples from the same 100 fish evaluated previously in the precision study (Clouthier et al. (2021b).

Populations	Accuracy dataset "a"			Accuracy dataset "b"		
	New fish (this study)	Precision study (set 1)	Total (per population)	New fish (this study)	Precision study (set 2)	Total (per population)
Low	–	20	20	–	20	20
Moderate	50	50	100	50	50	100
High	119	30	149	119	30	149
All	169	100	**269**	169	100	**269**

These 2 datasets of 269 fish are referred to as dataset "a" and dataset "b" (Fig. 1, Table 1). Dataset "a" consisted of outcomes for the first set of precision study samples (i.e. set 1) plus the 169 samples from this study. This dataset was selected as the primary test dataset for model analysis. Dataset "b" consisted of outcomes for the duplicate set of precision study samples (i.e. set 2) plus the 169 samples from this study. Dataset "b" was substituted for dataset "a" to assess the intra-fish variability in test accuracy estimates generated using duplicate samples from the same set of fish.

2.7.2. Model structure

The latent class models were structured based on the assumption that there were two latent classes of fish: infected and non-infected. All 3 diagnostic tests were considered to measure the same latent variable (i.e. virus presence/absence) even though the 2 RT-qPCR tests Q1G and Q2N detect SVCV nucleic acid whereas the VI test detects actively replicating SVCV.

2.7.3. Model construction, assumptions and parameters

Models were constructed using the binary test results from all 3 diagnostic assays performed with same-fish tissue samples from 269 koi from 3 different prevalence populations. The 3 tests were assumed to be conditionally independent within each of the two latent classes. Although the 2 RT-qPCR tests share the same target type (i.e. viral nucleic acid), they were considered to be independent given that they detect nucleotides encoding 2 different molecules: the nucleoprotein and the glycoprotein. The validity of this assumption was explored by adding covariance terms to the Bayesian models (see Section 2.8 for more details). For convenience, each new configuration of datasets, populations, priors and/or covariances were considered to represent a new model.

A total of 9 unknown parameters were estimated by the conditionally independent models. For each of the 3 diagnostic tests, estimates included one DSe which represents the probability to test positive in the latent class of infected fish, and one DSp which corresponds to the probability to test negative in the latent class of non-infected fish with an additional prevalence estimate generated for each of the 3 populations (i.e. unless informative priors were used in the models; see Section 2.8 and Supplement 2 for more details). The 2 former parameters were assumed to be constant for each test across the populations. Sufficient degrees of freedom is required to ensure model identifiability. The total degrees of freedom available for these models was 21 (i.e. $K(2^P-1)$ where $P = 3$ tests, $K = 3$ populations) (Hui and Walter, 1980).

2.7.4. Prior specification

Bayesian models were run with non-informative prior distributions on all parameters or an informative prior distribution on the low prevalence population that restricted the prevalence probability to zero. Forcing the prevalence of the low prevalence population to 0 was considered reasonable given the specific disease-free status of the koi obtained for the study and the negative results from our pre-testing of the population for SVCV and cyprinid herpesvirus 3 (Clouthier et al., 2021b).

2.8. Bayesian modelling

Bayesian estimates of population prevalence and diagnostic test DSe and DSp were generated using Markov Chain Monte Carlo (MCMC) implemented in WinBUGS (Lunn et al., 2000; Spiegelhalter et al., 2003). WinBUGS code for Bayes-1 and Bayes-2covGN is provided in Supplements 4 and 5. Posterior distributions were estimated using 50,000 iterations following a burn in of 10,000 iterations. Markov chain convergence diagnostics included trace plots, cumulative quantile plots and autocorrelation plots of MCMC samples. Point estimates of each parameter are presented as the posterior median (with their corresponding 95 % credible intervals).

A total of 8 models were evaluated in the Bayesian framework with all 3 tests and all 3 prevalence populations included in each model (Tables 2, S6−1 (Supplement 6)). A preliminary model was run with dataset "a" using a non-informative beta (1, 1) prior distribution for DSe, DSp and population prevalence (i.e. base model Bayes-1). This model was run with all tests independent of each other (Table 2).

The Bayes-1 model was refined by adding covariance terms to test for possible assay dependence between 2 diagnostic tests within each latent class of fish (i.e. infected and non-infected) (Gardner et al., 2000; Nerette et al., 2008). Conditional covariances between test data in both the infected and non-infected fish were calculated because dependence of test sensitivity (in infected fish) does not imply a similar dependence of test specificity (in non-infected fish). The magnitude of dependence was measured by expressing covariance as a percentage of the maximal possible value identified from the covariance boundaries (Gardner et al., 2000). This exploration of conditional dependence between each possible pair of tests (i.e. Q1G|Q2N, Q1G|VI, Q2N|VI) resulted in 3 additional models referred to as Bayes-2 covGN, Bayes-2 covGV and Bayes-2 covNV (Tables 2, S6−1 (Supplement 6)). Adjusting for dependence among tests was relevant for the RT-qPCR assays because they both target the same biological molecule (i.e. nucleic acid).

Model stability and robustness of the accuracy estimates generated with dataset "a" were evaluated with 2 sensitivity models in which an informative beta (1, 1 000 000) prior distribution restricted the prevalence of the low-prevalence population to 0 (Bayes-1 low) or the low- and moderate-prevalence populations were combined into a single population (Bayes-3) (Table 2). These 2 models were run with all tests independent of each other.

The intra-fish variability of accuracy estimates from models run using dataset "a" was tested by substituting dataset "a" with dataset "b" and re-running the models with a non-informative beta (1, 1) prior distribution (Bayes-1b) or with an informative beta (1, 1 000 000) prior distribution that restricted the prevalence of the low-prevalence population to 0 (Bayes-1b low) (Table 2). These 2 models were run with all tests independent of each other.

Models were evaluated using the deviance information criterion (DIC), Bayesian goodness-of-fit P-values (P-values near 0.5 implied better model fit) and parsimony (Spiegelhalter et al., 2002; Gelman et al., 2003; Nerette et al., 2008; Caraguel et al., 2012). Models with lower DIC values differing by at least 3 units were considered to be better models.

Table 2
Bayesian models used in the accuracy study to generate estimates of diagnostic sensitivity (DSe) and diagnostic specificity (DSp) for the Q1G (G), Q2N (N), and virus isolation (V) tests. Informative prior sets the prevalence for the low prevalence population to 0. cov, covariance. Bayes-3 was run using data from the low and moderate prevalence populations combined.

Variables	Base	Sensitivity analysis		Two-test dependence	Intra-fish variability analysis	
	Bayes-1	Bayes-1 low	Bayes-3	Bayes-2covGN	Bayes-1b	Bayes-1b low
Tests						
Q1G	✓	✓	✓	✓	✓	✓
Q2N	✓	✓	✓	✓	✓	✓
Virus isolation	✓	✓	✓	✓	✓	✓
Covariance						
No covariance	✓	✓	✓		✓	✓
DSe covariance Q1G|Q2N				✓		
DSp covariance Q1G|Q2N				✓		
Populations and priors						
Low: Non-informative prior	✓			✓	✓	
Low: Informative prior (prevalence low = 0)		✓				✓
Moderate	✓	✓		✓	✓	✓
High	✓	✓	✓	✓	✓	✓
Low-moderate combined			✓			
Datasets*						
"a", precision set 1 (n = 100), accuracy (n = 169)	✓	✓	✓	✓		
"b", precision set 2 (n = 100), accuracy (n = 169)					✓	✓

* Datasets "a" (n = 269) and "b" (n = 269) contained same-fish samples from the low (n = 20), moderate (n = 100) and high (n = 149) prevalence populations.

3. Results

3.1. Test results

Indeterminate results (e.g. test failure) or missing data were not observed in the test results reported for the Q2N, Q1G or VI (Fig. S7−1, Supplement 7). All combinations of positive or negative test results for each population and dataset are presented for Q2N, Q1G and VI in Table 3.

3.2. Gold standard reference population method

The DSp of the SVCV tests ranged from 0.95 to 1 for Q1G, 0.9–1.0 for Q2N and 0.85–1 for the VI test, depending on whether dataset "a" or "b" was used for the analysis. In this case, estimates of DSp for Q1G, Q2N and VI were generated using the binary test results from fish belonging to the gold standard negative reference population (i.e. the low prevalence population). These fish were tested twice in the precision study (Clouthier et al., 2021b) and the results from both datasets were used here (Fig. 1, Table 1).

3.3. Latent class modelling

Maximum likelihood and Bayesian latent class analysis methods were applied to generate estimates of DSe and DSp for Q1G, Q2N and VI. The 2 data sets used for this evaluation consisted of Q1G, Q2N and VI binary test results reported for duplicate sets of 100 fish from the precision study (Clouthier et al., 2021b) and an additional 169 fish sampled for this study (Fig. 1). Dataset "a" was selected for analysis whereas dataset "b" was substituted for the first set to assess intra-fish variability in test accuracy estimates generated using duplicate samples from the same set of fish.. For convenience, each new configuration of datasets, populations, priors and/or covariances is considered to represent a new model. Maximum likelihood estimates of population prevalence and test DSe and DSp are presented in Supplement 2.

3.3.1. Bayesian estimates

A total of 8 Bayesian models were tested (Tables 2, S6−1 (Supplement 6)). Models were run with1) non-informative prior distributions for each parameter, 2) an informative prior distribution for disease prevalence or fish grouped into 2 or 3 populations for sensitivity analyses, 3) conditional independence or covariance terms for conditional dependence between tests and 4) dataset "a" or "b" for intra-fish variability in parameter estimates (Tables 2, S6−1 (Supplement 6)). Stable history and quantile plots along with autocorrelation plots were obtained for all models suggesting that MCMC chains converged and attained a stationary distribution.

The model Bayes-1 low was selected as the model of choice for this study because it had the best Bayesian *P*-value and the lowest DIC value (Table 4). The model was run using dataset "a" with all tests

Table 3
All combinations of positive or negative test results for each population and dataset are presented for Q2N, Q1G and VI, respectively.

Dataset	Fish population	SVCV prevalence	Test results (Q2N, Q1G, VI)							Total number
			(+,+,+)	(+,+,-)	(+,-,+)	(+,-,-)	(-,+,+)	(-,+,-)	(-,-,+)	(-,-,-)	
a	I	low	0	0	0	2	0	0	0	18	20
	II	moderate	35	34	0	8	1	4	1	17	100
	III	high	93	47	0	1	1	2	0	5	149
	I to III	combined	128	81	0	11	2	6	1	40	269
b	I	low	0	0	0	0	0	1	3	16	20
	II	moderate	44	25	3	2	1	8	4	13	100
	III	high	97	43	0	1	1	2	1	4	149
	I to III	combined	141	68	3	3	2	11	8	33	269

Table 4
Diagnostic sensitivity and specificity of the Q1G, Q2N and virus isolation tests and spring viremia of carp virus prevalence for the three fish populations. Median posterior estimates (95 % credible interval) from Bayesian models described in Table 2. Cov, covariance; DIC, deviance information criterion; DSe, diagnostic sensitivity; DSp, diagnostic specificity; G, Q1G; N, Q2N; *, significant; ^, percentage of the maximum covariance for all significant parameters.

Parameters	Model category, name and median posterior estimates					
	Base	Sensitivity analysis		2-test dependence	Intra-fish variability analysis	
	Non-informative priors	Informative priors	2 populations	Q1G|Q2N	Non-informative priors	Informative priors
	Bayes-1	Bayes-1 low	Bayes-3	Bayes-2covGN	Bayes-1b	Bayes-1b low
DSe						
Q1G	0.995 (0.975, 1.000)	0.995 (0.974, 1.000)	0.995 (0.976, 1.000)	0.993 (0.966, 1.000)	0.980 (0.949, 0.998)	0.980 (0.949, 0.998)
Q2N	0.976 (0.944, 0.994)	0.976 (0.944, 0.994)	0.981 (0.951, 0.996)	0.968 (0.932, 0.991)	0.984 (0.951, 0.999)	0.984 (0.952, 0.999)
Virus isolation	0.609 (0.541, 0.674)	0.609 (0.541, 0.673)	0.615 (0.547, 0.681)	0.610 (0.542, 0.677)	0.670 (0.605, 0.732)	0.671 (0.605, 0.731)
DSp						
Q1G	0.915 (0.795, 0.991)	0.915 (0.795, 0.990)	0.880 (0.749, 0.969)	0.913 (0.759, 0.993)	0.795 (0.665, 0.905)	0.795 (0.663, 0.906)
Q2N	0.789 (0.667, 0.887)	0.790 (0.667, 0.889)	0.776 (0.646, 0.879)	0.766 (0.628, 0.871)	0.941 (0.820, 0.996)	0.941 (0.820, 0.997)
Virus isolation	0.969 (0.902, 0.996)	0.969 (0.901, 0.996)	0.967 (0.897, 0.996)	0.975 (0.906, 0.999)	0.815 (0.694, 0.907)	0.814 (0.692, 0.908)
Cov						
Se				0.003 (0.000, 0.022)*		
Sp				0.008 (-0.021, 0.087)		
Se%				4.6^		
Prevalence						
Low	0.033 (0.001, 0.162)	0 (0, 0)	0.571 (0.473, 0.663)	0.033 (0.001, 0.162)	0.033 (0.001, 0.162)	0 (0, 0)
Moderate	0.705 (0.604, 0.795)	0.704 (0.604, 0.794)		0.704 (0.588, 0.799)	0.726 (0.622, 0.817)	0.726 (0.622, 0.817)
High	0.952 (0.907, 0.981)	0.953 (0.907, 0.981)	0.950 (0.904, 0.979)	0.954 (0.908, 0.984)	0.951 (0.905, 0.981)	0.951 (0.952, 0.999)
Bayesian *P*-value	0.38	0.39	0.50	0.35	0.37	0.40
DIC	69.45	67.46	56.90	70.02	77.88	75.94

conditionally independent and prevalence restricted to zero for the negative reference population. The DIC value for the comparable model with non-informative prevalence prior distribution (i.e. Bayes-1) was within 3 units of the Bayes-1 low DIC value. Although the 2 values were similar, Bayes-1 low was ranked higher based on the history of the low prevalence population (i.e. fish were from a hatchery that historically tested negative for SVCV) and the negative results obtained from pre-testing each fish from the population with Q1G. Taken together, this information justified restricting SVCV prevalence to zero for the low prevalence population in the model. Refinement of Bayes-1 with the addition of covariance terms between Q1G and Q2N provided evidence of minor dependence with a DSe covariance of 0.003 (0.000, 0.022) (Table 4). Replacing dataset "a" with dataset "b" increased the DIC by 8.4 (Bayes-1 versus Bayes-1b) or 8.5 (Bayes-1 low versus Bayes-1b low) and shifted the DSp estimates (Table 4).

Bayesian posterior distributions for the models are provided in Table 4 with the exception of Bayes-2covGV and Bayes-2covNV in Table S6−2 (Supplement 6). Point estimates for DSe were similar across all models. Higher DSe values were observed for Q1G (i.e. 98–99%) and Q2N (i.e. 97–98%) relative to those for VI (i.e. 61–67%) (across all models). The credible intervals of the Q1G and Q2N estimates overlapped with each other but not with the VI test providing evidence that the DSe estimates for the RT-qPCR tests were statistically different relative to the VI test. The DSp for all 3 tests ranged from 77 to 97 % (across all models). The estimates varied between models according to dataset, with lower estimates for Q2N associated with dataset "a" (79–94%) and lower estimates for Q1G (79–91%) and VI (81–97%) associated with dataset "b" (Table 4; excluding Bayes-3 and covariance models).

The Bayesian *P*-values for the 8 models ranged from 0.29 to 0.50 and the DIC values varied between 56.90 and 77.88. Combining the low and moderate prevalence populations produced accuracy estimates with the Bayes-3 model (Table 4) similar to those obtained with 4 of the other 5 models generated with dataset "a" providing evidence of model stability and robustness. The Bayes-2covNV model that was adjusted for covariance between the Q2N and VI tests produced DSp estimates similar to those produced by models that were run using dataset "b" even though dataset "a" had been used (Table S6−2). In this case, even though the credible interval spanned zero, the addition of another (non-significant) parameter (i.e. covariance) to the model altered the estimates (Table S6−2). The DIC for Bayes-2covNV was significantly higher than the model of choice (i.e. Bayes-1 low) by 11.15 units.

Prevalence estimates were generated for the negative reference population in models where prevalence was not set at zero (3.2–3.5%) and for the 2 positive populations with moderate (70.4–72.6%) and high (94.6–95.4%) SVCV prevalence (Tables 4 and S6−2 (Supplement 6)). The prevalence estimate for the combined low and moderate populations used in the Bayes-3 model was 57 % (Table 4).

4. Discussion

Conventional two-class latent class models were built to generate estimates of DSp and DSe for Q1G, Q2N and VI, in the absence of a perfect reference test for detection of SVCV. The ML and Bayesian approaches were both successful in estimating the diagnostic performance of the 3 tests and yielded consistent results across the two frameworks with one exception. Unlike the Bayesian base model (Bayes-1), the ML base model (MLE-1) failed to converge likely due to mathematical constraints arising from the joint prediction of 0% prevalence for the low prevalence population by the Q1G and VI tests (Enoe et al., 2000). A comparison of the output from equivalent models (i.e. MLE-1 low, Bayes-1 low; MLE-1b, Bayes-1b; MLE-1b low, Bayes-1b low) revealed that the parameter values estimated by the 2 approaches differed by less than 0.03 units, the values for the corresponding lower and upper credible intervals by less than 0.16 units and the range for each credible interval by less than 0.31. Overall, the parameters estimated by both approaches showed the same pattern whereby the 3 populations had distinct prevalence levels, the DSe of the RT-qPCR tests was higher relative to the VI test, and the DSp of all tests varied depending on the dataset used.

Prevalence estimates generated for the 3 fish populations in our study confirmed the non-violation of one of the assumptions inherent to latent class models - that tested individuals are partitioned into populations with different prevalence levels (Hui and Walter, 1980). The

best fitting model was determined to be Bayes-1 low with a prevalence prior probability of 0 for the low prevalence population. For the other two populations, the model estimated SVCV prevalence at 70 % for the moderate prevalence population and 96 % for the high prevalence population. The prevalence of the negative population was estimated to be 3% from Bayes-1, a model with non-informative prevalence prior distributions. In this case, the difference in prevalence levels between populations was distinct, providing evidence that the populations were suitable for latent class modelling. Within the moderate and high prevalence populations, a range of values from 18 to 40 Cq was observed from Q1G pre-testing of the fish new to this study (n = 169) (Fig. 2). The average virus load was approximately 10^6 equivalent plasmid copies per μg total RNA and ranged from $10^{2.3}$ to $10^{7.3}$ equivalent plasmid copies per μg total RNA in the subset of 166 positive fish. These results in conjunction with those of Clouthier et al. (2021b) suggested that the full spectrum of infection status was represented by the samples - non-infected, infected with a low virus load and infected with a high virus load – an important consideration given the influence of biological factors such as the stage and severity of infection on the sensitivity of a diagnostic test (Greiner and Gardner, 2000).

No evidence of false positive or false negative results was observed in the study based on results reported for the positive and negative control samples that were run concurrently with the blinded test samples. However, evidence of misclassification of both test positive and test negative individuals was apparent for all tests based on positive results reported for some samples from the negative reference population consisting of SVCV-free koi and negative results reported for some of the koi from the moderate population that pretested positive for SVCV. These results suggest that both type I and II errors occurred in the study and reflect the uncertainty in estimating disease prevalence using imperfect diagnostic tests. The use of a single imperfect diagnostic test to investigate disease ecology can lead to biases in prevalence estimates that can impact our ability to understand aquatic animal disease dynamics (McDonald and Hodgson, 2018).

The higher DSe of Q1G (99 %) and Q2N (98 %) was credited to their higher analytical sensitivity relative to the VI test (61 %). The estimates of DSe were equivalent across models from both LCM estimation methods indicating that the estimates were stable and robust. Addition of covariance terms between Q1G and Q2N revealed minor dependence between the DSe of the two RT-qPCR tests which can be explained by their recognition of the same biological trait within the tissues. Because Q1G|Q2N covariance only accounted for 4.6 % of the total possible covariance, model DIC increased when covariance terms were added and DSe estimates did not change upon addition of the covariance parameter, the covariance term was not retained in the final model. Dependence among all 3 tests was not evaluated because the size of the sample was too small and because dependence between VI and the two RT-qPCR tests was not observed. The ability to test for conditional dependence between assays was an option available only through the Bayesian estimation framework and underscores its utility in calculating diagnostic test accuracy (Johnson et al., 2019).

Estimates of DSp for Q1G, Q2N and VI from LCM analyses were similar to those predicted using a gold standard negative reference population. However, the LCM estimates of DSp for all 3 tests were not consistent across populations due to the imperfect repeatability observed with datasets “a” and “b” (as initially reported in Clouthier et al., 2021b). Test performance has been shown to vary across populations of fish in different stages of infection (Dendukuri et al., 2009; Caraguel et al., 2012). The variation violates one of the assumptions inherent to LCMs that test accuracy is consistent across all populations (Hui and Walter, 1980). In the current study, the difference in DSp estimates between datasets supports the notion that the probability to test negative will vary among non-infected fish. Thus, the DSp value can be interpreted here to reflect an average across same-fish samples and reveals the potential range in test performance for one sample. In essence, the range in DSp estimates between models reflects the diagnostic uncertainty of each imperfect diagnostic test, highlighting the need to use the results of at least 2 tests to verify the positive disease status of a fish.

Test accuracy estimates should be re-evaluated in the future to determine if more than two latent classes of fish exist in SVCV-infected fish populations. We suspect, as described by Greiner and Gardner (2000), that diagnostic test performance will vary across populations, particularly those with fish at different stages of infection. For example, a third class of infection was investigated in Caraguel et al. (2012) and was characterized as fish that tested negative by virus isolation but positive by RT-PCR. This sub-group of ‘non-infected’ fish is similar to our findings in Clouthier et al. (2021b) as well as in this study whereby Q1G or Q2N positive koi may or may not test positive by virus isolation when the Cq values were >30 (Fig. S7−2, Supplement 7). These disparate results were primarily associated with fish collected early in the infection cycle (i.e. 10 dpc; data not shown). Similarly, the plaque assay produced negative results with samples from koi that tested positive with Q2N between 80 and 167 days following exposure to SVCV (Clouthier et al., 2021a). In this case, the fish were in the later phases of an SVCV infection cycle (i.e. following peak mortality) with the virus possibly shifting to a persistent type of infection in the host. The VI and plaque assays both measure the same biological trait of actively replicating virus. So, in ‘non-infected’ fish that test VI negative and RT-qPCR positive, the virus is either not replicating or is replicating but is being maintained at titers below the limit of detection of these assays. Which scenario is more plausible requires further investigation. Developing a better understanding of SVCV disease dynamics will improve how these tests are applied in a diagnostic setting.

Latent class models were evaluated in the 2 model-fitting frameworks to develop a better understanding of Q1G, Q2N and VI test accuracy. As expected, similar parameter estimates were obtained for a model fit under the two approaches when non-informative priors were placed on all parameters in the Bayesian method. The ML approach was found to be less robust and more restrictive relative to its Bayesian counterpart because it failed to fit the base model MLE-1 and fewer parameters could be estimated. In this respect, the Bayesian approach provided several advantages over its ML counterpart including the option to add an informative prior probability distribution, addition of covariance terms to the model to evaluate the assumption of conditional independence of tests within each latent class, the ease with which probability results from Bayesian analyses can be interpreted and the computation of credible intervals. The option to add informative prior distributions means that models that are not identifiable in the frequentist approach can be fit using a Bayesian framework. Finally, the ML method estimates the mode of the posterior distribution whereas the Bayesian approach provides estimates of the mean and median. An inherent problem with the ML estimation procedure is that it can converge on values that do not represent the best-fitting values and it is not possible to confirm whether the ML estimates represent the true global maximum (Nerette et al., 2008). The computational, interpretative and practical benefits of the Bayesian estimation framework underscore why Bayesian latent class analysis has become the preferred approach for estimating the diagnostic accuracy of test methods (Johnson et al., 2019). Nevertheless, cross-referencing the results from both methods increased confidence in the estimates generated for each parameter.

The use of multiple diagnostic tests is an effective approach for SVCV diagnosis because it reduces any potential bias associated with using a single imperfect test. The diagnostic workflow selected in any given situation can be guided by the intended use of each test. The OIE outlines the most common purposes which include demonstrating freedom from infection in individual animals or their populations, surveillance to establish pathogen-free zones, presumptive diagnosis or confirmation of a presumptive positive diagnosis (OIE (World Organization for Animal Health), 2019b). Ideally, a screening test will have a high DSe to reduce the occurrence of false-negative results and a test used to confirm an

initial positive result will have a high DSp to reduce the risk of false-positive results (OIE (World Organization for Animal Health), 2019b). When using more than one test either in parallel or in series, a trade off in diagnostic performance occurs when one of the tests has lower DSp and/or DSe than the other. The Q1G and Q2N tests are suitable candidates for use as diagnostic screening tests given their high analytical characteristics (Clouthier et al., 2021a), high precision (Clouthier et al., 2021b) and high DSe. Q2N would be the better screening test than Q1G given that the former assay is pan-specific for SVCV (Clouthier et al., 2021b). If VI is used to confirm the presence of actively replicating SVCV in a Q1G or Q2N positive sample, then the DSe of the workflow is limited to the lower performance of the VI test. Using 2 RT-qPCR tests that have similar diagnostic characteristics eliminates the performance disparity but introduces the paradigm of two tests targeting the same biological trait and their potential lack of independence.

5. Conclusions

The Q2N and Q1G tests are suitable for use as diagnostic assays and are fit for presumptive diagnosis, surveillance, and certification of populations or individuals as SVCV free. The DSe of Q1G (>98 %) and Q2N (>96 %) was substantial and higher than the DSe of VI (>60 %). The DSp of all three tests varied by 12–15 % (79–91 % for Q1G, 79–94 % for Q2N and 81–97 % for VI) across same-fish samples revealing the potential range in test performance for one fish. If VI is used to confirm the presence of actively replicating SVCV in a Q2N positive sample, then the DSe of the workflow is limited to the lower performance of the VI test. Using 2 RT-qPCR tests that have similar diagnostic characteristics eliminates the performance disparity but introduces the paradigm of two tests targeting the same biological trait.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgements

This study was supported by the Centre of Expertise for Aquatic Animal Health Research and Development, Fisheries and Oceans Canada. We thank Eveline Emmenegger for providing the 2002 North Carolina isolate of SVCV and Kyle Garver for SVCV isolate HHOcarp06. Special thanks are extended to Philip Byrne and members of his laboratory who conducted the SVCV exposure study of koi at the Gulf Biocontainment Unit-Aquatic Animal Health Laboratory (Charlottetown, Prince Edward Island, Canada). The authors also thank Régis Pouillot for kindly providing permission to reprint the R code for TAGS.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.prevetmed.2021.105338.

References

Ahne, W., 1977. Evidence for the systemic character of Rhabdovirus carpio infection. Bull. Off. Int. Epiz. 87, 435–436.
Ahne, W., Bjorklund, H.V., Essbauer, S., Fijan, N., Kurath, G., Winton, J.R., 2002. Spring viremia of carp (SVC). Dis. Aquat. Org. 52, 261–272.
Bachman, P.A., Ahne, W., 1973. Isolation and characterization of agent causing swim bladder inflammation in carp. Nature 244, 235–237.
Branscum, A.J., Gardner, I.A., Johnson, W.O., 2005. Estimation of diagnostic-test sensitivity and specificity through Bayesian modelling. Prev. Vet. Med. 68, 145–163.
Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin. Chem. 55, 611–622.
Caraguel, C., Stryhn, H., Gagne, N., Dohoo, I., Hammell, L., 2012. Use of a third class in latent class modelling for the diagnostic evaluation of five infectious salmon anaemia virus detection tests. Prev. Vet. Med. 104, 165–173.
Clouthier, S.C., Schroeder, T., Bueren, E.K., Anderson, E.D., Emmenegger, E., 2021a. Analytical validation of two RT-qPCR tests and their detection of spring viremia of carp virus (SVCV) in persistently infected koi *Cyprinus carpio* L. Dis. Aquat. Org. https://doi.org/10.3354/dao03564.
Clouthier, S.C., McClure, C., Schroeder, T., Aldous, S., Allen, J., Collette-Belliveau, C., Li, S., Lindsay, M., Anderson, E.D., 2021b. Measures of diagnostic precision (repeatability and reproducibility) for three test methods designed to detect spring viremia of carp virus. Prev. Vet. Med. https://doi.org/10.1016/j.prevetmed.2021.105288.
Dendukuri, N., Hadgu, A., Wang, L., 2009. Modeling conditional dependence between diagnostic tests: a multiple latent variable model. Stat. Med. 28, 441–461.
Dixon, P.F., 2008. In: Eiras, J.C., Segner, H., Wahli, T., Kapoor, B.G. (Eds.), Virus Diseases of Cyprinids In: Fish Diseases, vol. 1. Science Publishers, Enfield, New Hampshire, USA, pp. 87–184.
Dohoo, I.R., Martin, W., Stryhn, H., 2009. Veterinary Epidemiologic Research, 2nd ed. VER Inc., Charlottetown, PEI, Canada.
Dybkaer, R., 1995. Result, error and uncertainty. Scand. J. Clin. Lab. Invest. 55, 97–118.
Enoe, C., Georgiadis, M.P., 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.
Fijan, N., Petrinec, Z., Sulimanovic, D., Zwillenberg, L.O., 1971. Isolation of the viral causative agent from the acute form of infectious dropsy of carp. Vet. Arh. 41, 125–138.
Fijan, N., Sulimanović, D., Bearzotti, M., Mužinić, D., Zwillenberg, L.O., Chilmonczyk, S., Vautherot, J.F., de Kinkelin, P., 1983. Some properties of the epithelioma papulosum cyprinid (EPC) cell line from carp *Cyprinus carpio*. Ann. Inst. Pasteur Virol. 134E, 207–220.
Gardner, I.A., Stryhn, H., Lind, P., Collins, M.T., 2000. Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. Prev. Vet. Med. 45, 107–122.
Gardner, I.A., Whittington, R.J., Caraguel, C.G.B., Hick, P., Moody, N.J.G., Corbeil, S., Garver, K.A., Warg, J.V., Arzul, I., Purcell, M.K., Crane, M.St.J., Waltzek, T.B., Olesen, N.J., Lagno, A.G., 2016. Recommended reporting standards for test accuracy studies of infectious diseases of finfish, amphibians, molluscs and crustaceans: the STRADAS-aquatic checklist. Dis. Aquat. Org. 118, 91–111.
Gelman, A.E., Carlin, J., Stern, H., Rubin, D., 2003. Bayesian Data Analysis, 2nd ed. Chapman and Hall/CRC Press, Boca Raton.
Greiner, M., Gardner, I.A., 2000. Epidemiologic issues in the validation of veterinary diagnostic tests. Prev. Vet. Med. 30, 3–22.
Hui, S.L., Walter, S.D., 1980. Estimating the error rates of diagnostic tests. Biometrics 36, 167–171.
Ip, H.S., Lorch, J.M., Blehert, D.S., 2016. Detection of spring viraemia of carp virus in imported amphibians reveals an unanticipated foreign animal disease threat. Emerg. Microbes Infect. 5, e97.
Johnson, W.O., Gardner, I.A., Metoyer, C.N., Branscum, A.J., 2009. On the interpretation of the test sensitivity in the two-test two-population problem: assumptions matter. Prev. Vet. Med. 91, 116–121.
Johnson, W.O., Jones, G., Gardner, I.A., 2019. Gold standards are out and Bayes is in: implementing the cure for imperfect reference tests in diagnostic accuracy studies. Prev. Vet. Med. 167, 113–127.
Kostoulas, P., Nielsen, S.S., Branscum, A.J., Jognson, W.O., Dendukuri, N., Dhand, N.K., Toft, N., Gardner, I.A., 2017. STARD-BLCM: standards for the reporting of diagnostic accuracy studies that use Bayesian latent class models. Prev. Vet. Med. 138, 37–47.
Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS – A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing 10, 325–337. https://doi.org/10.1023/A:1008929526011.
McClure, C.A., Hammell, K.L., Stryhn, H., Dohoo, I.R., Hawkins, L.J., 2005. Application of surveillance data in evaluation of diagnostic tests for infectious salmon anemia. Dis. Aquat. Org. 63, 119–127.
Nerette, P., Stryhn, H., Dohoo, I., Hammell, L., 2008. Using pseudogold standards and latent-class analysis in combination to evaluate the accuracy of three diagnostic tests. Prev. Vet. Med. 85, 207–225.
OIE (World Organization for Animal Health), 2019a. Aquatic Animal Health Code 2019, Chap 1.3. Diseases Listed by the OIE. Available at http://www.oie.int/en/international-standard-setting/aquatic-code/access-online/ (Accessed December 4, 2019)..
OIE (World Organization for Animal Health), 2019b. Manual of Diagnostic Tests for Aquatic Animals 2019, Chap 1.1.2. Principles and Methods of Validation of Diagnostic Assays for Infectious Diseases. Available at www.oie.int/international-standard-setting/aquatic-manual/access-online/ (accessed April 14, 2020)..

OIE (World Organization for Animal Health), 2019c. Manual of Diagnostic Tests for Aquatic Animals 2019, Chap 2.3.9. Spring Viraemia of Carp. Available at www.oie.int/international-standard-setting/aquatic-manual/access-online/ (accessed December 4, 2019)..

Pouillot, R., Guillaume, G., Gardner, I.A., 2002. "TAGS", a program for the evaluation of test accuracy in the absence of a gold standard. Prev. Vet. Med. 53, 67–81.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Lind, A., 2002. Bayesian measures of model complexity and fit. J. R. Stat. Soc. Ser. B 64, 583–639.

Spiegelhalter, D.J., Thomas, A., Best, N.G., Lunn, D., 2003. WinBUGS Version 1.4 User Manual. MRC Biostatistic Unit. http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml.

Toft, N., Jorgensen, E., Hojsgaard, S., 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.

Winton, J., Batts, W., deKinkelin, P., LeBerre, M., Bremont, M., Fijan, N., 2010. Current lineages of the epithelioma papulosum cyprinid (EPC) cell line are contaminated with fathead minnow, *Pimephales promelas* cells. J. Fish Dis. 33, 701–704.

Wolf, K., 1988. Spring viremia of carp. Fish Viruses and Fish Viral Diseases. Cornell University Press, Ithaca, USA, pp. 191–216.