Systematic Review and Meta-Analysis of Diagnostic Accuracy of Serum Refractometry and Brix Refractometry for the Diagnosis of Inadequate Transfer of Passive Immunity in Calves

S. Buczinski (iD), E. Gicquel, G. Fecteau, Y. Takwoingi, M. Chigerwe, and J.M. Vandeweerd

Background: Transfer of passive immunity in calves can be assessed by direct measurement of immunoglobulin G (IgG) by methods such as radial immunodiffusion (RID) or turbidimetric immunoassay (TIA). IgG can also be measured indirectly by methods such as serum refractometry (REF) or Brix refractometry (BRIX).

Objectives: To determine the accuracy of REF and BRIX for assessment of inadequate transfer of passive immunity (ITPI) in calves.

Design: Systematic review and meta-analysis of diagnostic accuracy studies.

Methods: Databases (PubMed and CAB Abstract, Searchable Proceedings of Animal Science) and Google Scholar were searched for relevant studies. Studies were eligible if the accuracy (sensitivity and specificity) of REF or BRIX was determined using direct measurement of IgG by RID or turbidimetry as the reference standard. The study population included calves <14 days old that were fed with natural colostrum (colostrum replacement products were excluded). Quality assessment was performed by the QUADAS-2 tool. Hierarchical models were used for meta-analysis.

Results: From 1,291 references identified, 13 studies of 3,788 calves were included. Of these, 11 studies evaluated REF and 5 studies evaluated BRIX. The median (range) prevalence of ITPI (defined as calves with IgG <10 g/L by RID or TIA) was 21% (1.3–56%). Risk of bias and applicability concerns were generally low or unclear. For REF, summary estimates were obtained for 2 different cutoffs: 5.2 g/dL (6 studies) and 5.5 g/dL (5 studies). For the 5.2 g/dL cutoff, the summary sensitivity (95% CI) and specificity (95% CI) were 76.1% (63.8–85.2%) and 89.3% (82.3–93.7%), and 88.2% (80.2–93.3%) and 77.9% (74.5–81.0%) for the 5.5 g/dL cutoff. Due to the low number of studies using the same cutoffs, summary estimates could not be obtained for BRIX.

Conclusions and Clinical Importance: Despite their widespread use on dairy farms, evidence about the optimal strategy for using refractometry, including the optimal cutoff, are sparse (especially for BRIX). When using REF to rule out ITPI in herds, the 5.5 g/dL cutoff may be used whereas for ruling in ITPI, the 5.2 g/dL cutoff may be used.

Key words: Accuracy; IgG; Refractometry; Sensitivity; Specificity.

The newborn calf is highly dependent on colostrum intake to acquire adequate passive immunity during the neonatal period. The quality of transfer of passive immunity is most often practically assessed using serum immunoglobulin G (IgG) concentration 1–6 days after birth.[1] Inadequate transfer of passive immunity (ITPI) is diagnosed when serum IgG concentration is below a particular threshold (cutoff). Several cutoffs have been reported by different authors such as 10 g/L,[2,3] 12 g/L,[4] 16 g/L,[5] and 27 g/L.[6] These different cutoffs were

From the Department of clinical sciences, Faculté de Médecine Vétérinaire, Université de Montréal, Saint-Hyacinthe, QC Canada (Buczinski, Gicquel, Fecteau); Institute of Applied Health Research, University of Birmingham, Birmingham, UK (Takwoingi); Department of Medicine and Epidemiology, University of California-Davis, Davis, CA (Chigerwe); Integrated Veterinary Research Unit, Department of Veterinary Medicine, Faculty of Sciences, Research Institute for Life Sciences, Université de Namur, Namur, Belgium (Vandeweerd).

Corresponding author: S. Buczinski, Department of Clinical Sciences, Faculté de Médecine Vétérinaire, Université de Montréalz, Saint-Hyacinthe, 3200, rue Sicotte, J2S 2M2, QC, Canada; e-mail: s.buczinski@umontreal.ca.

Abbreviations:

BRIX	Brix refractometry
FN	false negative
FP	false positive
HSROC	hierarchical summary receiver operating characteristic
IgG	immunoglobulin G
PI	prediction interval
QUADAS	QUAlity Assessment of Diagnostic Accuracy Studies
REF	serum refractometry
SROC	summary receiver operating characteristic
RID	radial immunodiffusion
TIA	turbidimetry immunoassay
TN	true negative
TP	true positive
UREM	univariate random effects logistic regression model

obtained from different study populations (beef and dairy breeds) with different clinical definitions of inadequate immune transfer and by a data-driven approach. Various negative outcomes have been associated with lower IgG concentration in calves.[7] Increased risk of mortality, overall neonatal morbidity, as well as diarrhea, and respiratory disease have been observed.[7] Stochastic model risk analysis estimated the total cost per calf with ITPI as €60 ($72 using 2017, September 11 currency exchange rate (95% prediction interval [PI]: €10–€109 [$12–131]) and €80 ($96); 95% PI €20–€139 ($24–167) for dairy and beef calves, respectively.[7] Therefore, to assess the quality of transfer of passive

immunity in calves on a farm, regular measurements of serum IgG concentrations on calves less than a week old is a recommended practice.[8]

Serum IgG concentration can be determined by several diagnostic tests such as the radial immunodiffusion assay (RID) and turbidimetric immunoassay (TIA). The RID is the gold standard method for assessing IgG,[1] but TIA can also be considered a reference technique.[9] Serum IgG <10 g/L has been traditionally used as an acceptable threshold for defining ITPI.[1,8,10] Both RID and TIA are reliable for measuring IgG concentrations but are not practical on farms, are expensive, and the results are typically available after 24–36 hours.[11] For these reasons, other practical methods to estimate serum IgG concentration have been investigated.[12,13] Among these methods, optical or digital refractometry (REF) and Brix refractometry (BRIX) have been suggested as methods to assess serum IgG concentration. Both methods use serum samples, and the results are numerical and continuous (in g/dL for REF also referred to as total solids or % for BRIX). One BRIX % is equivalent to the refractance of a 1% sugar water solution. Several cutoffs have been proposed for REF (5.0 g/dL, 5.2 g/dL, or 5.5 g/dL)[1,10] and BRIX (8.4% Brix degrees),[12] with values below a cutoff indicating a positive test result for ITPI. Using cutoffs is a practical way to dichotomize results and can be easily understood by farmers although categorization of a continuous marker has some limitations.[14]

The diagnostic accuracy of REF and BRIX has been estimated in various studies quantifying test sensitivity (proportion of calves with ITPI who are test positive) and specificity (proportion of calves without ITPI who are test negative). However, comparing the results of different studies without a formal, systematic approach is problematic. The evidence-based approach is increasingly used in veterinary medicine.[15] This approach involves identification, appraisal, and synthesis of relevant studies on a specific topic by a reproducible step-by-step approach; results are interpreted taking into account between-study variability, risk of bias (internal validity), and applicability (external validity) of the studies. Diagnostic test accuracy can also be assessed by an evidence-based approach to summarize available test accuracy data and to investigate potential sources of heterogeneity.[16,17] Recently, we used this approach to determine the accuracy of BRIX to assess IgG concentration of colostrum (defining a good quality colostrum when IgG ≥50 g/L by RID or near-infrared spectroscopy as a reference standard).[18] The diagnostic accuracy of REF and BRIX for assessing ITPI needs to be determined to provide evidence of the likely performance of these tests in practice.

Therefore, our objective was to perform a systematic review and meta-analysis of the 2 index tests REF and BRIX for diagnosing ITPI as defined by IgG quantification. The specific question we were interested in answering was the following: "What is the sensitivity and specificity of serum REF and BRIX for the diagnosis of ITPI using IgG concentration (determined by RID or TIA) as a reference standard test in calves <14 days of age?"

Materials and Methods

The protocol of this systematic review has been published.[19] Two minor changes concerning the study population and statistical analysis were made to the protocol. A systematic search of the literature was performed for studies published between 1986 and June 1st, 2016. We searched PubMed, CAB Abstract, and Searchable Proceedings of Animal Conferences (S-PAC) to identify relevant studies reporting the accuracy of at least one of the 2 index tests (REF or BRIX) versus the reference standard (serum IgG measurement by either RID or TIA). The search strategy was published in the protocol. The reference list of each selected article was further screened to identify other potentially relevant articles and gray literature. Another search was performed by Google Scholar with the strategy: "Brix refractometer failure of passive transfer in calves" to identify published studies not retrieved from the other databases. The Google Scholar search was stopped after 40 consecutive references were judged not to be related to the review question.

All references were imported into freely available software,[a] and duplicates were removed. Titles were then screened for their relevance to the review question by 2 authors (SB and EG). A second screening was performed by reading abstracts of the studies selected. We excluded studies that were not written in English, French, or Spanish (that could not be read by at least 1 author); did not evaluate the accuracy of REF and/or BRIX; the reference standard was not RID or TIA; or were review manuscripts.

The final selection of studies was completed after reading the manuscripts. Studies were included if they reported the accuracy of REF and/or BRIX against RID or TIA as reference standard, and if 2 by 2 tables of one of the index tests against IgG concentration could be retrieved from the manuscript. In the original protocol, we had planned to focus on calves ≤8 days old. However, although some studies mostly included calves ≤8 days old, they also included calves up to 10 or 13 days old. These studies were included in this systematic review. Data were collected independently by 2 review authors (SB and EG) and checked for consistency. The data extracted were authors' name, year of publication, study design, population of calves (dairy, beef, and mixed), age of calves sampled, reference standard used (RID or TIA), IgG cutoff used for defining ITPI, proportion of calves below reported cutoff, refractometer or Brix refractometer used, and sample storage before performing the index or reference standard test. We recorded study design as single-gate or 2-gate.[20] Briefly, a single-gate study design is where cases (calves with ITPI) and noncases (calves with adequate TPI) were sampled using a single set of eligibility criteria (i.e., 1–gate). In contrast, a 2-gate study is a study where cases and noncases were sampled using a different set of eligibility criteria (i.e., cases and noncases enter the study through separate gates). The 2-gate study design is at higher risk of spectrum bias and can overestimate test sensitivity and specificity.[20]

The risk of bias and applicability of the included studies were assessed by the QUality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool.[21] This tool assesses the internal validity of each study (i.e., risk of bias) as well as their external validity (i.e., applicability of the study with respect to the question of the review). The assessment of applicability differs from eligibility screening because the included studies are assessed in terms of how well the study population and setting, index test, and reference standard match the review question. Two review authors (SB and EG) independently assessed each study and disagreements were resolved by discussion with a third review author. We recorded the degree of agreement between the review authors.

Statistical analysis

Descriptive statistics were obtained for the main study characteristics including the number of calves, prevalence of ITPI (defined as the proportion of calves below the IgG cutoff assessed by RID or TIA [reference standard test]), and proportion of studies with specific characteristics. For each study, 2×2 tables of the number of true positives (TP: index test positive and ITPI present), true negatives (TN: index test negative and no ITPI), false positives (FP: index test positive but no ITPI), and false negatives (FN: index test negative but ITPI present) were obtained for all cutoffs reported in the study. Using these tables, sensitivity and specificity, and their 95% confidence intervals (CI), were calculated.

The hierarchical summary receiver operating characteristic (HSROC) model[22] was planned for meta-analysis of a pair of sensitivity and specificity from each study.[19] The HSROC model is one of the models recommended for diagnostic accuracy meta-analysis because it accounts for both within- and between-study variation.[23] In the HSROC model, the number of positive test results in the jth group and ith study follows binomial distributions with the probability of a positive test given by

$$\text{logit}\left(\pi_{ij}\right) = \left(\theta_i + \alpha_i X_{ij}\right)\exp\left(-\beta X_{ij}\right),$$

where π_{ij} is the proportion of test positives, and n_{ij} is the number in group j in the ith study. For the adequate TPI group, $j = 0$ and X_{ij} is coded as -0.5. For the ITPI group, $j = 1$ and X_{ij} is coded as 0.5. The implicit threshold θ_i models the trade-off between true and false-positive fractions, while α_i (accuracy parameter) measures the difference between the true and false-positive fractions. Both θ_i and α_i are modeled as random effects with independent normal distributions. The shape parameter, β, allows for asymmetry in the shape of the summary receiver operating characteristic (SROC) curve.

We chose the HSROC model so that we could estimate SROC curves because we expected studies to use different cutoffs. For this analysis, if a study reported more than 1 cutoff, we randomly
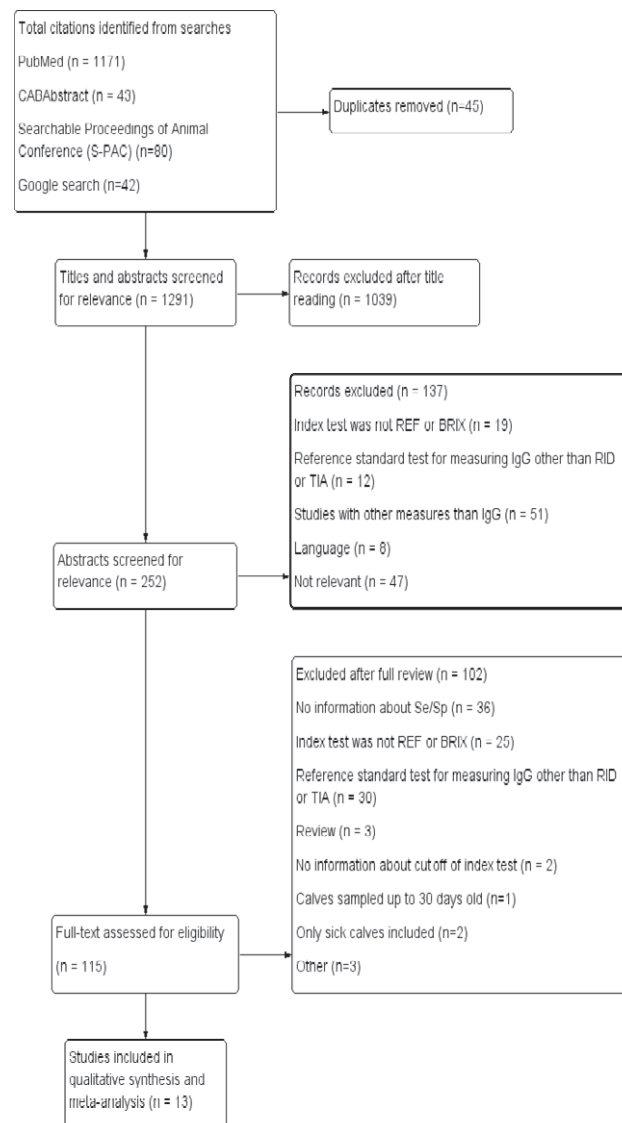

Fig 1. Flow of the study selection process. The search was performed on June 1, 2016. CAB, Commonwealth Agricultural Bureau; TP, total protein; REF, refractometer in g/L; BRIX, Brix refractometry (%); RID, radial immunodiffusion; TIA, turbidimetric immunoassay; IgG, immunoglobulin G; Se, sensitivity; Sp, specificity.

Table 1. Characteristics of studies reporting the accuracy of refractometer or Brix refractometry (BRIX) for the diagnosis of failure transfer of passive immunity in calves.

Study[a], Year	Population	Age of Calves (Days)	Reference Standard	IgG Cutoff (g/L)	Refractometer	Refractometer Type	Sample Storage for Index Test	Sample Storage for IgG Measurement
Optical or digital refractometry (REF)								
Calloway, 2002[b]	Beef and dairy calves	<10	RID	10	> Reichert Medical instrument, Buffalo, NY (1) TS Meter, Leica, Buffalo, NY [2]) Westover RHC-2000 handheld refractometer, Woodinville, WA (3)	Optical	Frozen	Frozen
Dawes, 2002	Beef and dairy calves	<14	RID	10	Leica TS meter refractometer, model 10400A, Leica, Buffalo, NY	Optical	Fresh	Frozen at −20°C
Deelen, 2014	Holstein calves	3–6	RID	10	PA202X-003-105, MISCO, Cleveland, OH	Digital	Frozen at −20°C	Frozen at −20°C
Elsohaby, 2015	Dairy calves in majority	1–11	RID	10	Westover RHC-200ATC handheld refractometer, Woodinville, WA	Optical	Frozen at −20°C	Frozen at −20°C
Gungor, 2004	Holstein-Friesian calves	1	RID	8	NA	NA	Frozen at −18°C	Frozen at −18°C
Hernandez, 2016[c]	Holstein calves	2–6	TIA	10	Kernco Instruments Company Clinical refractometer, NOW, Tokyo, Japan	Optical	Refrigerated at 4°C	Refrigerated at 4°C
Lee, 2008	Dairy calves	0–10	RID	10		NA	Frozen at −80 °C	Frozen at −80 °C
McVicker, 2002	Holstein male calves	4–8	TIA	10	TS Meter, American Optical, Buffalo, NY	Optical	Frozen at −20°C	Frozen at −20°C
Perino, 1993	4 crossbreeds beef calves	0–1	RID	8	NA	NA	NA	NA
Priestley, 2013	Holstein female calves	1–7	RID	10	NA	NA	Fresh	Frozen at −25°C
Tyler, 1996	Calves	1–8	RID	10	TS Meter; American Optical, Buffalo, NY	Optical	Refrigerated at 4°C	Refrigerated at 4°C
BRIX								
Chamorro, 2015	Dairy calves	1–5	RID	10	NA	Digital	NA	NA
Deelen, 2014	Holstein dairy calves	3–6	RID	10	PA202X-003-105, MISCO, Cleveland, OH	Digital	Frozen at −20°C	Frozen at −20°C
Elsohaby, 2015	Dairy calves in majority	1–11	RID	10	PAL-1 digital Brix refractometer, Atago Co Ltd; Bellevue, WA	Digital	Frozen at −20°C	Frozen at −20°C
Hernandez, 2016[c]	Holstein calves	2–6	TIA	10	Palm Abbe PA201, MISCO, 2013	Digital	Refrigerated at 4°C	Refrigerated at 4°C
Morill, 2013	Holstein calves	1	RID	10	Model 300034; Sper Scientific, Scottsdale, AZ	Digital	Frozen at −20°C	Frozen at −20°C

RID, radial immunodiffusion assay; TIA, turbidimetric immunoassay; NA, not available.
[a]All studies were single-gate studies.
[b]This particular study used 3 different refractometers on the same samples.
[c]The accuracy measures came from the 1st trial reported in this study.

selected 1 cutoff so that only a pair of sensitivity and specificity was included from each study. Only studies that defined ITPI using a cutoff of 10 g/L for the reference standard were included in all meta-analyses. The HSROC model was fitted using the NLMIXED procedure in the SAS[b] software package.[24]

As several cutoffs have been recommended depending on the objective of maximizing sensitivity or specificity,[1,10] we also estimated summary estimates of sensitivity and specificity at these cutoffs (using 5.0–5.2 g/dL and 5.5 g/dL for REF, and 8.4% for BRIX). When analyses using the HSROC model failed to converge due to the small number of studies, because summary estimates were the focus of these analyses, we used univariate random effects logistic regression models (UREM) which are recommended when data are sparse.[25] This model is a simplification of the bivariate model by assuming the covariance is zero as follows:

$$\begin{pmatrix} \mu_{A_i} \\ \mu_{B_i} \end{pmatrix} \sim \left(N \begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix}, \Sigma_{AB} \right) \text{with } \Sigma_{AB} = \begin{pmatrix} \sigma_A^2 & 0 \\ 0 & \sigma_B^2 \end{pmatrix}.$$

The logit sensitivity (μ_{Ai}) and logit specificity (μ_{Bi}) of the ith study follow normal distributions with mean μ_A and variance σ_A^2 for the logit sensitivities, and mean μ_B and variance σ_B^2 for the logit specificities. A binomial likelihood is used for modeling within-study variability.

Heterogeneity was investigated by visual examination of forest plots and SROC plots. We planned to formally investigate heterogeneity by adding a covariate to the HSROC model (i.e., meta-regression) for each potential source of heterogeneity. Factors of interest included the type of refractometer used (digital versus optical), peer-reviewed versus nonpeer-reviewed studies, and low (<20%) versus high (≥20%) prevalence of ITPI.[8] We planned to perform the analyses only if there were at least 5 studies for each subgroup of a covariate. We did not assess publication bias. Although the Deeks' test for detecting funnel plot asymmetry was developed specifically for systematic reviews of diagnostic accuracy studies, the test has low power when there is heterogeneity as is typically observed in diagnostic accuracy reviews.[26,27]

We created a summary of findings table to illustrate the implications of our meta-analytic findings for a hypothetical population of 1,000 calves with expected prevalence of ITPI of 10% (low-risk group), 20% (moderate-risk group), and 50% (high-risk group). These values represent variation of ITPI in different clinical settings (e.g., herds with good, average, or poor performance in terms of ITPI) and were obtained from the included studies. Using these hypothetical populations, the summary sensitivity and specificity, and their lower and upper 95% confidence limits, we calculated the number of TP, FN, FP, and TN,[c] as well as positive and negative predictive values.

Results

The flow diagram summarizing the flow of studies through the selection process is shown in Figure 1. After combining search results from the different sources and removing duplicates, we identified a total of 1,291 publications. Of these, 115 full-text papers were assessed for eligibility. A total of 13 test accuracy studies (3,788 calves), comprising 11 studies (1,814 calves) of REF[9,12,13,28–34] and 5 studies (2,881 calves) of BRIX,[11,12,29,30,35] were included in this systematic review. The characteristics of the studies are presented in Table 1. All studies were single-gate studies. Three studies[12,29,30] evaluated REF and BRIX in the same calves (Fig S1).

The risk of bias and applicability concern of each study are presented in Figure 2. With the exception of the reference standard domain, most studies had an unclear risk of bias. Applicability concerns were mainly unclear in the patient selection domain, while all studies were of low concern in the index test domain. In 5 of 13 studies, there was perfect agreement of the QUADAS-2 assessments performed by the 2 review authors. Disagreements were often due to scoring studies as low or unclear risk of bias in the flow and timing domain.

The reference standard used was RID in 11 studies and TIA in 2 studies. Eleven studies used a cutoff of 10 g/L of serum IgG concentration to define ITPI. In 2 studies, the authors defined ITPI as IgG <8 g/L.[28,33] The median prevalence of ITPI was 21% and ranged from 1.3[30] to 56%.[9] For each study, estimates of sensitivity and specificity, and the cutoff(s) used, are shown for REF and BRIX in Figures 3, 4, respectively.

We obtained SROC curves for REF and BRIX tests as shown in Figures S2 and S3. For meta-analyses at specific cutoffs, analyses using the HSROC model failed to converge. Using UREM, summary sensitivities and

Fig 2. Risk of bias and applicability concerns for each included study.

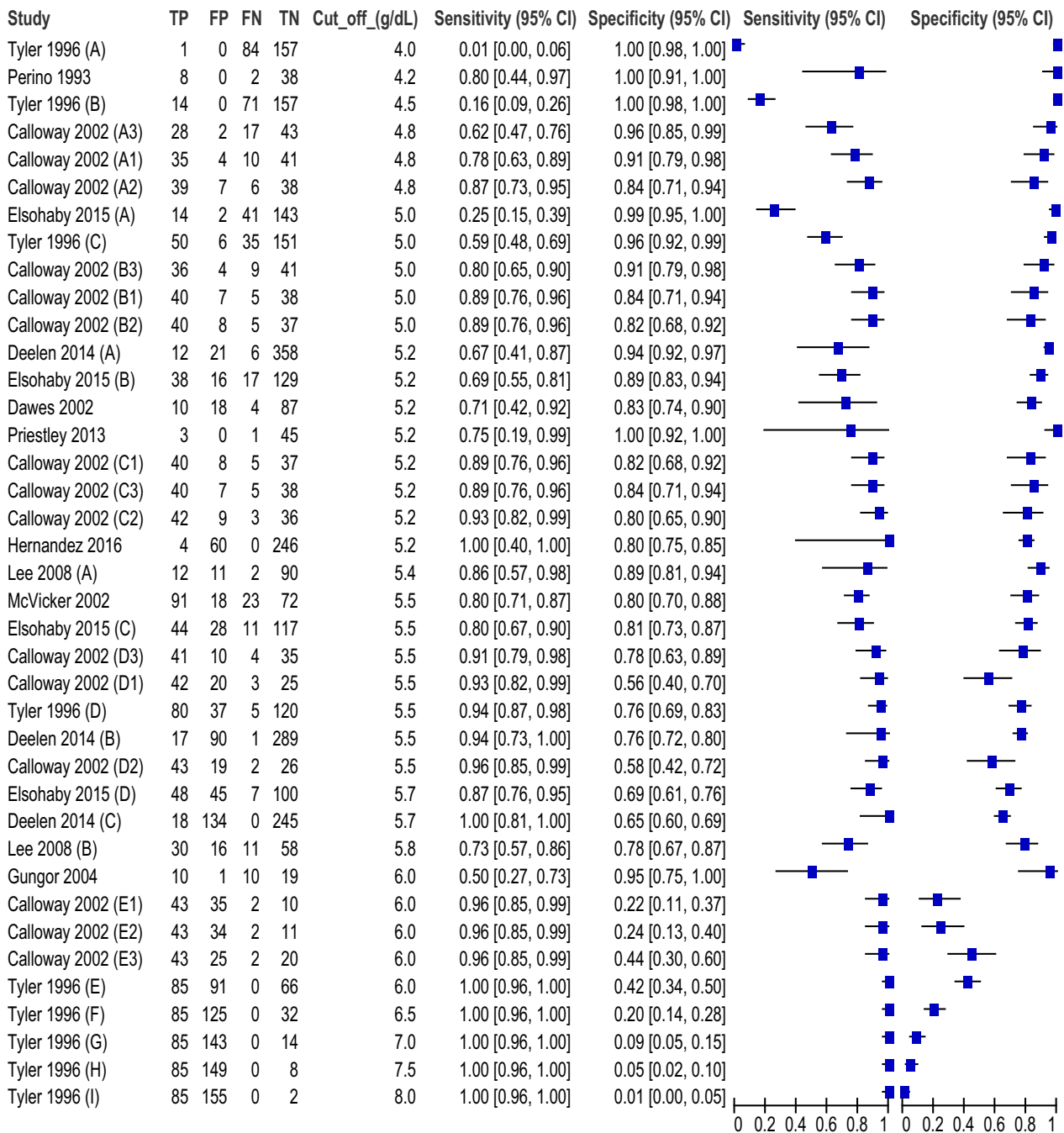| Study	TP	FP	FN	TN	Cut_off_(g/dL)	Sensitivity (95% CI)	Specificity (95% CI)
Tyler 1996 (A)	1	0	84	157	4.0	0.01 [0.00, 0.06]	1.00 [0.98, 1.00]
Perino 1993	8	0	2	38	4.2	0.80 [0.44, 0.97]	1.00 [0.91, 1.00]
Tyler 1996 (B)	14	0	71	157	4.5	0.16 [0.09, 0.26]	1.00 [0.98, 1.00]
Calloway 2002 (A3)	28	2	17	43	4.8	0.62 [0.47, 0.76]	0.96 [0.85, 0.99]
Calloway 2002 (A1)	35	4	10	41	4.8	0.78 [0.63, 0.89]	0.91 [0.79, 0.98]
Calloway 2002 (A2)	39	7	6	38	4.8	0.87 [0.73, 0.95]	0.84 [0.71, 0.94]
Elsohaby 2015 (A)	14	2	41	143	5.0	0.25 [0.15, 0.39]	0.99 [0.95, 1.00]
Tyler 1996 (C)	50	6	35	151	5.0	0.59 [0.48, 0.69]	0.96 [0.92, 0.99]
Calloway 2002 (B3)	36	4	9	41	5.0	0.80 [0.65, 0.90]	0.91 [0.79, 0.98]
Calloway 2002 (B1)	40	7	5	38	5.0	0.89 [0.76, 0.96]	0.84 [0.71, 0.94]
Calloway 2002 (B2)	40	8	5	37	5.0	0.89 [0.76, 0.96]	0.82 [0.68, 0.92]
Deelen 2014 (A)	12	21	6	358	5.2	0.67 [0.41, 0.87]	0.94 [0.92, 0.97]
Elsohaby 2015 (B)	38	16	17	129	5.2	0.69 [0.55, 0.81]	0.89 [0.83, 0.94]
Dawes 2002	10	18	4	87	5.2	0.71 [0.42, 0.92]	0.83 [0.74, 0.90]
Priestley 2013	3	0	1	45	5.2	0.75 [0.19, 0.99]	1.00 [0.92, 1.00]
Calloway 2002 (C1)	40	8	5	37	5.2	0.89 [0.76, 0.96]	0.82 [0.68, 0.92]
Calloway 2002 (C3)	40	7	5	38	5.2	0.89 [0.76, 0.96]	0.84 [0.71, 0.94]
Calloway 2002 (C2)	42	9	3	36	5.2	0.93 [0.82, 0.99]	0.80 [0.65, 0.90]
Hernandez 2016	4	60	0	246	5.2	1.00 [0.40, 1.00]	0.80 [0.75, 0.85]
Lee 2008 (A)	12	11	2	90	5.4	0.86 [0.57, 0.98]	0.89 [0.81, 0.94]
McVicker 2002	91	18	23	72	5.5	0.80 [0.71, 0.87]	0.80 [0.70, 0.88]
Elsohaby 2015 (C)	44	28	11	117	5.5	0.80 [0.67, 0.90]	0.81 [0.73, 0.87]
Calloway 2002 (D3)	41	10	4	35	5.5	0.91 [0.79, 0.98]	0.78 [0.63, 0.89]
Calloway 2002 (D1)	42	20	3	25	5.5	0.93 [0.82, 0.99]	0.56 [0.40, 0.70]
Tyler 1996 (D)	80	37	5	120	5.5	0.94 [0.87, 0.98]	0.76 [0.69, 0.83]
Deelen 2014 (B)	17	90	1	289	5.5	0.94 [0.73, 1.00]	0.76 [0.72, 0.80]
Calloway 2002 (D2)	43	19	2	26	5.5	0.96 [0.85, 0.99]	0.58 [0.42, 0.72]
Elsohaby 2015 (D)	48	45	7	100	5.7	0.87 [0.76, 0.95]	0.69 [0.61, 0.76]
Deelen 2014 (C)	18	134	0	245	5.7	1.00 [0.81, 1.00]	0.65 [0.60, 0.69]
Lee 2008 (B)	30	16	11	58	5.8	0.73 [0.57, 0.86]	0.78 [0.67, 0.87]
Gungor 2004	10	1	10	19	6.0	0.50 [0.27, 0.73]	0.95 [0.75, 1.00]
Calloway 2002 (E1)	43	35	2	10	6.0	0.96 [0.85, 0.99]	0.22 [0.11, 0.37]
Calloway 2002 (E2)	43	34	2	11	6.0	0.96 [0.85, 0.99]	0.24 [0.13, 0.40]
Calloway 2002 (E3)	43	25	2	20	6.0	0.96 [0.85, 0.99]	0.44 [0.30, 0.60]
Tyler 1996 (E)	85	91	0	66	6.0	1.00 [0.96, 1.00]	0.42 [0.34, 0.50]
Tyler 1996 (F)	85	125	0	32	6.5	1.00 [0.96, 1.00]	0.20 [0.14, 0.28]
Tyler 1996 (G)	85	143	0	14	7.0	1.00 [0.96, 1.00]	0.09 [0.05, 0.15]
Tyler 1996 (H)	85	149	0	8	7.5	1.00 [0.96, 1.00]	0.05 [0.02, 0.10]
Tyler 1996 (I)	85	155	0	2	8.0	1.00 [0.96, 1.00]	0.01 [0.00, 0.05]

Fig 3. Forest plot of refractometry for diagnosing inadequate transfer of passive immunity in calves. TP, true positives; FP, false positives; FN, false negatives; TN, true negatives; CI, confidence interval. The studies are ordered by ascending cutoff (in g/dL), sensitivity, and study name. Study names with suffixes A to I reported test accuracy at multiple cutoffs. The study by Calloway et al., 2002, used 3 different refractometers, identified by suffix numbers 1 to 3 (1 = Reichert Medical instrument, Buffalo, NY, 2 = TS Meter, Leica, Buffalo, NY, and 3 = Westover RHC-2000 handheld refractometer, Woodinville, WA). The information from the 2 × 2 tables Calloway (C3) and Calloway (D3) was used for summary accuracy assessment using 5.2 g/dL and 5.5 g/dL threshold, respectively (these tables were selected randomly among the 3 different refractometers datasets). With the exception of Perino 1993 and Güngör 2004, the studies used a cutoff of 10 g/L serum IgG concentration to define inadequate transfer of passive immunity. The 2 studies used a cutoff of 8 g/L and were excluded from all meta-analyses.

specificities were obtained for REF cutoffs of 5.2 g/dL (6 studies) and 5.5 g/dL (5 studies). For the 5.2 g/dL cutoff, the sensitivities from the 6 studies (140 ITPI cases out of 1,165 calves) ranged between 67 and 100%, and the specificities ranged between 83 and 100%; the summary sensitivity (95% CI) and specificity (95% CI) were 76.1% (63.8–85.2%) and 89.3% (82.3–93.7%). For the 5.5 g/dL cutoff, the sensitivities from the 5 studies
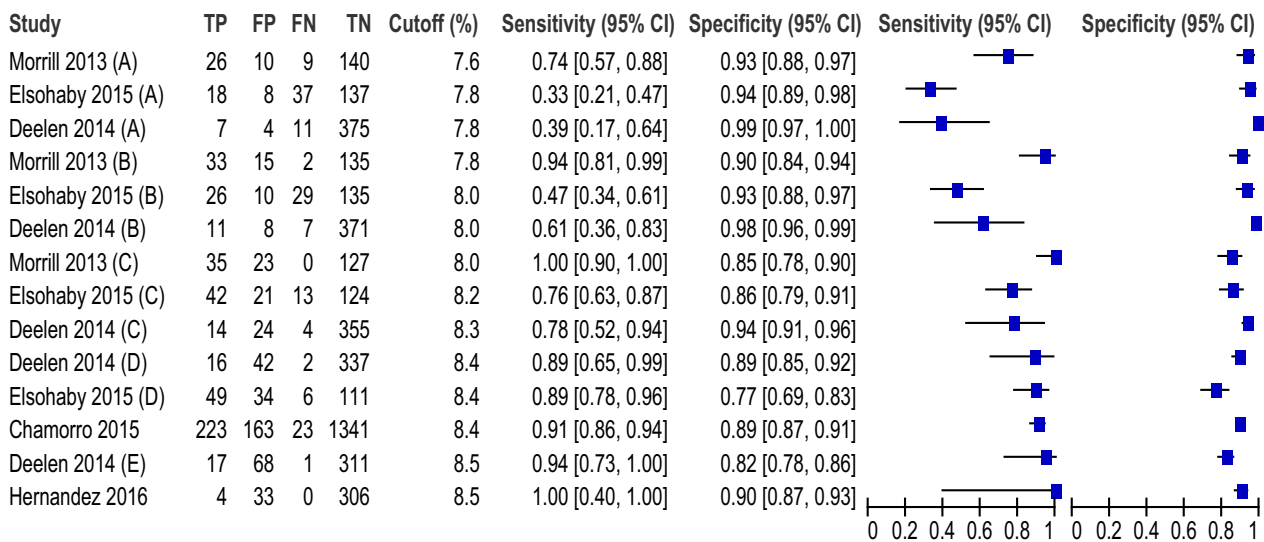
Study	TP	FP	FN	TN	Cutoff (%)	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Morrill 2013 (A)	26	10	9	140	7.6	0.74 [0.57, 0.88]	0.93 [0.88, 0.97]		
Elsohaby 2015 (A)	18	8	37	137	7.8	0.33 [0.21, 0.47]	0.94 [0.89, 0.98]		
Deelen 2014 (A)	7	4	11	375	7.8	0.39 [0.17, 0.64]	0.99 [0.97, 1.00]		
Morrill 2013 (B)	33	15	2	135	7.8	0.94 [0.81, 0.99]	0.90 [0.84, 0.94]		
Elsohaby 2015 (B)	26	10	29	135	8.0	0.47 [0.34, 0.61]	0.93 [0.88, 0.97]		
Deelen 2014 (B)	11	8	7	371	8.0	0.61 [0.36, 0.83]	0.98 [0.96, 0.99]		
Morrill 2013 (C)	35	23	0	127	8.0	1.00 [0.90, 1.00]	0.85 [0.78, 0.90]		
Elsohaby 2015 (C)	42	21	13	124	8.2	0.76 [0.63, 0.87]	0.86 [0.79, 0.91]		
Deelen 2014 (C)	14	24	4	355	8.3	0.78 [0.52, 0.94]	0.94 [0.91, 0.96]		
Deelen 2014 (D)	16	42	2	337	8.4	0.89 [0.65, 0.99]	0.89 [0.85, 0.92]		
Elsohaby 2015 (D)	49	34	6	111	8.4	0.89 [0.78, 0.96]	0.77 [0.69, 0.83]		
Chamorro 2015	223	163	23	1341	8.4	0.91 [0.86, 0.94]	0.89 [0.87, 0.91]		
Deelen 2014 (E)	17	68	1	311	8.5	0.94 [0.73, 1.00]	0.82 [0.78, 0.86]		
Hernandez 2016	4	33	0	306	8.5	1.00 [0.40, 1.00]	0.90 [0.87, 0.93]		

0 0.2 0.4 0.6 0.8 1   0 0.2 0.4 0.6 0.8 1

Fig 4. Forest plot of Brix refractometry for diagnosing inadequate transfer of passive immunity in calves. TP, true positives; FP, false positives; FN, false negatives; TN, true negatives; CI, confidence interval. The studies are ordered by ascending cutoff (in % of Brix), sensitivity, and study name. Study names with suffixes A to E reported test accuracy at multiple cutoffs.

(317 ITPI cases out of 1,133 calves) ranged between 80 and 94%, and the specificities ranged between 76 and 81%; the summary sensitivity (95% CI) and specificity (95% CI) were 88.2% (80.2–93.3%) and 77.9% (74.5–81.0%). Due to a limited number of studies with the same cutoff, we were unable to determine summary estimates for BRIX.

The summary of findings is presented in Table 2. The impact of these different REF cutoffs is obvious for high prevalence of ITPI (50%) with a positive predictive value of 75.6% (378/500) for the 5.2 g/dL cutoff or 87.8% (439/500) for the 5.5 g/dL cutoff.

Discussion

There was variability in diagnostic accuracy between studies. The performance of REF depends on the choice of cutoff; REF was more specific but less sensitive at the 5.2 g/dL cutoff compared to the 5.5 g/dL cutoff. This implies that the lower cutoff may be used when ruling in ITPI (e.g., trying to avoid a calf that would be falsely diagnosed as having ITPI). In contrast, the higher cutoff may be used when ruling out ITPI (lower false-negative rate). It was not possible to obtain summary estimates for BRIX due to limited data. Despite the common use of these diagnostic tests (especially REF) in clinics, data on their accuracy were limited as indicated by the low number of studies that were identified.

Systematic reviews of diagnostic test accuracy are seldom performed in veterinary medicine. However, they are important to understand the clinical performance of a diagnostic test.[36] To the authors' knowledge, this is the first systematic review on the diagnostic accuracy of REF and BRIX for diagnosing ITPI in calves. The review was performed using methods recommended in the Cochrane Handbook for Diagnostic Test Accuracy Studies (http://methods.cochrane.org/sdt/dta-review-

author-training assessed October 4, 2016).[37] Therefore, the key strength of this review is that it represents a methodologically robust overview of the currently available evidence on the accuracy of REF and BRIX for assessing ITPI.

The reporting quality of the available studies hampered the assessment of the risk of bias. For most of the studies, risk of bias in the index test domain was unclear as it was not reported whether REF or BRIX was interpreted without knowledge of IgG measurement results. In addition, the choice of the optimal cutoff used to define sensitivity and specificity was frequently data-driven which is known to inflate test accuracy.[38] We graphically explored heterogeneity in test accuracy, but due to limited data, we were unable to formally investigate potential sources of heterogeneity such as age and breed of calves, technical characteristics of the index tests (e.g., digital versus optical refractometry), and sample conservation method.

The definition of the reference standard to diagnose ITPI was based on a specific cutoff of serum IgG concentration determined by RID or TIA in healthy calves.[8] In all but 2 studies included in this systematic review, the serum IgG concentration cutoff used was 10 g/L although we did not limit our search to this specific cutoff. A 10 g/L cutoff is generally accepted as an industry benchmark, but the association between IgG levels and ITPI risk is complex. To the authors' knowledge, there is no serum IgG cutoff that would be 100% sensitive and specific for ITPI definition due to the dichotomization of a biologic process most likely better represented on a continuous scale. Most likely, the higher the serum IgG concentration, the better is the transfer of passive immunity. Dichotomizing the serum IgG values means that we suppose that a calf just below the used cutoff is different from the calf barely above the cutoff which we know is not true.

Table 2. Summary of findings table for refractometry (REF) to assess transfer of passive immunity in calves using immunoglobulin G concentration of 10 g/L by either radial immunodiffusion or turbidimetric immunoassay as reference standard.

REF Cutoff	Sensitivity % (95% CI)	Specificity % (95% CI)	Prevalence of FPT (%)	True Positives (TP) (Range)	False Negatives (FN) (Range)	False Positives (FP) (Range)	True Negatives (TN) (Range)	PPV (Range %)	NPV (Range %)
<5.2 g/dL	76.1 (63.8–85.2)	89.3 (82.3–93.7)	10	76 (64–85)	24 (15–36)	99 (57–159)	801 (741–843)	43.4 (34.8–52.9)	97.1 (95.9–98.0)
			20	151 (128–170)	49 (30–72)	88 (50–142)	712 (658–750)	63.2 (54.5–71.9)	93.6 (91.2–95.6)
			50	378 (319–426)	122 (74–181)	55 (32–89)	445 (412–479)	87.3 (82.7–90.9)	78.5 (72.6–84.8)
<5.5 g/dL	88.2 (80.2–93.3)	77.9 (74.5–81.0)	10	88 (80–93)	12 (7–20)	199 (171–230)	701 (671–729)	30.7 (28.8–31.9)	98.3 (97.3–99.0)
			20	176 (160–187)	24 (13–40)	177 (152–204)	623 (596–648)	49.9 (47.8–51.3)	96.3 (94.2–97.9)
			50	439 (401–467)	61 (34–99)	111 (95–128)	389 (373–405)	79.8 (78.5–80.8)	86.4 (80.4–91.6)

Three different prevalence scenarios of inadequate transfer of passive immunity (ITPI, defined as serum IgG <10 g/L) were used based on the distribution of prevalence of ITPI in included studies. The summary sensitivity and specificity for each cutoff were used for obtaining the number of TP, FN, FP, and TN in a hypothetical population of 1,000 calves. The positive (PPV) and negative (NPV) predictive values were computed using these numbers. The ranges around all the numbers were obtained using the lower and upper 95% confidence limits of the summary sensitivities and specificities.

Recent studies have shown that serum IgG >20 g/L and ≤25 g/L could be suggested as potential cutoffs above which heifer calves were at lower risk of dying during the first 4 months of life.[39] The optimal concentration of IgG to protect a calf may also depend on the specific risk factors which are variable from one farm to another. The same concentration of IgG in a calf may have different health impact if the calf is in a well-managed farm with low infectious challenges versus a farm with high known morbidity/mortality.

Very high levels of IgG could also be associated with ongoing illness (detected or not) and dehydration. Serum refractance can be affected by dehydration which can concentrate blood components and actual protein content. A sick calf could have an ongoing inflammatory process (e.g., septicemia, diarrhea, or respiratory disease) and be falsely classified as having adequate transfer of passive immunity due to increased serum refractance (increased blood inflammatory marker). In this case, the risk of misclassification concerning passive transfer of immunity is increased.[40] This is the reason for excluding studies exclusively describing sick animals (especially calves with scours) from our review.

The benefits of improving and monitoring transfer of passive immunity are evident because it is an important driver of calf health and morbidity in beef and dairy production,[7] as well as in veal production.[41] The REF and Brix refractometers are convenient for daily use. In addition, the BRIX scale can be used for other purposes such as colostrum quality assessment[18] or for assessing milk replacer total solids concentration.[42] The IgG levels are not the only colostrum component that could benefit calf long-term health and productivity. Cellular or humoral factors contained in the colostrum are also of importance to calf health.[43–45] However, unlike serum refractance which is a proxy of IgG concentration, the other constituents are not easy to measure. This association between serum IgG and serum refractance can also be affected by the type of colostrum fed. Colostral replacer or supplement has different protein profiles and therefore would give different serum refractance for treated calves than calves treated with maternal colostrum.[34,46] For this reason, we excluded studies where a colostrum replacement or supplement product was used. In these situations, the accuracy of REF or BRIX estimated in this review is unlikely to be applicable. In these cases, it would be interesting to examine the specific relationship between serum refractance and IgG concentration in order to determine the optimal way to diagnose ITPI.

Although there was paucity of data, we found more studies on REF than on BRIX for assessing ITPI. No summary estimates were obtained for any BRIX cutoff. Therefore, further research on the diagnostic accuracy of REF and BRIX is needed. Future studies of REF and BRIX should be reported using the STAndards for Reporting Diagnostic accuracy studies (STARD) guidelines to facilitate appraisal of their methodological quality.[47] Better reporting will also ensure availability of test accuracy data and information on how reported cutoffs were selected. This review identified only 3

studies that evaluated both REF and BRIX. Therefore, there was insufficient evidence to determine whether one refractance scale is more accurate than the other. Future research should address the comparative accuracy of the 2 tests to determine which test is more accurate, and therefore if one type of refractometer scale should be preferred when assessing ITPI.

In conclusion, there was a paucity of data on the accuracy of REF and BRIX when evaluated against RID or TIA as a reference standard, and study quality was often unclear. The main objective of ITPI investigation at the herd level is to identify cases of ITPI and/or improve colostrum management strategies. For minimizing the number of false-negative cases (i.e., the number of calves with ITPI not detected by the test), a cutoff of 5.5 g/dL instead of 5.2 g/dL for REF appears to be a better threshold for ruling out ITPI.

Footnotes

Acknowledgments

References

1. Weaver DM, Tyler JW, VanMetre DC, et al. Passive transfer of colostral immunoglobulins in calves. J Vet Intern Med 2000;14:569–577.

2. Besser TE, Gay CC, Pritchett L. Comparison of three methods of feeding colostrum to dairy calves. J Am Vet Med Assoc 1991;198:419–422.

3. Furman-Fratczak K, Rzasa A, Stefaniak T. The influence of colostral immunoglobulin concentration in heifer calves' serum on their health and growth. J Dairy Sci 2011;94:5536–5543.

4. Robison JD, Stott GH, DeNise SK. Effects of passive immunity on growth and survival in the dairy heifer. J Dairy Sci 1988;71:1283–1287.

5. Wittum TE, Perino LJ. Passive immune status at postpartum hour 24 and long-term health and performance of calves. Am J Vet Res 1995;56:1149–1154.

6. Dewell RD, Hungerford LL, Keen JE, et al. Association of neonatal serum immunoglobulin G1 concentration with health and performance in beef calves. J Am Vet Med Assoc 2006;228:914–921.

7. Raboisson D, Trillat P, Cahuzac C. Failure of passive immune transfer in calves: A meta-analysis on the consequences and assessment of the economic impact. PLoS ONE 2016;11: e0150452.

8. Godden S. Colostrum management for dairy calves. Vet Clin North Am Food Anim Pract 2008;24:19–39.

9. McVicker JK, Rouse GC, Fowler MA, et al. Evaluation of a lateral-flow immunoassay for use in monitoring passive transfer of immunoglobulins in calves. Am J Vet Res 2002;63:247–250.

10. McGuirk SM, Collins M. Managing the production, storage, and delivery of colostrum. Vet Clin North Am Food Anim Pract 2004;20:593–603.

11. Morrill KM, Polo J, Lago A, et al. Estimate of serum immunoglobulin G concentration using refractometry with or without caprylic acid fractionation. J Dairy Sci 2013;96:4535–4541.

12. Deelen SM, Ollivett TL, Haines DM, et al. Evaluation of a Brix refractometer to estimate serum immunoglobulin G concentration in neonatal dairy calves. J Dairy Sci 2014;97:3838–3844.

13. Tyler JW, Hancock DD, Parish SM, et al. Evaluation of 3 assays for failure of passive transfer in calves. J Vet Intern Med 1996;10:304–307.

14. Collins GS, Ogundimu EO, Cook JA, et al. Quantifying the impact of different approaches for handling continuous predictors on the performance of a prognostic model. Stat Med 2016;35:4124–4135.

15. Vandeweerd JM, Gustin P, Buczinski S. Evidence-based practice? An evolution is necessary for bovine practitioners, teachers, and researchers. Vet Clin North Am Food Anim Pract 2012;28:133–139.

16. O'Connor A, Evans RB. Critically appraising studies reporting assessing diagnostic tests. Vet Clin North Am Small Anim Pract 2007;37:487–497.

17. Leeflang MM, Deeks JJ, Takwoingi Y, et al. Cochrane diagnostic test accuracy reviews. Syst Rev 2013;2:82.

18. Buczinski S, Vandeweerd JM. Diagnostic accuracy of refractometry for assessing bovine colostrum quality: A systematic review and meta-analysis. J Dairy Sci 2016;99:7381–7394.

19. Buczinski S, Fecteau G, Chigerwe M, et al. Diagnostic accuracy of refractometer and Brix refractometer to assess failure of passive transfer in calves: Protocol for a systematic review and meta-analysis. Anim Health Res Rev 2016;17:3–8.

20. Rutjes AW, Reitsma JB, Vandenbroucke JP, et al. Case–control and two-gate designs in diagnostic accuracy studies. Clin Chem 2005;51:1335–1341.

21. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. Annals Intern Med 2011;155:529–536.

22. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med 2001;20:2865–2884.

23. Macaskill P, Gatsonis C, Deeks J, et al. Chapter 10: Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C, ed. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. London: The Cochrane Collaboration, 2010:1–16.

24. Takwoingi Y, Deeks J. METADAS: A SAS macro for meta-analysis of diagnostic accuracy studies. User guide version 1.3.2010 July. Available from : http://srdta.cochrane.org

25. Takwoingi Y, Guo B, Riley RD, et al. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. Stat Meth Med Res 2017;26:1896–1911.

26. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol 2005;58:882–893.

27. van Enst WA, Ochodo E, Scholten RJ, et al. Investigation of publication bias in meta-analyses of diagnostic test accuracy: A meta-epidemiological study. BMC Med Res Methodol 2014;14:70.

28. Perino L, Sutherland R, Woollen N. Serum gamma-glutamyltransferase activity and protein concentration at birth and after suckling in calves with adequate and inadequate passive transfer of immunoglobulin G. Am J Vet Med 1993;54:56–59.

29. Elsohaby I, McClure JT, Keefe GP. Evaluation of digital and optical refractometers for assessing failure of transfer of passive immunity in dairy calves. J Vet Intern Med 2015;29:721–726.

30. Hernandez D, Nydam DV, Godden SM, et al. Brix refractometry in serum as a measure of failure of passive transfer compared to measured immunoglobulin G and total protein by refractometry in serum from dairy calves. Vet J 2016;211:82–87.

31. Calloway CD, Tyler JW, Tessman RK, et al. Comparison of refractometers and test endpoints in the measurement of serum protein concentration to assess passive transfer status in calves. J Am Vet Med Assoc 2002;221:1605–1608.

32. Dawes ME, Tyler JW, Hostetler D, et al. Evaluation of a commercially available immunoassay for assessing adequacy of passive transfer in calves. J Am Vet Med Assoc 2002;220:791–793.

33. Güngör Ö, Bastan A, Erbil M. The usefulness of the γ-glutamyltransferase activity and total proteinemia in serum for detection of the failure of immune passive transfer in neonatal calves. Rev Med Vet 2004;155:27–30.

34. Priestley D, Bittar JH, Ibarbia L, et al. Effect of feeding maternal colostrum or plasma-derived or colostrum-derived colostrum replacer on passive transfer of immunity, health, and performance of preweaning heifer calves. J Dairy Sci 2013;96:3247–3256.

35. Chamorro MF, Sargent R, Haines DM, et al. Evaluation of the Brix refractometer for measuring colostral and serum IgG concentrations. J Dairy Sci 2015;98:236.

36. Leeflang MM. Systematic reviews and meta-analyses of diagnostic test accuracy. Clin Microbiol Infect 2013;20:105–113.

37. Macaskill P, Gatsonis C, Deeks J, et al. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 09 0. London: The Cochrane Collaboration; 2010.

38. Leeflang MM, Moons KG, Reitsma JB, et al. Bias in sensitivity and specificity caused by data-driven selection of optimal cutoff values: Mechanisms, magnitude, and solutions. Clin Chem 2008;54:729–737.

39. Chigerwe M, Hagey JV, Aly SS. Determination of neonatal serum immunoglobulin G concentrations associated with mortality during the first 4 months of life in dairy heifer calves. J Dairy Res 2015;82:400–406.

40. Fecteau G, Arsenault J, Paré J, et al. Prediction of serum IgG concentration by indirect techniques with adjustment for age and clinical and laboratory covariates in critically ill newborn calves. Can J Vet Res 2013;77:89–94.

41. Pardon B, Alliet J, Boone R, et al. Prediction of respiratory disease and diarrhea in veal calves based on immunoglobulin levels and the serostatus for respiratory pathogens measured at arrival. Prev Vet Med 2015;120:169–176.

42. Floren HK, Sischo WM, Crudo C, et al. Technical note: Use of a digital and an optical Brix refractometer to estimate total solids in milk replacer solutions for calves. J Dairy Sci 2016;99:7517–7522.

43. Ontsouka EC, Albrecht C, Bruckmaier RM. Invited review: Growth-promoting effects of colostrum in calves based on interaction with intestinal cell surface receptors and receptor-like transporters. J Dairy Sci 2016;99:4111–4123.

44. Reber AJ, Donovan DC, Gabbard J, et al. Transfer of maternal colostral leukocytes promotes development of the neonatal immune system I. Effects on monocyte lineage cells. Vet Immunol Immunopathol 2008;123:186–196.

45. Reber AJ, Donovan DC, Gabbard J, et al. Transfer of maternal colostral leukocytes promotes development of the neonatal immune system Part II. Effects on neonatal lymphocytes. Vet Immunol Immunopathol 2008;123:305–313.

46. Quigley JD 3rd, Kost CJ, Wolfe TM. Absorption of protein and IgG in calves fed a colostrum supplement or replacer. J Dairy Sci 2002;85:1243–1248.

47. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. Radiology 2015;277:826–832.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1. Forest plot and summary receiver operating characteristic plot of studies reporting the accuracy of refractometry and Brix refractometry in the same calves for the diagnosis of inadequate transfer of passive immunity in calves.

Figure S2. Forest plot and hierarchical receiver operating characteristic (HSROC) curve of studies reporting the accuracy of refractometry for the diagnosis of inadequate transfer of passive immunity in calves.

Figure S3. Forest plot and hierarchical receiver operating characteristic curve of studies reporting the accuracy of Brix refractometry for the diagnosis of inadequate transfer of passive immunity in calves.