Using pseudogold standards and latent-class analysis in combination to evaluate the accuracy of three diagnostic tests

Pascale Nérette [*], Henrik Stryhn, Ian Dohoo, Larry Hammell

Department of Health Management, Atlantic Veterinary College, University of PEI, Charlottetown, PEI C1A-4P3, Canada

Abstract

We previously reported our use of latent-class models to estimate the sensitivity (Se) and specificity (Sp) for each of three tests used to monitor farmed salmon for infectious salmon anaemia virus (ISAv). Those tests were reverse-transcriptase polymerase chain reaction (RT-PCR), virus isolation (VI), and an indirect immunofluorescent-antibody test (IFAT). We used tissues from 403 salmon from four populations presumed to have different prevalence of ISAv. However, no formal evaluation of the assumptions of conditional independence and constant accuracy had been carried out.

In our present study, we "adjusted" that and used two pseudogold standards (a composite reference standard and a study pseudogold), as indicative of the true health status of each fish. The assumption of constant accuracy across populations was evaluated using separate random-effects logistic-regression models for fish classified as D+ or D− (disease positive or negative, according to the pseudogold standards) with study population included in the model to determine if it affected the probability of a positive test result. Where there was evidence of variation in test accuracy across populations, the issue was further investigated using separate latent-class models with informative priors for each study population. Our results suggested that only one PCR test had an accuracy that varied across populations.

The assumption of conditional independence among tests was first evaluated using log-linear models of D+ and D− fish with significant interaction between test results indicative of conditional dependence. Latent-class models which incorporated up to two pairs of between-test dependencies were also fit using Bayesian methods. The two approaches showed considerable evidence of dependence between IFAT and VI and some evidence of dependence between one PCR and IFAT. Results obtained from both maximum-likelihood models and from Bayesian analyses of models allowing for conditional dependence between two pairs of tests were consistent with those obtained with the pseudogold standards. The results suggest that

* Corresponding author at: 954, rue Rabelais, Mont St-Hilaire, Québec J3H-4V6, Canada. Tel.: +1 450 281 0594; fax: +1 450 281 0594.

E-mail address: gregm@videotron.ca (P. Nérette).

pseudogold standards can help in choosing a correct dependence structure and should be used in combination with latent-class models.

1. Introduction

Infectious salmon anaemia (ISA) virus (ISAv) is one of the most important viral pathogens threatening farmed Atlantic salmon in New Brunswick, Canada. Diagnostic tests commonly used for the detection of ISAv in surveillance programs in New Brunswick include virus isolation (VI) on the salmon head kidney cell line (Dannevig et al., 1995), reverse-transcriptase polymerase chain reaction (RT-PCR) (Mjaaland et al., 1997) and indirect fluorescent-antibody testing (IFAT) (Falk and Dannevig, 1995; Falk et al., 1998).

The lack of an appropriate gold-standard test for ISAv makes it difficult to validate the above-mentioned tests. Two approaches for the estimation of the accuracy of diagnostic tests in the absence of a gold standard are: the use of an explicit (but imperfect) pseudogold standard or the use of a latent-class model in which the true disease status is a latent (unknown) variable (Dendukuri and Joseph, 2001).

Two previously described pseudogold standard procedures are discrepant resolution (DR) and composite reference standard (CRS). Both methods involve two stages of testing with a reference test and a resolver test. With DR, initially, the results obtained from the new test are compared to those of the reference standard. Subsequent testing is then performed to resolve the discordant results. The fundamental problem with DR lies in the fact that disease status measurement is conditional upon the test being evaluated (Miller, 1998; Hagdu, 1999). Miller and Hagdu argued that although DR had been widely used for estimating test performance, it should not be adopted for the evaluation of any diagnostic test. To overcome this problem, Alonzo and Pepe (1999) proposed the use of a CRS. With this approach, subsequent testing with a resolver test is performed on all samples that tested negative on the reference test.

Latent-class analysis is a statistical technique that allows the evaluation of the accuracy of diagnostic tests when there is no gold standard. In its standard and most commonly used form, the model involves three assumptions: (i) the target population should consist of two (or more) subpopulations with different prevalences; (ii) the sensitivity (Se) and specificity (Sp) (together referred to as the ''accuracy'') of the diagnostic tests should be constant across subpopulations (i.e. the ability of a test to detect infected individuals should be the same regardless of whether the test is used in a population with a high prevalence of infection or one with low prevalence); and (iii) the tests should be conditionally independent given the disease status (Enøe et al., 2000). Both maximum-likelihood (ML) and Bayesian estimation procedures can be used to fit latent-class models (see e.g. Hui and Zhou, 1998; Enøe et al., 2000, for reviews of the earlier literature). Briefly, ML estimates are a set of parameter estimates that were most likely to have generated the observed data and are obtained by maximizing the likelihood function. Alternatively, the Bayesian approach can be used to combine a priori scientific knowledge about unknown parameters with the information contained in the likelihood based on the observed data. Latent-class models can be extended to account for dependence among tests (Espeland and Handelman, 1989; Qu et al., 1996; Yang and Becker, 1997) to relax the assumption of conditional