This document is a postprint version of an article published in Preventive Veterinary Medicine © Elsevier after peer review. To access the final edited and published work see https://doi.org/10.1016/j.prevetmed.2019.02.003

Document downloaded from:
Diagnostic performance of direct and indirect methods for assessing failure of transfer of passive immunity in dairy calves using latent class analysis

Ibrahim Elsohaby\textsuperscript{a,b,*}, Marshal M. Mweu\textsuperscript{c}, Yasser S. Mahmod\textsuperscript{b,d,e}, J. Trenton McClure\textsuperscript{a}, Greg P. Keefe\textsuperscript{a}

\textsuperscript{a} Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, Prince Edward Island, Canada

\textsuperscript{b} Department of Animal Medicine, Division of Infectious Diseases, Faculty of Veterinary Medicine, Zagazig University, Zagazig City, Sharkia Province, Egypt

\textsuperscript{c} School of Public Health, College of Health Sciences, University of Nairobi, Nairobi, Kenya.

\textsuperscript{d} IRTA, Centre de Recerca en Sanitat Animal (CReSA, IRTA), Campus de la Universitat Autònoma de Barcelona, 08193-Bellaterra, Cerdanyola del Vallès, Barcelona, Spain

\textsuperscript{e} Universitat Autònoma de Barcelona, 08193-Bellaterra, Cerdanyola del Vallès, Barcelona, Spain

*Corresponding author: I. Elsohaby, Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, Prince Edward Island, Canada C1A 4P3; Tel: +1 902 566-6063; Fax: +1 902 566-0823; e-mail: ielsohaby@upei.ca
Abstract

Accurate diagnosis of failure of transfer of passive immunity (FTPI) in newborn calves is an essential component of dairy farm management plan. Several methods (direct and indirect) are available for diagnosis of FTPI in dairy calves. However, the indirect methods offer an advantage over the direct methods in not requiring an experienced veterinarian, rapid, cost efficient and can be performed under field-setting. The objective of this study was to estimate the diagnostic performance of radial immunodiffusion (RID) assay, transmission infrared (TIR) spectroscopy and digital Brix refractometer for diagnosis of FTPI in dairy calves using latent class models at four cut-off values of digital Brix refractometer. Holstein calves (n = 691) from 40 commercial dairy farms in the four Atlantic Canada provinces were blood-sampled and tested for detection of FTPI. Results showed that the number of calves with FTPI was 253 (36.6%) by RID, 194 (28.1%) by TIR and 204 (29.5%) by Brix refractometer at cut-off value of 8.2%. Estimates of Se_{RID} was higher than Se_{TIR} and Se_{Brix}, at all Brix refractometer cut-offs, but with increase of Brix refractometer cut-off from 8.2 to 8.5%, Se_{RID} and Se_{TIR} were decreased from 96.0% (95% PCI: 88.0-99.0) and 79.0% (95% PCI: 70.0-85.0), to 92.0% (95% PCI: 77.0-99.0) and 74.0% (95% PCI: 61.0-82.0), respectively. Sp_{RID} and Sp_{TIR} were always higher than Sp_{Brix} at all tested cut-offs and were above 92.0%, and 96.0%, respectively. With increasing the cut-off of Brix refractometer from 8.2 to 8.5%, Se_{Brix} estimate has remarkably increased from 79.0% (95% PCI: 70.0-96.0) to 95.0% (95% PCI: 87.0-100.0), respectively. Whilst, Sp_{Brix} was decreased from 95.0% (95% PCI: 91.0-98.0) at cut-off 8.2% to 84.0% (95% PCI: 78.0-94.0) at cut-off 8.5%. In conclusion, RID has a higher Se than TIR and Brix, if the latter is used with cut-offs of 8.2% or 8.3%. However, the higher the cut-off, the more comparable sensitivities of RID and digital Brix refractometer. The median estimate of Sp_{TIR} were always higher than Sp_{RID} and Sp_{Brix} at all tested cut-offs. However, the 95% confidence intervals estimates of the three tests were overlapping across the tested cut-offs of digital Brix refractometer reflecting the inability to prefer a test over the other based on the Sp estimate.
Keywords: Calves, FTPI, RID, Infrared spectroscopy, Refractometer, Latent class analysis

1. Introduction

Newborn calves are born hypogammaglobulinemic, without circulating immunoglobulin G (IgG), due to the diffuse epitheliochorial structure of their placentation that does not allow for the passage of IgG from the dam to the fetus (Smith, et al., 1964). Therefore, newborn calves rely upon the acquisition of passive immunity via transfer of maternal IgG through colostrum intake (Butler, 1983; Godden, 2008). Inadequate transfer of colostral IgG (threshold <1000 mg/dl) to newborn calves within 48h of birth is defined as failure of transfer of passive immunity (FTPI) (Godden, 2008). There is a recognized association between FTPI and reduced calf growth rate as well as increased risk of neonatal infectious diseases (Tyler, et al., 1999). Effects of FTPI might extend beyond the neonatal period, affecting long-term productivity and resulting in decreased milk yields and increased culling rates during the first lactation period (DeNise, et al., 1989; Heinrichs and Heinrichs, 2011). The reported prevalence of FTPI in Canada is relatively high, estimated at 25 to 37% (Wallace, et al., 2006; Trotz-Williams, et al., 2008), 12% in the United States (Dairy, 2007), and 38% in Australia (Vogels, et al., 2013). Although early diagnosis of FTPI is an important part of the dairy herd management practices to ensure its timely detection and thus implementation of interventional measures, only 6% of dairy farms are routinely screened for FTPI (USDA, 2016).

Several methods have been developed to assess FTPI in dairy calves by measuring circulating IgG concentration either directly or indirectly. Direct methods include radial immunodiffusion (RID) assay (McBeath, et al., 1971), transmission infrared (TIR) spectroscopy (Elsohaby, et al., 2014), enzyme-linked immunosorbent assay (ELISA) (Filteau, et al., 2003) and an automated turbidimetric immunoassay (Alley, et al., 2012). The indirect methods include the sodium sulfate turbidity test (Pfeiffer and McGuire, 1977), zinc sulfate (ZnSO4) turbidity test (McEwan, et al., 1970), glutaraldehyde coagulation test (Tyler et al., 1996) and refractometry (Naylor and Kronfeld, 1977).
The RID assay is the classical direct and quantitative reference method for measuring IgG concentrations (McBeath, et al., 1971). The RID assay is expensive and utilizes reagents with a limited shelf life. The RID procedure takes 18 to 24h to obtain the results and requires advanced technical skill to not only perform the assay but also measure the zones of precipitation accurately (Riley, et al., 2007; Bielmann, et al., 2010). Further, imprecision in RID assay replicates from the same sample have been noted and are attributed to inconsistencies in the assay standards (Ameri and Wilkerson, 2008). Recently, TIR spectroscopy has been touted as a suitable alternative for measuring bovine serum IgG and, as such, for assessing FTPI in dairy calves (Elsohaby, et al., 2014) owing to its rapidity, low cost and its minimal requirement in sample preparation (Shaw, et al., 1998). Refractometers are an effective indirect tool for assessing FTPI in dairy calves under field settings, because immunoglobulins are the major constituents of total protein in neonatal calf blood, besides albumin; and albumin is relatively stable in healthy individuals (Deelen, et al., 2014; Cuttance, et al., 2017). The correlation between RID-determined IgG concentrations and serum total protein has been determined in previous studies (Deelen, et al., 2014; Elsohaby, et al., 2015; Cuttance, et al., 2017). However, the results of refractometry are affected by instrument quality, ambient temperature, calf age and health status (Wallace, et al., 2006; Thornhill, et al., 2015).

The diagnostic test characteristics of TIR spectroscopy and digital Brix refractometer have formerly been evaluated using an imperfect RID assay as the reference standard (Deelen, et al., 2014; Elsohaby, et al., 2016). Owing to the imperfection of reference tests, Bayesian latent class analysis models (BLCMs) present a suitable option for the simultaneous estimation of Se and Sp of two or more tests without any assumption about the underlying true disease status of each subject (Hui and Walter, 1980). Basically, these models are premised on three key assumptions: (1) the target population should consist of two or more subpopulations with different prevalences, (2) the sensitivity and specificity of the index tests should be constant across the subpopulations and (3) the tests should be conditionally independent (CID) given the disease status (Hui and Walter, 1980). To
the best of our knowledge, there are no published studies that apply a BLCM framework to quantify the accuracy of RID, TIR-spectroscopy and digital Brix refractometer for the assessment of FTPI in dairy calves. Therefore, the objective of this study was to estimate the diagnostic performance of RID, TIR spectroscopy and digital Brix refractometer for diagnosis of FTPI in dairy calves within a Bayesian framework.

2. Materials and methods

2.1. Study population

Forty commercial Holstein dairy herds were selected to participate in a study to investigate colostrum and calf health management practices in Atlantic Canada dairy herds (Prince Edward Island (PE; n = 5), Nova Scotia (NS; n = 21), New Brunswick (NB; n = 8) and Newfoundland (NL; n = 6)). The herds were selected by the veterinary clinics in the study area. To be eligible for inclusion in the study, herds had to provide at least 10 blood samples.

2.2. Sample collection

A total of 691 holstein calves from PE (n = 203), NS (n = 218), NB (n = 210) and NL (n = 60) were blood-sampled between June 2013 and September 2015. Sampling was conducted in accordance with the Canadian Council on Animal Care guidelines (Care, Canadian Council On Animal, 2009) under a protocol approved by the Animal Care Committee at the University of Prince Edward Island (UPEI; protocol #6006206). Specifically, within each farm, whole blood was collected from 1 to 14 day-old calves by jugular venipuncture, using a 20-gauge, 1-inch hypodermic needle (BD Vacutainer Precision Glide, Becton Dickinson Co., Franklin Lakes, NJ), into a sterile, plastic vacutainer tube (BD Vacutainer, Becton Dickinson Co.) without anticoagulant. Samples were labeled with a calf identification number, dated and then stored on the farm at 20°C refrigeration until transported to the Maritime Quality Milk Laboratory, UPEI. Serum was separated by
centrifugation at $1500 \times g$ for 10 min at $\sim20^\circ C$. Serum samples were divided into 3 aliquots and stored at $-80^\circ C$.

2.3. Direct diagnostic tests

2.3.1. RID assay

Serum samples were allowed to thaw at room temperature (20 to 24°C) and vortexed at a maximum of 2700 rpm for 10 s. Subsequently, IgG was measured by a commercial RID assay (Bovine IgG RID Kit, Triple J Farms, Bellingham, WA). The RID assay was performed according to the manufacturer’s instructions, using $5 \, \mu l$ of undiluted serum sample in each well. The diameter of precipitated rings was measured using a hand-held caliper after 18 to 24 h of incubation at room temperature. Each sample and assay standard was tested in replicates of two. The averages of the 2 replicates of the assay standards were used to build a calibration curve that was subsequently used to determine IgG concentrations for the serum samples. The final IgG concentration for each sample was determined by calculating the average of the 2 replicates. Serum samples with IgG concentrations greater than the manufacturer’s stated performance range for the assay (>3000 mg/dl) were diluted (1:1) with deionized sterile water and retested. The IgG concentration of 1000 mg/dl was used as a cut-point to differentiate calves with and without FTPI (Weaver, et al., 2000; Godden, 2008).

2.3.2. TIR spectroscopy

Thawed serum samples were diluted (1:1) with deionized sterile water and vortexed at a maximum of 2700 rpm for 10 s. Each diluted serum sample was tested in replicates of 6 by evenly spreading 10-μl aliquots into 5-mm diameter wells within an adhesive-masked, 96-well silicon microplate (Riley et al., 2007). An empty well served as the background reference for each
microplate. The loaded microplates were allowed to dry at room temperature (20–24°C) for 2 h, to produce dried, thin films.

For collection of the spectra, the microplates were inserted into a multisampler (HTS-XT Autosampler, Bruker Optics, Milton, ON, Canada) interfaced with an infrared (IR) spectrometer (Tensor 37, Bruker Optics) equipped with a deuterium tryglycine sulfate detector and controlled by proprietary software (OPUS ver. 6.5, Bruker Optics). A total of 4146 (691 samples × 6 replicates) spectra were collected over the wavenumber range between 4000 and 400 cm⁻¹ with a nominal resolution of 4 cm⁻¹, with 512 scans collected for data acquisition. Collected spectra were converted into a printable format using manufacturer’s software (GRAMS/AI ver.7.02, Thermo Fisher Scientific Inc., Waltham, MA). The printable format spectral data were imported into MATLAB (MathWorks R2016a, Natick, MA), and then preprocessed using the techniques previously described for developing the partial least squares (PLS) regression model (Elsohaby et al., 2014).

A previously developed PLS model built for prediction of serum IgG concentration from IR spectra (Elsohaby et al., 2014) was used to predict serum IgG concentrations. The IgG concentration was predicted from each spectrum; and, subsequently, the IgG concentration for each serum sample was calculated as the average of the 6 replicate IgG values. The IgG concentration of 1000 mg/dl was used as a cut-point to differentiate calves with and without FTPI (Godden, 2008; Weaver et al., 2000).

2.4. Indirect diagnostic test

2.4.1. Digital Brix refractometer

Thawed serum samples were vortexed at a maximum of 2700 rpm for 10 s and then tested using a digital Brix refractometer (PAL-1, Atago Co. Ltd., Bellevue, WA), with a scale from 0 to 52% Brix. Approximately, 250 μl of serum were used, with the Brix score of the liquid determined by shining a light through the sample in the prism followed by measuring the index of refraction and...
reading off the percentage Brix on a digital scale. The refractometer was cleaned and re-calibrated with distilled water at room temperature before each analysis. Previous studies have reported that Brix scores of 8.2% (Morrill, et al., 2013), 8.3% (Elsohaby, et al., 2015), 8.4% (Deelen, et al., 2014) and 8.5% (Hernandez, et al., 2016), may be used to identify FTPI in dairy calves.

2.5. Population classification

Dairy calf data derived from the four Canadian provinces (PE, NB, NS, and NL) constituted four calf populations that were presumed to have different FTPI prevalences. These populations formed the basis for the estimation of the Se and Sp of the three diagnostic tests.

2.6. Statistical analysis

A Bayesian latent class model fitted in OpenBUGS v3.2.2 (Thomas et al., 2006) was used to infer the Se and Sp of the three tests (RID, TIR and digital Brix refractometer), as well as the four population prevalences as per the standards for reporting diagnostic accuracy studies that use BLCMs (Kostoulas et al., 2017). The Bayesian model was implemented in R version 3.5.1 using (BRugs) package. In particular, the accuracy of the diagnostic tests was assumed to be similar across the study populations, i.e. Se and Sp constancy. However, granted that the two direct tests (RID and TIR) are based on related measurement mechanisms for FTPI but independent of the indirect digital Brix test i.e. are conditionally correlated, we allowed for dependence between the two tests by adding two conditional covariance parameters, $\phi_{11}$ and $\phi_{22}$, between pairs of the Se and Sp of the tests respectively as specified by (Gardner, et al., 2000). Notably, the performance of the digital Brix test was evaluated based on four pre-identified cut-off values: 8.2% (Morrill, et al., 2013), 8.3% (Elsohaby, et al., 2015), 8.4% (Deelen, et al., 2014) and 8.5% (Hernandez, et al., 2016), thus resulting in four cut-off-specific models.
Counts \((Q_k)\) of the different test combinations (e.g. \(+,+,+\)) were assumed to follow a multinomial distribution of the form:

\[Q_k \mid S_{t_k} P_k \sim \text{multinomial}(\text{prob}_k, n_k)\]

Where \(S_{t_k}\) and \(P_k\) represent the respective test characteristics for test \(t\) \((t = 1, 2, 3)\) and \(P_k\) is the specific prevalence for the \(k\)th \((k = 1, 2, 3, 4)\) population. \(\text{prob}_k\) is a vector of probabilities of observing the different combinations of test results, whereas \(n_k\) reflects the total number of calves tested for the \(k\)th population. For instance, in the 1st population for an individual testing positive to each of the three tests, \(\text{prob}_1\) is given by:

\[
\text{prob}_1 = (\mu_1(1 - T_1^+)T_2^+T_3^+) + \mu_1(1 - T_1^+)T_2^-T_3^-) = (S_{t_1} + c_{t_2})S_{t_2}P_1 + [(1 - S_{t_2})(1 - S_{t_3}) + c_{t_3}][(1 - S_{t_2})(1 - S_{t_3})[1 - P_1]]
\]

Given four populations, the available data furnished 28 degrees of freedom sufficient to estimate 12 parameters (Se and Sp of the three tests, four prevalences and two conditional covariances), essentially yielding an identifiable model. Non-informative priors \(\text{beta}(1, 1)\) were used to fit the Bayesian model since no reliable prior information was available for any of the aforementioned parameters. Notably, the hypothesis: \(H_0\) \(c_{xy} = 0\), was evaluated using a Bayesian \(P\) -value.

The model was initialised with two Markov Chain Monte Carlo chains with different values. Each chain comprised 70000 samples, with the first 20000 being discarded as the burn-in (supplementary file). Convergence of the chains was evaluated by visual appraisal of the time series plots of selected variables and the Gelman-Rubin diagnostic plots. The posterior distribution of the population prevalences, the Se and Sp of the three tests, as well as the conditional covariances were reported as the median and the corresponding 95% posterior credibility intervals (PCI).
3. Results

The cross-classified counts of the three tests outcomes at each of the digital Brix cut-off levels and specific calf populations are displayed in Table 1. Based on these tabulated data, the Se and Sp for RID, TIR and digital Brix refractometer, together with the conditional covariance parameters between the RID and TIR test characteristics were derived. The true prevalences of FTPI in dairy calves in the four Canadian provinces are presented in Table 2.

RID showed the best Se and Sp estimates at all the tested cut-offs of digital Brix refractometer, but the higher the cut-off, the more comparable sensitivities of RID and digital Brix refractometer became until a cut-off of 8.5%. With increasing digital Brix refractometer cut-off, Se of RID assay and TIR spectroscopy decreased from 96.0% (95% PCI: 88.0-99.0) and 79.0% (95% PCI: 70.0-85.0), (at lower cut-off 8.2%) to 92.0% (95% PCI: 77.0-99.0) and 74.0% (95% PCI: 61.0-82.0), (at higher cut-off 8.5%), respectively. With increasing the cut-off of digital Brix refractometer from cut-off 8.2% to 8.5%, the Se estimate remarkably increased from 79.0% (95% PCI: 70.0-96.0) to 95.0% (95% PCI: 87.0-100.0), which is a good improvement. However, Sp estimate of digital Brix refractometer sharply decreased from 95.0% (95% PCI: 91.0-98.0), (at cut-off 8.2%) to 84.0% (95% PCI: 78.0-94.0), (at cut-off 8.5%). Sp of RID assay and TIR spectroscopy were always higher than Sp of digital Brix refractometer at all the tested cut-offs and were above 92.0%, and 96.0%, respectively (Table 2). In the model, since the parameter, $\alpha_{2p}$ was significant (Bayesian P-value=), the dependence model was retained for subsequent analyses.

4. Discussion

Characteristics of tests estimates

Using a Bayesian framework, we estimated the Se and Sp of RID, TIR spectroscopy and digital Brix refractometer to determine presence or absence of calf FTPI at different cut-off values of the digital Brix refractometer without the assumption of a reference standard. Previous studies have
compared TIR spectroscopy and digital Brix refractometer with the assumption of RID as gold standard for measuring serum IgG concentration and diagnosis of FTPI in calves (Deelen, et al., 2014; Elsohaby, et al., 2016). The analysis showed that RID has a higher Se than TIR and Brix, if the latter is used with cut-offs of 8.2% or 8.3%. However, the higher the cut-off, the more comparable sensitivities of RID and digital Brix refractometer. The possible explanation for high Se_{RID} is that RID assay measures IgG concentration through antigen-antibody precipitation (McBeath, et al., 1971). However, the TIR spectroscopy and Brix refractometer measure serum proteins and then either use mathematical models to directly extract the IgG concentration for TIR spectroscopy (Elsohaby, et al., 2016) or relate the Brix scores to IgG concentration for refractometry (Deelen, et al., 2014). The TIR spectroscopy demonstrated lower Se (79%) and similar Sp (97%) at cut-off 8.2% than that previously reported by (Elsohaby, et al., 2016) [Se = 0.87% (0.76–0.95) and Sp = 0.97% (0.92–0.99)], which could be argued by the using of imperfect reference standard. However, the Se and Sp of the digital Brix refractometer fall within the range of those reported previously by (Deelen, et al., 2014) [Se = 88.9% and Sp = 88.9%], (Elsohaby, et al., 2015) [Se = 85.5% and SP = 83.5%], and (Hernandez, et al., 2016) [Se = 100% and Sp = 89.2%].

Various cut-off values have been recommended as an appropriate cut-off for diagnosis of calves with FTPI using digital Brix refractometer (Deelen, et al., 2014; Elsohaby, et al., 2015). In this study, our findings showed that the test estimates are changed at the different cut-offs of BRIX test, especially the Se estimates, meanwhile the 95% CI of the Sp estimates of the three tests at the different cut-offs were overlapping. Previous studies recommended different cut-offs for diagnosis of FTPI in newborn calves, for example, Morrill, et al. (2013) recommended cut-off value of 8.2%. In contrast, other studies recommended other cut-offs 8.2% (Elsohaby, et al., 2015), 8.4% (Deelen, et al., 2014), 8.5% (Hernandez, et al., 2016) and 8.8% (Cuttance et al., 2017). The possible reasons for this difference could be related to the use of refractometers from different manufacturers, as well as age, breed and health status of calves involved in these studies (Wallace, et al., 2006; Thornhill, et
al., 2015). In addition, previous studies used traditional diagnostic evaluation method, which may biased the findings due to the imperfect gold standard tests (Toft et al., 2005). The Se and Sp of the RID and TIR spectroscopy for diagnosis of calves with FTPI have been shown to be consistent (RID: Se estimates centered around 95% and Sp estimates centered around 93%, while, TIR: Se estimates centered around 76% and Sp estimates centered around 96%), regardless of the digital Brix refractometer cut-off values. However, the $Se_{\text{Brix}}$ increased from 79% to 95% and $Sp_{\text{Brix}}$ decreased from 95% to 84% with increasing the cut-off value from 8.2 to 8.5%, which is in accordance with previous studies (Deelen, et al., 2014; Elsohaby, et al., 2015; Hernandez, et al., 2016). In those studies, estimation of digital Brix refractometer diagnostic characteristics was based on the assumption of an existing perfect reference standard. It was notable that the median values of the Sp estimate of TIR were always higher than Sp estimates of RID and Brix at all tested cut-offs. However, the 95% confidence intervals estimate of the RID, TIR and Brix tests was overlapping across the tested cut-offs of digital Brix refractometer reflecting the inability to prefer a test over the other based on the Sp estimate.

The uncertainty associated with the Se estimates of those assays is a reflection of the varying number of truly infected herds for each assay at different cut-off values that are used in the tests’ Se estimation. At the highest cut-off value of the digital Brix refractometer (8.5%), the RID and TIR spectroscopy have the largest uncertainty around its Se estimate, whereas digital Brix refractometer had the smallest. At the lowest cut-off (8.2%), the reverse was true. The uncertainty associated with the changes in Se estimates of those tests at different cut-offs was probably attributed to the changing of target condition, number of true positive and negative values at those cut-offs for the different tests, and to the greater difference in the prevalences amongst the populations studied (Mweu et al., 2012).

Model assumptions
Test characteristics (Se and Sp) for RID, TIR spectroscopy and digital Brix refractometer were estimated at four different cut-offs using Bayesian LCA. The LCA analysis was based on a modified version of the LCA introduced by Hui and Walter (1980) for evaluation of diagnostic tests in the absence of a gold standard. We evaluated the Hui-Walter model assumptions and it has been fulfilled. Different prevalences in populations are fundamental to LCA models (Kostoulas et al., 2017). In this study, location of the study participants was regarded as a variable for study population stratifier to classify them into four provinces. The apparent prevalence of FTPI in dairy calves from the different provinces (PE, NB, NS, and NL) showed a wide variation. Subsequently, it was assumed that the test characteristics could be considered constant across the populations to fulfill the first model assumption. For RID and TIR spectroscopy, serum samples with IgG concentration greater than 1000 mg/dl was considered “negative” and those with IgG concentration less than 1000 mg/dl was considered “positive” (Godden, 2008; Weaver et al., 2000; Elsohaby et al., 2014).

As for the second model assumption, the RID and TIR were considered conditionally dependent (COC) given the same disease status because both of them are direct test measuring the same target (IgG) in the serum samples while, digital Brix refractometer measuring the IgG in the serum samples indirectly through the index score of refraction. To confirm our assumption about COC between RID and TIR, therefore, we ran different models such as CID between the three tests and COC models (digital Brix refractometer and RID) and (digital Brix refractometer and TIR). Based on the DIC, we found the COC (RID and TIR) model is giving the smallest DIC value hence; it has been considered the best fitting model and was presented in this study. The third model assumption was achieved by repeating the model analysis with exclusion of each of the subpopulation, one at a time. The test estimate results showed unsubstantial changes, which supports that the assumption was not violated.

**Conclusions**
We have estimated the Se and Sp of RID, TIR spectroscopy and digital Brix refractometer for
diagnosis of FTPI without the assumption of an existing reference standard. RID has a higher Se than
TIR and Brix, if the latter is used with cut-offs of 8.2% or 8.3%. However, the higher the cut-off, the
more comparable sensitivities of RID and digital Brix refractometer. The median estimate of Sp\textsubscript{TIR}
were always higher than Sp\textsubscript{RID} and Sp\textsubscript{Brix} at all tested cut-offs. However, the 95% confidence
intervals estimates of the three tests were overlapping across the tested cut-offs of digital Brix
refractometer reflecting the inability to prefer a test over the other based on the Sp estimate.

Acknowledgments

The authors thank summer program students, participating dairy farmers and veterinarians, as
well as Theresa Andrews, Cynthia Mitchell, and Natasha Robinson (Maritime Quality Milk
Laboratory, Charlottetown, PEI, Canada) for their technical assistance and data collection. Authors
also would like to thank Dr. Emilie Laurin (Atlantic Veterinary Collage) for review the manuscript.
This research was funded by the Atlantic Canada Opportunities Agency (Charlottetown, PEI,
Canada); Milk 2020 (Sussex, NB, Canada) and Zoetis (Kirkland, QC, Canada). I. Elsohaby is
supported by a Mitacs Elevate Postdoctoral Fellowship (IT09473).

Conflict of interests

None of the authors have financial or personal relationships with other people or organizations
that could inappropriately influence or bias the content of the paper.

References

immunoglobulin G concentrations using an automated turbidimetric immunoassay in dairy
calves. J. Dairy Sci. 95, 4596-4599.


**Table 1**: Cross-tabulated results for combinations of three diagnostic tests for diagnosis of failure of transfer of passive immunity (FTPI) in dairy calves (n = 691) from four different provinces in Atlantic Canada.

<table>
<thead>
<tr>
<th>Cut-offs of Brix refractometer</th>
<th>Population based on location</th>
<th>Tests combinations (RID $T_1$; TIR $T_2$; Brix $T_3$)</th>
<th>Total (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+++</td>
<td>++</td>
</tr>
<tr>
<td>Cut-off 8.2%</td>
<td>NB</td>
<td>35</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>NL</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>NS</td>
<td>51</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>PE</td>
<td>36</td>
<td>12</td>
</tr>
<tr>
<td>Total (n)</td>
<td></td>
<td>142</td>
<td>46</td>
</tr>
<tr>
<td>Cut-off 8.3%</td>
<td>NB</td>
<td>38</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>NL</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>NS</td>
<td>53</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>PE</td>
<td>40</td>
<td>8</td>
</tr>
<tr>
<td>Total (n)</td>
<td></td>
<td>153</td>
<td>35</td>
</tr>
<tr>
<td>Cut-off 8.4%</td>
<td>NB</td>
<td>41</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>NL</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>NS</td>
<td>58</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>PE</td>
<td>44</td>
<td>4</td>
</tr>
<tr>
<td>Total (n)</td>
<td></td>
<td>166</td>
<td>22</td>
</tr>
<tr>
<td>Cut-off 8.5%</td>
<td>NB</td>
<td>41</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>NL</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>NS</td>
<td>61</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>PE</td>
<td>45</td>
<td>3</td>
</tr>
<tr>
<td>Total (n)</td>
<td></td>
<td>170</td>
<td>18</td>
</tr>
</tbody>
</table>

$^a$ $T_1$ = radial immunodiffusion (RID) assay; $T_2$ = transmission infrared (TIR) spectroscopy; $T_3$ = digital Brix refractometer.
Table 2: Posterior median and 95% posterior credibility interval (PCI) of tests sensitivity (Se) and specificity (Sp) estimates and true prevalence of FTPI at four cut-offs for digital Brix refractometer (8.2%, 8.3%, 8.4%, and 8.5%) in dairy calves (n = 691) from four different Canadian provinces.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Cut-offs for digital Brix refractometer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cut-off 8.2%</td>
</tr>
<tr>
<td></td>
<td>Median 95% PCI a</td>
</tr>
<tr>
<td>SeRID</td>
<td>0.96 0.88 0.99</td>
</tr>
<tr>
<td>SeTIR</td>
<td>0.79 0.70 0.85</td>
</tr>
<tr>
<td>SeBrix</td>
<td>0.79 0.70 0.96</td>
</tr>
<tr>
<td>SpRID</td>
<td>0.93 0.84 0.97</td>
</tr>
<tr>
<td>SpTIR</td>
<td>0.97 0.90 1.00</td>
</tr>
<tr>
<td>SpBrix</td>
<td>0.95 0.91 0.98</td>
</tr>
<tr>
<td>Prevalence in NS</td>
<td>0.29 0.21 0.37</td>
</tr>
<tr>
<td>Prevalence in NL</td>
<td>0.54 0.39 0.68</td>
</tr>
<tr>
<td>Prevalence in NB</td>
<td>0.38 0.27 0.47</td>
</tr>
<tr>
<td>Prevalence in PE</td>
<td>0.26 0.18 0.34</td>
</tr>
<tr>
<td>G_{SeRID_TIR}</td>
<td>0.01 -0.003 0.069</td>
</tr>
<tr>
<td>G_{SpRID_TIR}</td>
<td>0.02 0.0003 0.079</td>
</tr>
</tbody>
</table>

a 95% Posterior credibility interval (PCI).