Preventive Veterinary Medicine xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

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

An approach to evaluating the reliability of diagnostic tests on pooled groups of infected individuals

L.M. Hall*, L.A. Munro, I.S. Wallace, R. McIntosh, K. MacNeish, A.G. Murray

Marine Scotland Science, Aberdeen, AB11 9DB, UK

ARTICLE INFO

Article history:
Received 4 March 2013
Received in revised form 23 October 2013
Accepted 21 January 2014

Keywords:
Pools
Pancreas Disease
Salmonid alphavirus
Quantitative real-time polymerase chain-reaction
Virus isolation

ABSTRACT

An experimental design and statistical analysis providing information on the reliability of pooled test procedures is described. It involves estimating the relationship between the probability of a positive pooled test result (dependent variable) and the expected number of infected individuals in a pool (explanatory variable). The intercept is an estimate of the proportion of false positives (1 – pooled specificity) and pooled sensitivities can be estimated for indicative prevalences of infected individuals. Simulations for a theoretical infection are used to investigate the advantages and limitations of the approach. The approach is used to evaluate the reliability of a virus isolation and qRT-PCR test procedure detecting *Salmonid alphavirus* the pathogenic agent necessary for the development of Pancreas Disease in Atlantic salmon (*Salmo salar*).

1. Introduction

Field evaluations of the reliability of veterinary diagnostic test procedures (hereafter referred to as tests) which detect infection, illness or disease (hereafter referred to as infection) are becoming increasingly available. Within the context of salmonid aquaculture, for example, evaluations of tests detecting infection with *Nucleospora salmonis* and *Infectious salmon anemia virus* have been published (e.g. Georgiadis et al., 1998; Nérette et al., 2005). These evaluations provide estimates of diagnostic sensitivity and specificity as measures of the probability that a test will correctly identify either infected individuals as being infected (Se) or uninfected individuals as being uninfected (Sp) respectively. Such estimates are regarded as primary test performance indicators by the World Organisation for Animal Health (OIE) (World Organisation for Animal Health, 2009).

The pooling of samples from several individuals for a single test has long been advocated as a way of reducing the cost and effort of diagnostic testing (Dorfman, 1943). In a veterinary context pooling has been used for the identification of infected individuals (e.g. Kennedy, 2006) and populations (e.g. Kinde et al., 1996) and to estimate the prevalence of infected individuals (e.g. Raizman et al., 2011) and populations (e.g. McBeath et al., 2009). The OIE recognises the utility of testing pools (World Organisation for Animal Health, 2009, 2013) and, with regard to internationally listed diseases, stipulates that the results of such tests are interpreted using estimates of diagnostic sensitivity and specificity for pools (PSe and PSp respectively) (World Organisation for Animal Health, 2013).

Experimental designs and statistical methods used to evaluate the reliability of tests on individuals in the veterinary context are well established (Enøe et al., 2000; Branscum et al., 2005). In contrast approaches using pools are less developed with, to the best of our knowledge, no published standardised approach available which provides information on pooled test reliability of a type consistent with OIE requirements. The primary purpose of this report is, therefore, to promote debate on the experimental

* Corresponding author.
E-mail address: malcolm.hall@scotland.gsi.gov.uk (L.M. Hall).

0167-5877/$ – see front matter.
http://dx.doi.org/10.1016/j.prevetmed.2014.01.021

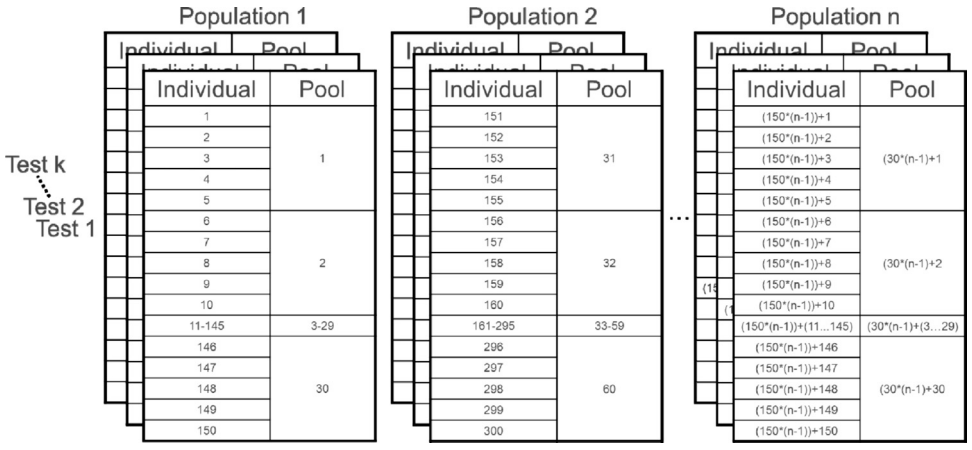

Fig. 1. Modified experimental design for estimating test reliability for individuals and pools.

designs and statistical analyses most suited to evaluating the reliability of pooled tests. This is done by describing a simple approach for evaluating the reliability of pooled tests and illustrating this using both an hypothetical and a real infection.

2. Outline of the approach

This is a modification to an experimental design formalised by Hui and Walter (1980) and widely used to evaluate individual test reliabilities. The original design involves sampling individuals from two populations characterised by different within-population individual infection prevalences (TP) and testing the individuals using two tests detecting the same infection. This original design has since been generalised to more than two populations and/or tests (Branscum et al., 2005). The modification described in this report involves the additional random assignment of individuals to groups comprising a small fixed number of individuals (pools) and subjecting the pools to the same diagnostic tests; the experimental structure is shown in Fig. 1. This modified experimental design therefore involves the testing of each individual both as an individual and as a part of a pool. It is advisable to select populations characterised by substantially different TP.

The statistical analysis starts with the familiar estimation of individual Se and Sp and TP (Branscum et al., 2005). Estimates of these parameters can be obtained from the posterior distributions of Markov Chain Monte Carlo (MCMC) chains generated by slice sampling (Neal, 1997) using vague priors utilising the statistical model of either Hui and Walter (1980) for tests which are assumed to be conditionally independent or, providing there are sufficient degrees of freedom, the model of Dendukuri and Joseph (2001) for tests which are likely to be conditionally dependent. The subsequent analysis for pooled tests, introduced in this report, involves calculating the probability of each individual being infected using the estimates of Se, Sp and TP and thereafter enumerating the expected number of infected individuals in each pool. The relationship between the probability of a positive pooled test result as the dependent variable and the expected number of infected individuals in a pool as an explanatory variable is then modelled with a Generalised Linear Model assuming a binomial error distribution and utilising the logistic link function. This is fitted by penalised maximum likelihood (Firth, 1993) which, in addition to reducing asymptotic bias, is stable and provides estimates for data with complete or quasi-separation. The regression parameters are then used to predict values for PSe at different TP and also PSp. The ability of a test to detect pools containing infected individuals can also be characterised by the median effective dose (ED_{50}), defined as the number of infected individuals required for a pool of fixed size to test positive 50% of the time.

There are several problems with the approach described in the previous section. One of these is that ED_{50} is not a preferred OIE test reliability parameter (World Organisation for Animal Health, 2013) and estimates of PSe therefore need to be standardised using defined TP. This is overcome by using the OIE test validation scenarios which include testing for freedom from disease and confirmatory diagnosis of suspect clinical cases (World Organisation for Animal Health, 2009). Testing for freedom from disease requires the detection of a low TP recommended to be two percent in the absence of reliable information (World Organisation for Animal Health, 2013). Confirmatory diagnosis of suspected infection usually involves targeted sampling of suspect individuals within a suspect population and an indicative TP of 0.80 has been set by the authors. Given these indicative prevalences it is then straight-forward to estimate the proportion of pools containing different numbers of infected individuals and generate estimates of PSe relevant to the problem at hand. The abbreviations $PSe_{0.02}$ and $PSe_{0.80}$ are assigned to the probability of a positive test result for pools containing one or more infected individuals from populations characterised by TP of 0.02 and 0.80 respectively. Other problems with the approach will become apparent throughout the report. The process, as described in this section, is summarised in Fig. 2.

3. Results for a theoretical infection

Simulated data for imaginary populations infected with an imaginary pathogen at different TP and tested

Please cite this article in press as: Hall, L.M., et al., An approach to evaluating the reliability of diagnostic tests on pooled groups of infected individuals. PREVET (2014), http://dx.doi.org/10.1016/j.prevetmed.2014.01.021