

## Article

# Predicting Positive ELISA Results in Dairy Herds with a Preferred Status in a Paratuberculosis Control Program

Maarten F. Weber<sup>1,2,\*</sup>, Marian Aalberts<sup>1</sup>, Thomas Dijkstra<sup>1</sup> and Ynte H. Schukken<sup>1,2,3</sup>

<sup>1</sup> Royal GD, P.O. Box 9, 7400 AA Deventer, The Netherlands; m.aalberts@gdanimalhealth.com (M.A.); t.dijkstra@gdanimalhealth.com (T.D.); y.schukken@gdanimalhealth.com (Y.H.S.)

<sup>2</sup> Department of Population Health Sciences, Faculty of Veterinary Medicine, Utrecht University, Yalelaan 7, 3584 CL Utrecht, The Netherlands

<sup>3</sup> Quantitative Veterinary Epidemiology, Department of Animal Sciences, Wageningen University, P.O. Box 338, 6700 AH Wageningen, The Netherlands

\* Correspondence: m.weber@gdanimalhealth.com

**Simple Summary:** In many paratuberculosis control programs, test-negative herds are assigned a preferred herd status. This may induce the false belief that these herds are free of paratuberculosis. Hence, farmers may refrain from control measures that could prevent the spread of any undetected infections. The aim of the present study was to develop a predictive model to alert farmers with test-negative herds and a preferred status in the Dutch paratuberculosis control program if they are at an increased risk of positive ELISA results in a subsequent 30-month period. On the basis of the results of this study, we conclude that discrimination of herds with high (52%) and low (17%) risks of positive ELISA results is feasible. This might help farmers with the highest risks of future positive results to make better informed decisions regarding the need to take additional control measures to prevent the spread of any undetected Map infections.



**Citation:** Weber, M.F.; Aalberts, M.; Dijkstra, T.; Schukken, Y.H. Predicting Positive ELISA Results in Dairy Herds with a Preferred Status in a Paratuberculosis Control Program. *Animals* **2022**, *12*, 384. <https://doi.org/10.3390/ani12030384>

Academic Editor: Susanne Eisenberg

Received: 29 December 2021

Accepted: 26 January 2022

Published: 4 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Dairy herds participating in the Dutch milk quality assurance program for paratuberculosis are assigned a herd status on the basis of herd examinations by ELISA of individual serum or milk samples, followed by an optional confirmatory fecal PCR. Test-negative herds are assigned Status A; the surveillance of these herds consists of biennial herd examinations. Farmers falsely believing that their Status A herds are Map-free may inadvertently refrain from preventive measures. Therefore, we aimed to develop a predictive model to alert Status A farmers at increased risk of future positive ELISA results. Using data of 8566 dairy herds with Status A in January 2016, two logistic regression models were built, with the probabilities of  $\geq 1$  or  $\geq 2$  positive samples from January 2017–June 2019 as dependent variables, and province, soil type, herd size, proportion of cattle born elsewhere, time since previous positive ELISA results, and the 95th percentile of the S/P ratios in 2015–2016, as explanatory variables. As internal validation, both models were applied to predict positive ELISA results from January 2019–June 2021, in 8026 herds with Status A in January 2019. The model predicting  $\geq 1$  positive sample had an area under the receiver operating characteristics curve of 0.76 (95% CI: 0.75, 0.77). At a cut-off predicted probability  $\pi_c = 0.40$ , 25% of Status A herds would be alerted with positive and negative predictive values of 0.52 and 0.83, respectively. The model predicting  $\geq 2$  positive samples had lower positive, but higher negative, predictive values. This study indicates that discrimination of Status A herds with high and low risks of future positive ELISA results is feasible. This might stimulate farmers with the highest risks to take additional measures to control any undetected Map infections.

**Keywords:** paratuberculosis; dairy cattle; control program; predictive model

## 1. Introduction

Paratuberculosis is an inflammatory bowel disease, primarily affecting ruminants, that is caused by *Mycobacterium avium* subsp. *paratuberculosis* (MAP). Paratuberculosis

is of concern to dairy cattle industries worldwide because of its economic impact on dairy production [1,2] and uncertainty about the potential involvement of MAP in human disease [3,4]. Control programs for paratuberculosis have been developed in at least 22 countries [5]. In several programs for paratuberculosis, test-negative herds or herds with a high probability of absence of the infection are assigned an assurance score or preferred herd status [5,6].

In 2006, a milk quality assurance program (MQAP) for paratuberculosis in Dutch dairy herds was initiated. The aim of this MQAP is to reduce the concentration of Map in milk delivered to the milk processing industry by controlling Map in the participating herds [6–8]. Herds participating in the MQAP are assigned a herd status (A, B, or C) on the basis of the results of herd examinations [9]. In short, an initial assessment of the participating herds consists of a single herd examination. Test-negative herds enter a surveillance procedure and are assigned Status A. The surveillance of herds with Status A consists of biennial herd examinations. Test-positive herds at the initial assessment or during the surveillance procedure, enter a control procedure and are assigned Status B (if all test-positive cattle have been removed from the herd) or Status C (if any test-positive cattle are retained in the herd). If an annual herd examination in a herd with Status B yields negative results only, then the herd progresses to Status A. Each herd examination consists of either testing individual milk samples of all lactating cattle, or serum samples of all cattle  $\geq 3$  years of age, by an ELISA for antibodies against Map. The confirmation of positive ELISA results by individual fecal PCR assay or culture is optional.

The test scheme of the MQAP was developed to provide assurance that herds with Status A have a high probability of either the absence, or a low concentration, of Map in the milk ( $<10^3$  organisms per liter) delivered to the milk processors [7,8,10]. This does not necessarily mean that a Status A herd is Map-free. Thus, the program can be run at considerably lower costs than a program aiming to certify herds as Map-free [8,11]. However, undetected infections may contribute to the rather high rate of loss of Status A that has been observed [12]. Herds that were assigned Status A at the initial assessment had a probability of 43% of losing this status within ten years after the initial assessment. Herds that were assigned Status B or Status C at the initial assessment and progressed to Status A at a later stage, subsequently had a probability of 81% of losing Status A within eight years [12].

Preventive management measures may reduce the risk that undetected Map infections in Status A herds result in the loss of status A. However, many farmers with Status A herds indicate, in personal contacts, that they refrain from taking preventive management measures on the basis of their belief that their herds are Map-free. Therefore, the aim of the present study was to develop a predictive model to alert farmers with Status A herds if they are at an increased risk of positive ELISA results.

## 2. Materials and Methods

### 2.1. Laboratory Assays

For the participating herds, the results of serum samples and individual milk samples submitted from January 2006 onwards were retrieved from the laboratory information system of Royal GD. The individual milk samples and serum samples were tested by commercial ELISAs at varying cut-off values (Table 1). However, for the development and validation of the predictive model, a cut-off value of 1.0 was used, irrespective of the sample type (milk or serum), the ELISA kit (A or B), or the date of submission of the sample. The cut-off value in the present study was chosen in line with the cut-off of 1.0 used in the MQAP for milk samples tested with ELISA Kit A (Table 1), which constituted the majority of samples included in the training dataset (see below, Section 3.1). Thus, for the purpose of this study, all valid ELISA results with a sample-to-positive ratio (S/P)  $<1.0$  were considered negative, and all ELISA results with an S/P  $\geq 1.0$  were considered positive.

**Table 1.** ELISA kits and cut-off values used in the MQAP for individual milk and serum samples.

Sample	S/P	ELISA A <sup>1</sup>			ELISA B <sup>2</sup>
		Jan 2006–Aug 2007	Aug 2007–May 2018	May 2018–Dec 2019	Dec 2019–Present
Serum	≤0.90	Negative	Negative	Negative	Negative
	>0.9 and <1.1	Suspect	Suspect	Positive	
	≥1.1 and <1.2	Positive	Positive		
	≥1.2				Positive
Milk	<0.25	Negative	Negative	Negative	Negative
	≥0.25 and <1.0	Positive			
	≥1.0 and <1.1		Positive	Positive	
	≥1.1				Positive

<sup>1</sup> IDEXX Paratuberculosis Screening Ab Test; <sup>2</sup> ID Screen Paratuberculosis Indirect Screening Test (IDvet).

## 2.2. Development of the Predictive Model

To explore whether a high risk of loss of the preferred herd status (Status A) could be adequately predicted, a logistic regression analysis was performed on data of herds that had Status A on 1 January 2016. The primary outcome of interest was whether the herd had at least one ELISA result with an S/P  $\geq 1.0$  of a sample submitted in a 30-month period between 1 January 2016 and 30 June 2018. This 30-month frame was chosen to ensure that at least one herd examination was performed in each herd, given that the interval between the herd examinations in Status A herds is approximately 24 months. The secondary outcome of interest was whether the herd had at least two samples with an S/P  $\geq 1.0$  submitted in the same 30-month period. These outcome variables were chosen because, other than the loss of Status A, they are independent of farmers' decisions on confirmatory fecal testing of cattle with positive ELISA results.

The retrospective data collection and analyses to develop the model were restricted to all herds that matched the following selection criteria: (a) The initial assessment of the herd was prior to 1 January 2011; (b) The herd had a bulk milk tank number and was participating in the MQAP on 1 January 2014, 1 January 2016, and 1 July 2018; (c) The herd was assigned Status A by 1 January 2016; (d)  $\geq 50$  ELISA results of samples submitted from the herd between 1 January 2014 and 31 December 2015; (e) On 1 January 2016, the herd had at least 50 adult cattle ( $\geq 2$  years of age); (f) The S/P could be retrieved for at least one ELISA result of samples submitted from the herd between 1 January 2016 and 30 June 2018; (g) The predominant soil type of the postal code area of the herd could be retrieved.

Putative explanatory variables that were included in the full logistic regression model were: the province of the Netherlands; the predominant soil type in the upper 80-cm layer of soil in the terrestrial surface area of the 4-digit postal code of the herd (sand, bog, sandy loam, clay, or other); the herd size (number of cattle  $\geq 2$  years of age on 1 January 2016); the proportion of cattle present in the herd on 1 January 2016 that were born in another herd; the time elapsed since the last ELISA result with an S/P  $\geq 1.0$ ; and the 95th percentile of the distribution of the S/P ratios between 1 January 2014 and 31 December 2015. The 95th percentile of the S/P ratios was chosen because, in comparison with the 1st, 5th, 10th, 25th, 50th, 75th, and 90th percentiles, preliminary univariable receiver-operating characteristics (ROC) analyses indicated a higher area under the curve for prediction of at least one future positive ELISA result based on the 95th percentile. Furthermore, the 95th percentile had fewer missing values than the 99th percentile, as the 99th percentile is not meaningful in herds with less than 100 samples tested. To avoid overfitting, none of the interactions between the explanatory variables were tested. Preliminary analyses indicated nonlinearity of the effects of the time elapsed since the last ELISA result with an S/P  $\geq 1.0$  and the proportion of cattle that were born in another herd. Therefore, these variables were categorized (Table 1). Multicollinearity between the putative explanatory variables was

evaluated by linear regression [13]. A tolerance <0.25 and a variance inflation factor >4 were considered indicative of multicollinearity.

For both outcomes of interest, a final model was obtained in a backward elimination procedure, with a threshold probability for stepwise removal of 0.10 in the likelihood ratio test. Confounding was monitored by the changes in the regression coefficients. If elimination of a variable resulted in a change in the estimated regression coefficient of any other variable exceeding 25% or 0.1, in case of an estimate between  $-0.4$  and  $0.4$ , the eliminated variable was considered a potential confounder and was re-entered into the model. The model fit was evaluated with the Hosmer–Lemeshow goodness-of-fit test. To correct for overfitting, the model parameters were multiplied with a heuristic shrinkage factor [14]:

$$SF_{heuristic} = \frac{(\text{model } \chi^2 - \text{df})}{\text{model } \chi^2} \quad (1)$$

after which the intercept was calibrated to align the average predicted probability of at least one (primary outcome of interest) or two (secondary outcome of interest) samples with an S/P  $\geq 1.0$  between 1 January 2016 and 30 June 2018, with the observed probabilities. The predicted probability  $\pi_i$  of herd  $i$  to have at least one or two samples with an S/P  $\geq 1.0$  was calculated on the basis of the point estimates of the model coefficients. If the logistic regression model reads:

$$\text{logit}(\pi) = \ln\left(\frac{\pi}{1 - \pi}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (2)$$

then the  $\pi_i$  of herd  $i$  is calculated as:

$$\pi_i = \left[1 + e^{-(\beta_0, \text{adjusted} + SF_{heuristic}(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n))}\right]^{-1} \quad (3)$$

in which  $\beta_0, \text{adjusted}$  is the calibrated intercept, and  $\beta_1, \beta_2, \dots, \beta_n$  are the coefficients of the explanatory variables  $X_1, X_2, \dots, X_n$  in the logistic regression model. Categorical explanatory variables were included in the prediction model as dummies for the various levels of each variable, with the exception of the reference categories of the variables. The dummy variables were coded as 1 if applicable (true), and 0 if not applicable (false), for herd  $i$ . The performances of the predictive models were evaluated in ROC analyses on the predicted probabilities of at least one (primary outcome of interest) or two (secondary outcome of interest) samples with an S/P  $\geq 1.0$ , in comparison with the observed outcome in each herd. In the ROC analyses, the proportion of herds considered to be at an increased risk of positive ELISA results at various cut-off values  $\pi_c$ , as well as the sensitivity, specificity, and positive and negative predictive values of the prediction model at this cut-off value, were calculated by comparing the  $\pi_i$  of each herd in the dataset with the cut-off value  $\pi_c$ .

### 2.3. Internal Validation of the Predictive Models

As an internal validation, the predictive models were applied to a second dataset of herds with Status A on 1 January 2019, predicting whether the herd had at least one (primary outcome of interest) or two (secondary outcome of interest) ELISA results with an S/P  $\geq 1.0$  of a sample submitted in a 30-month period between 1 January 2019 and 30 June 2021.

The dataset used for this internal validation consisted of all herds that matched the following selection criteria: (a) The initial assessment of the herd was prior to 1 January 2013; (b) The herd had a bulk milk tank number and was participating in the MQAP on 1 January 2017, 1 January 2019, and 1 July 2021; (c) The herd was assigned Status A by 1 January 2019; (d)  $\geq 50$  ELISA results of the samples submitted between 1 January 2017 and 31 December 2018; (e) On 1 January 2019, the herd had at least 50 adult cattle ( $\geq 2$  years of age); (f) The S/P could be retrieved for at least one ELISA result of samples submitted from the herd between 1 January 2019 and 30 June 2021; (g) The predominant soil type of the postal code area of the herd could be retrieved.

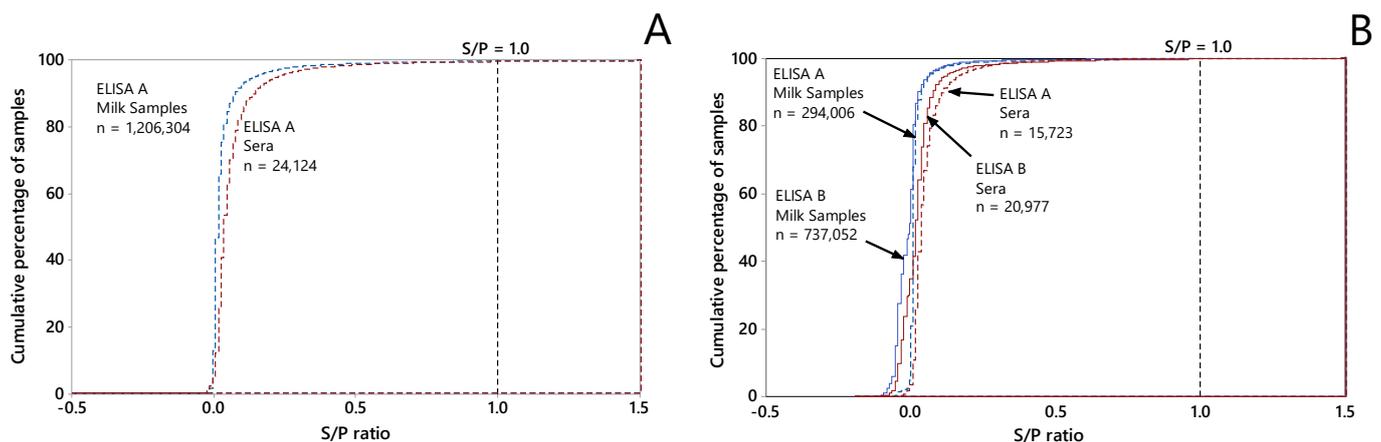
The data on the explanatory variables that were included in the predictive models were retrieved in a similar manner as in the development phase, with 1 January 2019 as a reference date. The province of the Netherlands, the predominant soil type, the herd size (number of cattle  $\geq 2$  years of age), the proportion of cattle present in the herd that were born in another herd, and the time elapsed since the last ELISA result with an  $S/P \geq 1.0$  were determined on 1 January 2019. The 95th percentile of the  $S/P$  ratios was calculated using ELISA results of samples submitted between 1 January 2017 and 31 December 2018.

Finally, the performance of the predictive models was evaluated in ROC analyses on the predicted probabilities of at least one (primary outcome of interest) or two (secondary outcome of interest) samples with an  $S/P \geq 1.0$ , in comparison with the observed outcome in each herd.

### 3. Results

#### 3.1. Development of the Predictive Models

A total of 8566 dairy herds with Status A on 1 January 2016 matched all the selection criteria for the training dataset. Between 1 January 2016 and 30 June 2018, 1,230,699 samples submitted from these herds were tested by ELISA. For 271 of these samples, the  $S/P$  could not be retrieved. The distribution of the  $S/P$  of the remaining 1,230,428 samples is shown in Figure 1. Only 5282 (0.4%) of the tested samples had an  $S/P \geq 1.0$ . These 5282 samples with an  $S/P \geq 1.0$  were clustered in 2347 (27%) of the 8566 dairy herds. In those 2347 herds, the mean proportion of samples with an  $S/P \geq 1.0$  was 1.3% (min: 0.07%; max: 13.2%). From 1126 (13%) of the 8566 herds, at least two samples with an  $S/P \geq 1.0$  were submitted.



**Figure 1.** Cumulative distributions of  $S/P$  ratios of serum and milk samples included in the outcome variable. (A) Training dataset: samples submitted between 1 January 2016 and 30 June 2018 from 8566 Dutch dairy herds. (B) Validation dataset: samples submitted between 1 January 2018 and 30 June 2021 from 8026 Dutch dairy herds.

All explanatory variables that were entered in the full logistic regression models were retained in the final logistic regression models (Tables 2 and 3). The numbers of adult cattle in the 8566 herds on 1 January 2016 ranged from 50 to 874 (median: 103; mean: 115) cattle. The 95th percentile of the  $S/P$  ratios between 1 January 2014 and 31 December 2015 ranged from 0.02 to 1.60 (median: 0.13; mean: 0.16). In both analyses, the  $SF_{heuristic}$  equalled 0.98 and, consequently, the adjusted beta values in the predictive models only marginally differed from the estimated beta values in the final logistic regression models (Tables 2 and 3).

**Table 2.** Logistic regression model <sup>†</sup> for the prediction of at least one ELISA result with an S/P  $\geq 1.0$  between January 2016 and June 2018 (training dataset). Data of dairy herds with Status A in the Dutch milk quality assurance program for paratuberculosis on 1 January 2016 (training dataset: 8566 herds) and on 1 January 2019 (validation dataset: 8026 herds).

Variable	Level	Training Dataset					Validation Dataset	
		Number of Herds	$\beta$ (S.E.)	$p$ (LR Test)	$p$ (Wald)	OR (95% CI)	Adjusted $\beta$ in Predictive Model	Number of Herds
Province	Gelderland	1504	Reference	<0.001				1408
	Drenthe	404	0.91 (0.13)		<0.001	2.48 (1.91; 3.23)	0.89	379
	Flevoland	139	0.51 (0.21)		0.016	1.66 (1.10; 2.51)	0.50	131
	Fryslân	1202	0.90 (0.10)		<0.001	2.46 (2.01; 3.00)	0.88	1145
	Groningen	375	0.74 (0.14)		<0.001	2.09 (1.58; 2.75)	0.72	374
	Limburg	297	0.18 (0.18)		0.317	1.20 (0.84; 1.72)	0.18	256
	Noord Brabant	1343	0.34 (0.10)		0.001	1.41 (1.15; 1.72)	0.34	1179
	Noord Holland	427	0.30 (0.14)		0.037	1.34 (1.02; 1.78)	0.29	409
	Overijssel	1664	0.41 (0.10)		<0.001	1.51 (1.24; 1.83)	0.40	1588
	Utrecht	541	0.67 (0.13)		<0.001	1.95 (1.51; 2.52)	0.66	500
	Zuid Holland	565	0.30 (0.14)		0.032	1.35 (1.03; 1.77)	0.29	554
	Zeeland	105	0.29 (0.25)		0.242	1.34 (0.82; 2.19)	0.29	103
Soil type	Sand	4780	Reference	0.002				4404
	Clay	1467	0.29 (0.09)		0.001	1.33 (1.13; 1.58)	0.28	1414
	Bog	1050	0.29 (0.09)		0.002	1.34 (1.11; 1.61)	0.29	1014
	Sandy loam	1133	0.17 (0.09)		0.056	1.19 (1.00; 1.41)	0.17	1071
	Other	136	0.42 (0.23)		0.073	1.52 (0.96; 2.40)	0.41	123
Number of adult cattle		8566	0.0048 (0.0005)	<0.001	<0.001	1.01 (1.00; 1.01)	0.0047	8026
Time since last ELISA result with S/P $\geq 1.0$ (days)	$\leq 730$	1587	1.20 (0.07)		<0.001	3.32 (2.87; 3.84)	1.18	1413
	>730 and $\leq 1460$	1492	1.41 (0.07)		<0.001	4.08 (3.55; 4.69)	1.38	1329
	>1460 and $\leq 2190$	718	0.79 (0.10)		<0.001	2.19 (1.82; 2.64)	0.77	859
	>2190 and $\leq 2920$	453	0.44 (0.12)		<0.001	1.55 (1.22; 1.97)	0.43	501
	> 2920 or never	4316	Reference	<0.001				3924
Proportion of cattle born elsewhere	0	3778	Reference	<0.001				3720
	>0 and $\leq 0.05$	2551	0.14 (0.06)		0.026	1.15 (1.02; 1.31)	0.14	2112
	>0.05 and $\leq 1.0$	2237	0.34 (0.07)		<0.001	1.41 (1.24; 1.60)	0.34	2194
95th percentile of S/P		8566	2.38 (0.24)	<0.001	<0.001	10.8 (6.71; 17.5)	2.34	8026
Intercept		8566	-3.32 (0.11)		<0.001		-3.28	8026

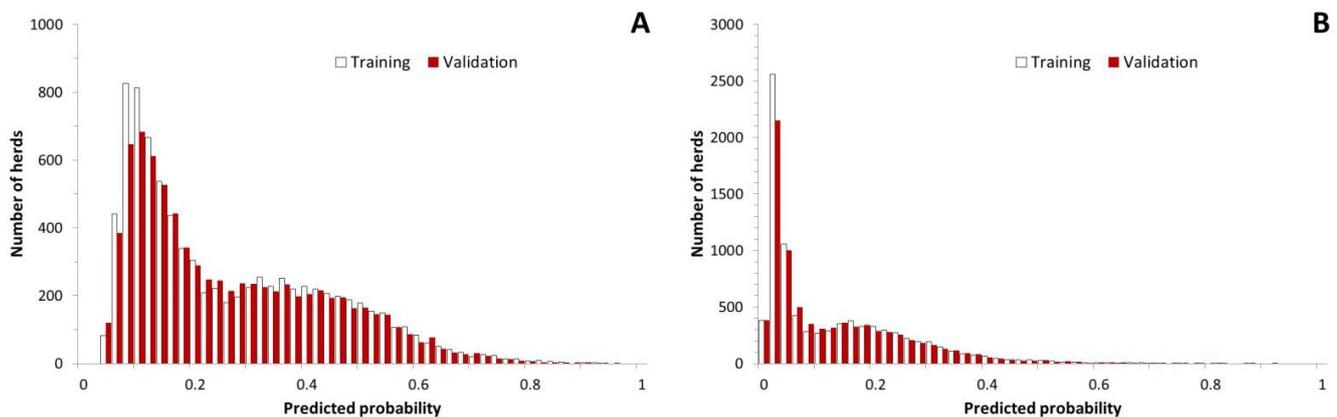
<sup>†</sup>  $-2 \log$  likelihood = 8707.2; Nagelkerke,  $R^2 = 0.21$ ; Hosmer–Lemeshow test,  $\chi^2 = 7.05$ ,  $df = 8$ ,  $p = 0.531$ .

**Table 3.** Logistic regression model <sup>†</sup> for the prediction of at least two ELISA results with an S/P  $\geq 1.0$  between January 2016 and June 2018 (training dataset). Data of dairy herds with Status A in the Dutch milk quality assurance program for paratuberculosis on 1 January 2016 (training dataset: 8566 herds) and on 1 January 2019 (validation dataset: 8026 herds).

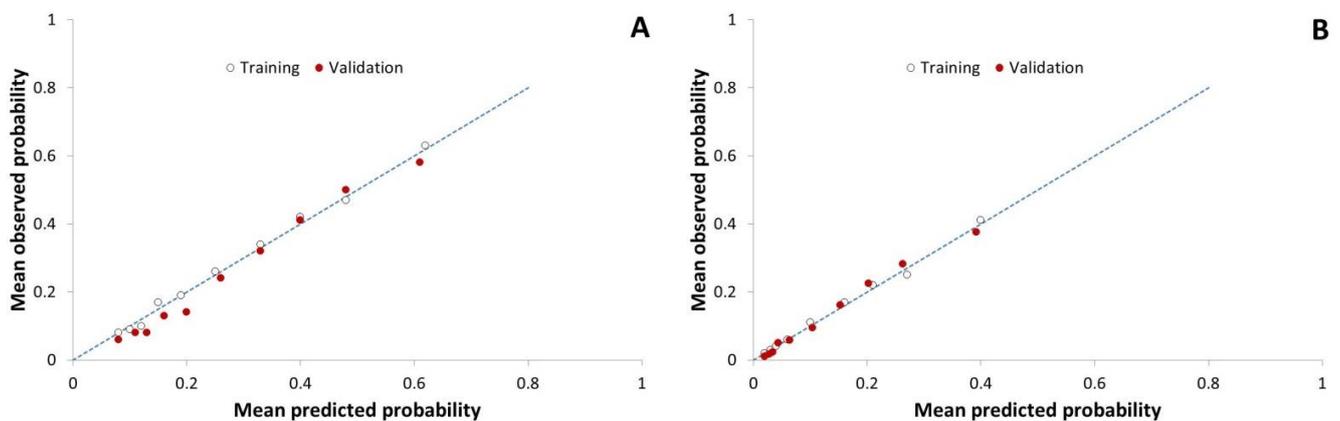
Variable	Level	Training Dataset					Adjusted $\beta$ in Predictive Model	Validation Dataset Number of Herds
		Number of Herds	$\beta$ (S.E.)	$p$ (LR Test)	$p$ (Wald)	OR (95% CI)		
Province	Gelderland	1504	Reference	<0.001				1408
	Drenthe	404	0.93 (0.18)		0.000	2.54 (1.80; 3.57)	0.91	379
	Flevoland	139	0.34 (0.27)		0.208	1.41 (0.83; 2.40)	0.34	131
	Fryslân	1202	0.96 (0.14)		0.000	2.60 (1.99; 3.40)	0.94	1145
	Groningen	375	0.65 (0.18)		0.000	1.92 (1.34; 2.74)	0.64	374
	Limburg	297	0.10 (0.26)		0.709	1.10 (0.66; 1.83)	0.09	256
	Noord Brabant	1343	0.33 (0.15)		0.021	1.40 (1.05; 1.86)	0.33	1179
	Noord Holland	427	0.29 (0.19)		0.132	1.33 (0.92; 1.94)	0.28	409
	Overijssel	1664	0.42 (0.14)		0.003	1.53 (1.16; 2.01)	0.41	1588
	Utrecht	541	0.56 (0.18)		0.002	1.75 (1.23; 2.49)	0.55	500
	Zuid Holland	565	0.31 (0.19)		0.102	1.37 (0.94; 1.99)	0.31	554
Zeeland	105	−0.28 (0.39)		0.474	0.76 (0.35; 1.62)	−0.27	103	
Soil type	Sand	4780	Reference	0.016				4404
	Clay	1467	0.32 (0.11)		0.004	1.37 (1.11; 1.70)	0.31	1414
	Bog	1050	0.15 (0.12)		0.227	1.16 (0.91; 1.46)	0.14	1014
	Sandy loam	1133	0.19 (0.12)		0.100	1.21 (0.96; 1.53)	0.19	1071
	Other	136	0.64 (0.30)		0.033	1.90 (1.05; 3.43)	0.63	123
Number of adult cattle		8566	0.01 (0.00)	<0.001	<0.001	1.01 (1.00; 1.01)		8026
Time since last ELISA result with S/P $\geq 1.0$ (days)	$\leq 730$	1587	1.74 (0.11)		<0.001	5.68 (4.60; 7.01)	1.70	1413
	>730 and $\leq 1460$	1492	1.87 (0.10)		<0.001	6.46 (5.26; 7.93)	1.83	1329
	>1460 and $\leq 2190$	718	1.33 (0.13)		<0.001	3.79 (2.92; 4.92)	1.30	859
	>2190 and $\leq 2920$	453	0.81 (0.18)		<0.001	2.24 (1.57; 3.20)	0.79	501
	>2920 or never	4316	Reference	<0.001			1.70	3924
Proportion of cattle born elsewhere	0	3778	Reference	<0.001				3720
	>0 and $\leq 0.05$	2551	0.20 (0.09)		0.019	1.22 (1.03; 1.45)	0.20	2112
	>0.05 and $\leq 1.0$	2237	0.33 (0.09)		<0.001	1.39 (1.18; 1.65)	0.32	2194
95th percentile of S/P		8566	2.52 (0.27)	<0.001	<0.001	12.4 (7.28; 21.1)	2.46	8026
Intercept		8566	−4.83 (0.16)		<0.001		−4.77	8026

<sup>†</sup>  $-2 \log$  likelihood = 5545.6; Nagelkerke,  $R^2 = 0.23$ ; Hosmer–Lemeshow test,  $\chi^2 = 7.11$ ,  $df = 8$ ,  $p = 0.525$ .

With both outcomes of interest, the distributions of the predicted probabilities in the training set were skewed: there were few herds with a high ( $\geq 0.8$ ) predicted probability of at least one ELISA result with an  $S/P \geq 1.0$  (Figure 2). However, the predicted probabilities of at least one (primary outcome of interest) or two (secondary outcome of interest) samples with an  $S/P \geq 1.0$  were close to the observed probabilities (Figure 3). In the training dataset, the predictive models for at least one (primary outcome of interest) or two (secondary outcome of interest) samples with an  $S/P \geq 1.0$  had areas under the ROC curve equal to 0.747 (95% CI: 0.735; 0.759) and 0.791 (0.778; 0.804), respectively (Figure 4).



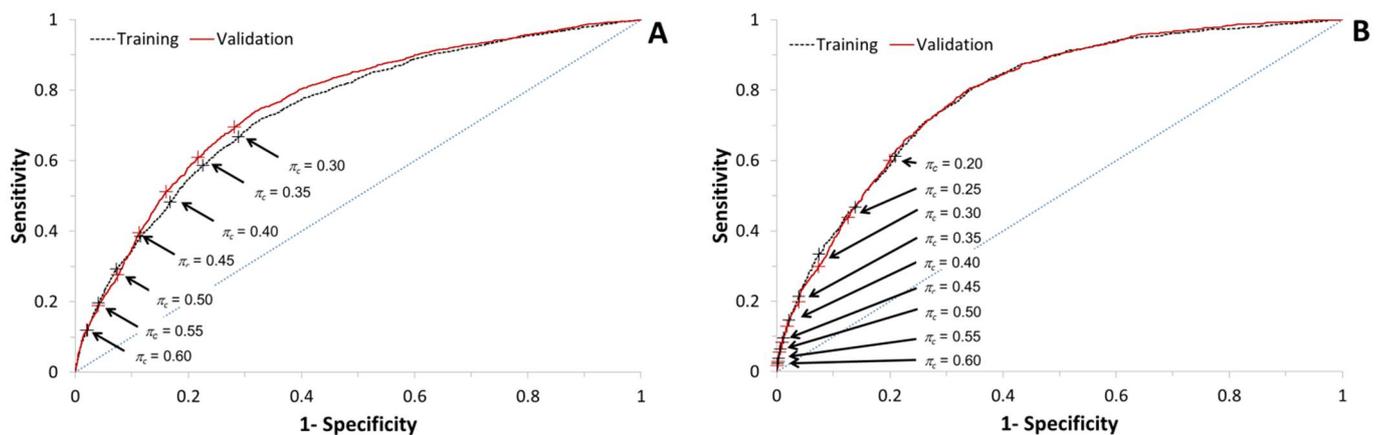
**Figure 2.** Distribution of predicted probabilities of (A) at least one, and (B) at least two ELISA result(s) with an  $S/P \geq 1.0$  in the training dataset (8566 dairy herds with Status A on 1 January 2016) and the validation dataset (8026 dairy herds with Status A on 1 January 2019).



**Figure 3.** Mean observed probabilities and mean predicted probabilities of (A) at least one, and (B) at least two ELISA result(s) with  $S/P \geq 1.0$  for observations in each of the ten deciles of the predicted probabilities in the training dataset (8566 dairy herds with Status A on 1 January 2016) and the validation dataset (8026 dairy herds with Status A on 1 January 2019). The dashed lines represent the identity line where the mean observed probability equals the mean predicted probability.

### 3.2. Internal Validation

A total of 8026 dairy herds with Status A on 1 January 2019 matched all selection criteria for the validation dataset. Between 1 January 2019 and 30 June 2021, 1,068,046 samples submitted from these herds were tested by ELISA. For 288 of these samples, the  $S/P$  ratio could not be retrieved. The distribution of the  $S/P$  ratios of the remaining 1,067,758 samples is shown in Figure 1. Only 4764 (0.4%) of these samples had an  $S/P \geq 1.0$ . These 4764 samples with an  $S/P \geq 1.0$  were clustered in 2036 (25%) of the 8026 dairy herds. In those 2036 herds, the mean proportion of samples with an  $S/P \geq 1.0$  was 1.3% (min: 0.08%; max: 50%). From 1041 (13%) of the 8026 herds, at least two samples with an  $S/P \geq 1.0$  were submitted.



**Figure 4.** Receiver-operating characteristics curve of the predictive models in the training dataset (8566 dairy herds with Status A in the Dutch milk quality assurance program for paratuberculosis on 1 January 2016) and the validation dataset (8026 dairy herds with Status A in the Dutch milk quality assurance program for paratuberculosis on 1 January 2016). **(A)** Predictive model for at least one ELISA result with an  $S/P \geq 1.0$ . **(B)** Predictive model for at least two ELISA results with an  $S/P \geq 1.0$ . Arrows and + signs indicate various cut-off values  $\pi_c$  of the predicted probability  $\pi$  of at least one or at least two ELISA result(s) with an  $S/P \geq 1.0$  on the ROC curves. The dashed lines represent the identity line.

The numbers of herds in the various levels of the predictive variables in the predictive model are shown in Table 2. The numbers of adult cattle in the 8026 herds on 1 January 2019 ranged from 50 to 762 (median: 97; mean: 110) cattle. The 95th percentile of the  $S/P$  ratios between 1 January 2017 and 31 December 2018 ranged from 0.00 to 1.12 (median: 0.14; mean: 0.17). The distribution of the predicted probabilities of at least one sample with an  $S/P \geq 1.0$  was right-skewed in the validation dataset as well (Figure 2). However, in contrast to the training set, in the validation dataset, the observed probabilities of at least one sample with an  $S/P \geq 1.0$  were slightly overestimated by the prediction model at the lower range of the predicted probabilities ( $<0.3$ ; Figure 3A). In the validation dataset, the predictive models for at least one (primary outcome of interest) or two (secondary outcome of interest) samples with an  $S/P \geq 1.0$  had areas under the ROC curve equal to 0.762 (95% CI: 0.750; 0.774) and 0.792 (0.779; 0.806), respectively (Figure 4).

### 3.3. Comparison of Model Performances in the Training and Validation Datasets

For each of the two models, the areas under the ROC curves in the training and validation datasets were comparable. Depending on the desired characteristics of the predictive models, a cut-off predicted probability  $\pi_c$  of at least one ELISA result with an  $S/P \geq 1.0$  can be chosen that corresponds to a data point on the ROC curves (Figure 4). The proportion of herds with a predicted probability exceeding the cut-off value, i.e., the proportion of herds that would receive an alert, and the positive and negative predictive values at the various cut-off values  $\pi_c$  of the predicted probability  $\pi$  are shown in Table 4.

For each of the two predictive models, the results obtained differed only marginally between the training and validation datasets. For example, at  $\pi_c = 0.40$  in the predictive model for at least one sample with an  $S/P \geq 1.0$ , 25% of the Status A herds would receive an alert in both datasets. Of the herds receiving an alert, 52% would indeed have at least one ELISA result with an  $S/P \geq 1.0$  in the next 30 months. Of the herds not receiving an alert, 81% (training set) to 83% (validation set) would receive no ELISA result with an  $S/P \geq 1.0$  in the next 30 months, while the remaining 17 to 19% would receive at least one ELISA result with an  $S/P \geq 1.0$  (Table 4).

**Table 4.** Proportions of herds with predicted probabilities larger than the cut-off value, positive predictive values (PVP), and negative predictive values (PVP) at various cut-off values  $\pi_c$  of the predicted probability  $\pi$  of at least one or at least two sample(s) with an S/P  $\geq 1.0$ . Data of 8566 dairy herds with Status A on 1 January 2016 (training dataset), and 8026 dairy herds with Status A on 1 January 2019 (validation dataset). Results discussed as examples in Sections 3.3 and 3.4 are printed in bold.

Predictive Model	Cut-Off Predicted Probability $\pi_c$	Training Dataset				Validation Dataset					
		Proportion of Herds with Predicted Probability > Cut-Off $\pi_c$	Se	Sp	PVP	PVN	Proportion of Herds with Predicted Probability > Cut-Off $\pi_c$	Se	Sp	PVP	PVN
$\geq 1$ sample with S/P $\geq 1.0$	0.30	0.39	0.67	0.71	0.47	0.85	0.39	0.69	0.72	0.46	0.87
	0.35	0.32	0.59	0.77	0.49	0.83	0.32	0.61	0.78	0.49	0.85
	<b>0.40</b>	<b>0.25</b>	<b>0.48</b>	<b>0.83</b>	<b>0.52</b>	<b>0.81</b>	<b>0.25</b>	<b>0.51</b>	<b>0.84</b>	<b>0.52</b>	<b>0.83</b>
	0.45	0.19	0.38	0.88	0.56	0.79	0.18	0.39	0.89	0.54	0.81
	0.50	0.13	0.29	0.93	0.60	0.78	0.13	0.28	0.92	0.55	0.79
	0.55	0.08	0.20	0.96	0.64	0.76	0.08	0.19	0.96	0.61	0.78
	0.60	0.05	0.12	0.98	0.67	0.75	0.05	0.12	0.98	0.67	0.77
$\geq 2$ samples with S/P $\geq 1.0$	<b>0.20</b>	<b>0.26</b>	<b>0.61</b>	<b>0.79</b>	<b>0.31</b>	<b>0.93</b>	<b>0.25</b>	<b>0.60</b>	<b>0.80</b>	<b>0.31</b>	<b>0.93</b>
	0.25	0.17	0.46	0.87	0.34	0.91	0.17	0.44	0.87	0.34	0.91
	0.30	0.11	0.33	0.92	0.40	0.90	0.10	0.30	0.93	0.37	0.90
	0.35	0.06	0.21	0.96	0.45	0.89	0.06	0.20	0.96	0.43	0.89
	0.40	0.04	0.15	0.98	0.50	0.88	0.03	0.13	0.98	0.51	0.88
	0.45	0.02	0.10	0.99	0.53	0.88	0.02	0.08	0.99	0.56	0.88
	0.50	0.01	0.06	0.99	0.58	0.88	0.01	0.05	0.99	0.61	0.88
	0.55	0.01	0.04	1.00	0.60	0.87	0.01	0.03	1.00	0.59	0.87
0.60	0.00	0.02	1.00	0.62	0.87	0.00	0.02	1.00	0.61	0.87	

### 3.4. Comparison of Performances of the Two Predictive Models

In both the training and validation sets, the predicted probability of a herd to have at least two positive ELISA results was always lower than the predicted probability of the herd to have at least one positive ELISA result. The area under the ROC curves (Figure 4) was higher for the model predicting  $\geq 2$  samples with an  $S/P \geq 1.0$ , in comparison to the model predicting  $\geq 1$  sample with an  $S/P \geq 1.0$ . At a given cut-off probability  $\pi_c$ , the model predicting  $\geq 2$  samples with an  $S/P \geq 1.0$  resulted in a lower proportion of herds with a predicted probability above the cut-off, a lower sensitivity and positive predictive value, and a higher specificity and negative predictive value than the model predicting  $\geq 1$  sample with an  $S/P \geq 1$  (Figure 4, Table 4).

However, at similar proportions of herds with a predicted probability above the cut-off (e.g., a cut-off predicted probability of 0.40, corresponding to 25% of herds in the model predicting  $\geq 1$  sample with an  $S/P \geq 1.0$ , and a cut-off predicted probability of 0.20, corresponding to 25 to 26% of herds in the model predicting  $\geq 2$  samples with an  $S/P \geq 1.0$ ), the model predicting  $\geq 2$  samples with an  $S/P \geq 1.0$  resulted in a lower positive predictive value and a higher negative predictive value than the model predicting  $\geq 1$  sample with an  $S/P \geq 1.0$  (Table 3).

## 4. Discussion

The aim of the present study was to develop a predictive model to alert farmers with Status A if they are at an increased risk of positive ( $S/P \geq 1.0$ ) ELISA results. The resulting predictive models were able to discriminate between herds with an increased risk (in the range of 31 to 67%, depending on the choice of model and cut-off, and shown as the positive predictive values in Table 4), and herds with a rather low risk (in the range of 7 to 25%, being the complement of the negative predictive values shown in Table 4). The results show that the performances of both models were remarkably stable in time, with the area under the ROC curves being even higher in the validation dataset than in the training dataset. Even though the model predicting  $\geq 2$  positive samples resulted in a slightly higher area under the ROC curve than the model predicting  $\geq 1$  positive sample, the choice between the two models should be primarily based on the required test characteristics. At similar proportions of herds with a predicted probability above the cut-off, the model predicting  $\geq 1$  positive sample resulted in a higher positive predictive value but a lower negative predictive value than the model predicting  $\geq 2$  samples. Because the aim of the present study was to identify herds with an increased risk of positive ELISA results (whether true or false positive), a higher positive predictive value would be preferred.

The predictive models were based on the data available in the central datasets and included, as explanatory variables, the province, soil type, number of adult cattle, proportion of cattle born elsewhere, time evolved since the last previous positive ELISA result and the 95th percentile of the  $S/P$  ratios of samples submitted in the previous 24 months. These explanatory variables were based on the results of previous studies, and their effects were considered biologically plausible. However, it is important that the outcome in the present study was not simply the presence or absence of Map infection, but, rather, test-positivity in herds that were previously test-negative and hence acquired the preferred herd Status A. Thus, effects observed in the present study may differ from the results of studies that focus on presence or absence of Map infection, or that focus on presence or absence of positive test results.

In line with the results of the present study, regional differences of the paratuberculosis prevalence have been described previously in the Netherlands [15], as well as in other countries [16–19]. In this study, in comparison to the sand soil type, clay was found to be associated with a higher probability of positive results in both models, whereas bog was significantly associated with the probability of  $\geq 1$  sample, but not the probability of  $\geq 2$  samples, with an  $S/P \geq 1.0$ . The effects of soil type on Map transmission may be confounded by effects on the Map transmission of farming practices that are associated with soil type. In previous studies, high soil iron content [20,21], low pH [16,21–24], low

silt content [25], high silt content [19], high loam content [19], high clay content [19,20], and high organic carbon content [20] were associated with Map prevalence or the incidence of clinical paratuberculosis. Thus, the inclusion of soil type in our predictive models is coherent with pre-existing evidence.

Several biological pathways may contribute to the observed association between the number of adult cattle and the probability of samples with an S/P  $\geq 1.0$ . Firstly, in random mixing conditions, the probability of infectious contacts between susceptible and infectious cattle increases with herd size, resulting in an increased transmission until an endemic equilibrium is reached [26,27], as well as in a reduced probability of fade-out of the infection [28]. Several field studies have observed a positive association between herd size and the incidence of paratuberculosis or the (apparent or true) prevalence of Map infection [19,29–36]. The effect of herd size may be partially counteracted by effects of herd size on herd management, given that larger herds more frequently have separate facilities for calf-rearing [37]. Secondly, the number of samples tested increases with herd size. Serum and milk ELISAs have imperfect sensitivities and specificities [38,39]. With an increasing number of samples tested, the probability of at least one false positive result as well as, in infected herds, the probability of at least one true positive result increase [40].

In a consensus statement of the International Association for Paratuberculosis, appropriately documenting the probability of the freedom from Map infection was considered an important aspect of herd classification programs [41]. In simulation modeling studies, this probability of freedom from Map depended on herd size, the number of introduced cattle, and the time during which only negative test results were obtained [40,42]. These results are in line with our observations of the effects of herd size, the proportion of introduced cattle, and the time since the last ELISA result with an S/P  $\geq 1.0$  on the probability of obtaining positive ELISA results.

Predictive modeling to predict the loss of a preferred herd status has been applied before, with varying success, to control programs for other infectious diseases, such as BVD and salmonellosis [43,44]. In their study on BVD, Gates et al. [43] suggested to annually test replacement heifers for BVD virus in order to reduce the effect of introduction of persistently infected heifers into the dairy herd. Likewise, the annual testing of nulli-parous heifers for antibodies against Map, or Map shedding, might enable the earlier detection of a Map infection in Status A herds, given that a considerable proportion of infected young stock become test-positive before adulthood [45–47].

In the present study, the proportion of cattle born in other herds was used as a proxy for potentially infectious contacts between herds through the introduction of cattle. Cattle purchases have been found to be a risk factor for Map infection in other studies as well [29,48,49]. However, the proportion of cattle born in other herds does not cover risks of other cattle movements between herds, such as rearing of young stock of multiple herds together in specialized young stock herds. No data on herd management were available, other than herd size and the proportion of cattle born in other herds. It is generally assumed that herd management has a major effect on the spread of Map. However, in both Danish and Minnesota dairy herds, herd management was found to be a poor predictor of paratuberculosis seroprevalence [48,50]. Presumably, there is a time lag of at least 4 to 5 years (a one-cow generation) before effects of changes in herd management become apparent [50]. Hence, in order to substantially improve our model predictions by adding data on (changes in) herd management, a long-term inventory of herd management in participating herds may be required, which may not be practically feasible.

In this study, a cut-off S/P ratio of 1.0 was used in both the dependent and explanatory variables in the models, whereas in the MQAP various cut-off values between 0.9 and 1.0 were used, depending on the ELISA kit, sample type (milk or serum), and date of submission of the sample. However, few samples had an S/P between 0.9 and 1.1 (Figure 1). Therefore, the proportion of positive results was considered rather insensitive to the exact choice of the cut-off value in this range of S/P ratios. The cut-off S/P ratio was chosen for the Dutch MQAP to increase the specificity and positive predictive value of the ELISA [12].

An important argument for this was that the objective of the program was to reduce the milk burden of MAP [7–10,12]. The MQAP, therefore, particularly aimed to identify high shedders that are known to have relatively high ELISA S/P values. Obviously, this decision also resulted in a relative low sensitivity and negative predictive value of the test.

The median values of the 95th percentile of the S/P ratios in the training and validation datasets were 0.13 and 0.14, respectively. At the low range of S/P ratios, some batch-to-batch variation may be expected, as verification of new batches of an ELISA kit is mainly focused on test performance around the cut-off values. However, the observed stability of the performance of our predictive models over time indicates that the model performance was rather insensitive to such batch-to-batch variation.

Our models were developed with annual predictions at a single point in time for all Status A herds in mind, irrespective of the date of the last herd examination in these herds. This would allow all data to be processed once per year, including an update of the coefficients of the model variables based on the most recent available data, followed by application of the model to predict which of the herds with Status A at the time have an increased risk of future positive results, and finally sending an alert by letter or email to the farmers of those herds. This approach requires selection of a cut-off value,  $\pi_c$ , to discriminate between herds with high and low risks of future positive results. The choice of this cut-off value needs to be based on careful judgement of the relationship between the cut-off value, the proportion of herds that will receive an alert, and the positive and negative predictive values of the alert (Table 4). By increasing the cut-off value, the positive predictive value can be increased, but this comes at the price of a lower proportion of herds receiving an alert and a lower negative predictive value (i.e., a higher proportion of herds that will have future positive results, despite not having received an alert).

An alternative approach would be to provide each farmer with a herd-specific prediction shortly after the results of a herd examination become available. To do so, the coefficients of the explanatory variables in the models should be re-estimated, as the distribution of some of these variables may be different (such as the distribution of the time since the last positive ELISA result). Given that results of herd examinations of participating herds become available at varying time points throughout the year, development of a software tool for the automated calculation of the model predictions shortly after each herd examination would seem to be a prerequisite for implementing such an alternative approach in a cost-efficient manner.

External validation of predictive models is generally preferred above internal validation [14]. In this study, external validation was not feasible as the milk quality assurance program and risk factors, such as province, were specific to the Netherlands. Nevertheless, the approach followed in this study may be applicable to other disease control programs in other cattle populations.

At present, the prediction models developed in this study are not yet operational. Available options will be discussed with the stakeholders and decision-makers involved in the control of paratuberculosis. The practical applications of our models include management of the expectations of farmers, veterinarians, and other stakeholders. Farmers with an increased risk of positive ELISA results may use the model predictions to make informed decisions on additional control measures, such as preventive management measures, to reduce the potential spread of Map in their herds. It remains to be seen whether farmers do consider a risk in the range of 31 to 67% of positive ELISA results (i.e., PVP values in Table 4) sufficiently high to warrant allocation of resources to such preventive management measures. In herds receiving an alert of an increased risk of positive results, environmental sampling may be a suitable approach, at low costs, to segregate herds with false positive alerts and herds with high risks of the spread of Map [51–55].

## 5. Conclusions

The aim of the present study was to develop a predictive model to alert farmers with the preferred Status A in the Dutch milk quality assurance program if they are at an increased risk of positive ELISA results in a subsequent 30-month period. On the basis of the results of this study, we conclude that discrimination of herds with high (52%) and low (17%) risks of positive ELISA results is feasible. This might help Status A farmers with the highest risks of positive results to make better informed decisions regarding the need to take additional control measures to prevent the spread of any undetected Map infections.

**Author Contributions:** Conceptualization, M.F.W. and Y.H.S.; methodology, M.F.W.; software, M.F.W.; validation, M.F.W.; formal analysis, M.F.W.; investigation, M.F.W.; resources, M.F.W., M.A., T.D. and Y.H.S.; data curation, M.F.W.; writing—original draft preparation, M.F.W.; writing—review and editing, M.F.W., M.A., T.D. and Y.H.S.; visualization, M.F.W.; supervision, Y.H.S.; project administration, M.F.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** No animals were sampled or used in any other form for this retrospective study. Therefore this study was not considered an animal experiment and ethical review and approval were waived for this study.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The herds or groups of herds included in this study may be identifiable from their unique combination of explanatory variables. Therefore, the datasets analyzed in this study cannot, and will not, be made available to readers.

**Acknowledgments:** The authors thank F. de Vries (Wageningen Environmental Research) for providing the data on the soil types.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Garcia, A.B.; Shalloo, L. Invited review: The economic impact and control of paratuberculosis in cattle. *J. Dairy Sci.* **2015**, *98*, 5019–5039. [[CrossRef](#)] [[PubMed](#)]
2. Hasanova, L.; Pavlik, I. Economic impact of paratuberculosis in dairy herds: A review. *Vet. Med.* **2006**, *51*, 193–211. [[CrossRef](#)]
3. Waddell, L.A.; Rajic, A.; Stark, K.D.; McEwen, S.A. The zoonotic potential of *Mycobacterium avium* ssp. *paratuberculosis*: A systematic review and meta-analyses of the evidence. *Epidemiol. Infect.* **2015**, *143*, 3135–3157. [[PubMed](#)]
4. Duffy, S.C.; Behr, M.A. Paratuberculosis and Crohn's Disease. In *Paratuberculosis. Organism, Disease, Control*, 2nd ed.; Behr, M.A., Stevenson, K., Kapur, V., Eds.; CABI: Wellingford, UK, 2020; pp. 29–44.
5. Whittington, R.; Donat, K.; Weber, M.F.; Kelton, D.; Nielsen, S.S.; Eisenberg, S.; Arrigoni, N.; Juste, R.; Saez, J.L.; Dhand, N.; et al. Control of paratuberculosis: Who, why and how. A review of 48 countries. *BMC Vet. Res.* **2019**, *15*, 198. [[CrossRef](#)]
6. Geraghty, T.; Graham, D.A.; Mullowney, P.; More, S.J. A review of bovine Johne's disease control activities in 6 endemically infected countries. *Prev. Vet. Med.* **2014**, *116*, 1–11. [[CrossRef](#)]
7. van Roermund, H.J.W.; Weber, M.F.; de Koeijer, A.A.; Velthuis, A.G.J.; de Jong, M.C.M. Development of a milk quality assurance program for paratuberculosis: From within- and between herd dynamics to economic decision analysis. In Proceedings of the 8th International Colloquium on Paratuberculosis, Copenhagen, Denmark, 14–18 August 2005; pp. 51–59.
8. Weber, M.F.; Nielen, M.; Velthuis, A.G.; van Roermund, H.J.W. Milk quality assurance for paratuberculosis: Simulation of within-herd infection dynamics and economics. *Vet. Res.* **2008**, *39*, 12. [[CrossRef](#)]
9. Weber, M.F.; van Schaik, G.; Aalberts, M.; Velthuis, A.G.J. Milk quality assurance for paratuberculosis: Progress obtained in the cohort of dairy herds that entered the program in 2006–2007. *Bull. Int. Dairy Fed.* **2014**, *475*, 55–61.
10. Weber, M.F. Milk quality assurance for paratuberculosis in the national Dutch dairy herd. *Paratuberculosis Newsl.* **2008**, *3*, 6.
11. Weber, M.F.; Groenendaal, H.; van Roermund, H.J.W.; Nielen, M. Simulation of alternatives for the Dutch Johne's disease certification-and-monitoring program. *Prev. Vet. Med.* **2004**, *62*, 1–17. [[CrossRef](#)]
12. Weber, M.F.; Aalberts, M.; Dijkstra, T.; de Roo, A. Results of milk quality assurance programme for paratuberculosis in Dutch dairy herds indicate reduced transmission of the infection. *Bull. Int. Dairy Fed.* **2018**, *493*, 23–34.
13. Field, A. Logistic regression. In *Discovering Statistics Using SPSS*; Field, A., Ed.; Sage Publications: London, UK, 2005; pp. 218–268.
14. Steyerberg, E.W. *Clinical Prediction Models. A Practical Approach to Development, Validation and Updating*; Springer: New York, NY, USA, 2009; p. 497.
15. Muskens, J.; Barkema, H.W.; Russchen, E.; Van Maanen, K.; Schukken, Y.H.; Bakker, D. Prevalence and regional distribution of paratuberculosis in dairy herds in The Netherlands. *Vet. Microbiol.* **2000**, *77*, 253–261. [[CrossRef](#)]

16. Kopecky, K.E. Distribution of paratuberculosis in Wisconsin, by soil regions. *J. Am. Vet. Med. Assoc.* **1977**, *170*, 320–324. [PubMed]
17. Bihmann, K.; Nielsen, S.S.; Toft, N.; Ersboll, A.K. Spatial differences in occurrence of paratuberculosis in Danish dairy herds and in control programme participation. *Prev. Vet. Med.* **2012**, *103*, 112–119. [CrossRef] [PubMed]
18. Bihmann, K.; Nielsen, S.S.; Ersboll, A.K. Spatial pattern in prevalence of paratuberculosis infection diagnosed with misclassification in Danish dairy herds in 2009 and 2013. *Spat. Spatiotemporal. Epidemiol.* **2016**, *16*, 1–10. [CrossRef]
19. Kanankege, K.S.T.; Machado, G.; Zhang, L.; Dokkebakken, B.; Schumann, V.; Wells, S.J.; Perez, A.M.; Alvarez, J. Use of a voluntary testing program to study the spatial epidemiology of Johne’s disease affecting dairy herds in Minnesota: A cross sectional study. *BMC Vet. Res.* **2019**, *15*, 429. [CrossRef]
20. Dhand, N.K.; Eppleston, J.; Whittington, R.J.; Toribio, J.A. Association of farm soil characteristics with ovine Johne’s disease in Australia. *Prev. Vet. Med.* **2009**, *89*, 110–120. [CrossRef]
21. Johnson-Ifeorlundu, Y.; Kaneene, J.B. Distribution and environmental risk factors for paratuberculosis in dairy cattle herds in Michigan. *Am. J. Vet. Res.* **1999**, *60*, 589–596.
22. Jansen, J. Paratuberculosis. *J. Am. Vet. Med. Assoc.* **1948**, *112*, 52–54.
23. Johnson-Ifeorlundu, Y.J.; Kaneene, J.B. Relationship between soil type and *Mycobacterium paratuberculosis*. *J. Am. Vet. Med. Assoc.* **1997**, *210*, 1735–1740.
24. Scott, H.M.; Sorensen, O.; Wu, J.T.; Chow, E.Y.; Manninen, K.; VanLeeuwen, J.A. Seroprevalence of *Mycobacterium avium* subspecies *paratuberculosis*, *Neospora caninum*, *Bovine leukemia virus*, and *Bovine viral diarrhoea virus* infection among dairy cattle and herds in Alberta and agroecological risk factors associated with seropositivity. *Can. Vet. J.* **2006**, *47*, 981–991.
25. Ward, M.P.; Perez, A.M. Association between soil type and paratuberculosis in cattle herds. *Am. J. Vet. Res.* **2004**, *65*, 10–14. [CrossRef] [PubMed]
26. Pouillot, R.; Dufour, B.; Durand, B. A deterministic and stochastic simulation model for intra-herd paratuberculosis transmission. *Vet. Res.* **2004**, *35*, 53–68. [CrossRef] [PubMed]
27. Collins, M.T.; Morgan, I.R. Epidemiologic Model of Paratuberculosis in Dairy-Cattle. *Prev. Vet. Med.* **1991**, *11*, 131–146. [CrossRef]
28. Lu, Z.; Schukken, Y.H.; Smith, R.L.; Grohn, Y.T. Stochastic simulations of a multi-group compartmental model for Johne’s disease on US dairy herds with test-based culling intervention. *J. Theor. Biol.* **2010**, *264*, 1190–1201. [CrossRef] [PubMed]
29. Wolf, R.; Barkema, H.W.; De, B.J.; Orsel, K. Dairy farms testing positive for *Mycobacterium avium* ssp. *paratuberculosis* have poorer hygiene practices and are less cautious when purchasing cattle than test-negative herds. *J. Dairy Sci.* **2016**, *99*, 4526–4536. [PubMed]
30. Barrett, D.J.; Mee, J.F.; Mullaney, P.; Good, M.; McGrath, G.; Clegg, T.; More, S.J. Risk factors associated with Johne’s disease test status in dairy herds in Ireland. *Vet. Rec.* **2011**, *168*, 410. [CrossRef]
31. Muskens, J.; Elbers, A.R.; van Weering, H.J.; Noordhuizen, J.P. Herd management practices associated with paratuberculosis seroprevalence in Dutch dairy herds. *J. Vet. Med. B Infect. Dis. Vet. Public Health* **2003**, *50*, 372–377. [CrossRef]
32. Jakobsen, M.B.; Alban, L.; Nielsen, S.S. A cross-sectional study of paratuberculosis in 1155 Danish dairy cows. *Prev. Vet. Med.* **2000**, *46*, 15–27. [CrossRef]
33. Tavornpanich, S.; Johnson, W.O.; Anderson, R.J.; Gardner, I.A. Herd characteristics and management practices associated with seroprevalence of *Mycobacterium avium* subsp *paratuberculosis* infection in dairy herds. *Am. J. Vet. Res.* **2008**, *69*, 904–911. [CrossRef]
34. Hirst, H.L.; Garry, F.B.; Morley, P.S.; Salman, M.D.; Dinsmore, R.P.; Wagner, B.A.; McSweeney, K.D.; Goodell, G.M. Seroprevalence of *Mycobacterium avium* subsp *paratuberculosis* infection among dairy cows in Colorado and herd-level risk factors for seropositivity. *J. Am. Vet. Med. Assoc.* **2004**, *225*, 97–101. [CrossRef]
35. Wells, S.J.; Wagner, B.A. Herd-level risk factors for infection with *Mycobacterium paratuberculosis* in US dairies and association between familiarity of the herd manager with the disease or prior diagnosis of the disease in that herd and use of preventive measures. *J. Am. Vet. Med. Assoc.* **2000**, *216*, 1450–1457. [CrossRef] [PubMed]
36. Wolf, R.; Barkema, H.W.; De, B.J.; Orsel, K. Sampling location, herd size, and season influence *Mycobacterium avium* ssp. *paratuberculosis* environmental culture results. *J. Dairy Sci.* **2015**, *98*, 275–287. [PubMed]
37. Hoe, F.G.; Ruegg, P.L. Opinions and practices of wisconsin dairy producers about biosecurity and animal well-being. *J. Dairy Sci.* **2006**, *89*, 2297–2308. [CrossRef]
38. van Weering, H.J.; van Schaik, G.; van der Meulen, A.; Waal, M.; Franken, P.; van Maanen, K. Diagnostic performance of the Pourquier ELISA for detection of antibodies against *Mycobacterium avium* subspecies *paratuberculosis* in individual milk and bulk milk samples of dairy herds. *Vet. Microbiol.* **2007**, *125*, 49–58. [CrossRef] [PubMed]
39. Nielsen, S.S.; Toft, N. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-gamma assay and faecal culture techniques. *Vet. Microbiol.* **2008**, *129*, 217–235. [CrossRef]
40. More, S.J.; Sergeant, E.S.; Strain, S.; Cashman, W.; Kenny, K.; Graham, D. The effect of alternative testing strategies and bio-exclusion practices on Johne’s disease risk in test-negative herds. *J. Dairy Sci.* **2013**, *96*, 1581–1590. [CrossRef]
41. Kennedy, D.J.; Benedictus, G.; Nielsen, S.S.; Lybeck, K.; Schwan, E.; Frossling, J.; Sergeant, E.; Kelton, D.; Nauholz, H. *Guidelines for Certification with Respect to the Movement of Livestock for Mycobacterium avium subsp. paratuberculosis (MAP) Infection*; International Association for Paratuberculosis, 2017; pp. 1–16. Available online: [https://www.paratuberculosis.net/documents/IAP\\_Guidelines\\_for\\_MAP\\_certification\\_for\\_livestock\\_FINAL.pdf](https://www.paratuberculosis.net/documents/IAP_Guidelines_for_MAP_certification_for_livestock_FINAL.pdf) (accessed on 28 December 2021).

42. Meyer, A.; McAloon, C.G.; Tratalos, J.A.; More, S.J.; Citer, L.R.; Graham, D.A.; Sergeant, E.S.G. Modeling of alternative testing strategies to demonstrate freedom from *Mycobacterium avium* ssp. *paratuberculosis* infection in test-negative dairy herds in the Republic of Ireland. *J. Dairy Sci.* **2019**, *102*, 2427–2442. [[CrossRef](#)]
43. Gates, M.C.; Evans, C.A.; Heuer, C.; Voges, H.; Weston, J.F. Temporal trends in bulk tank milk antibody ELISA and PCR test results for bovine viral diarrhoea in New Zealand pastoral dairy herds. *N. Z. Vet. J.* **2021**, *69*, 73–82. [[CrossRef](#)]
44. Stockmarr, A.; Bodker, R.; Nielsen, L.R. Dynamic changes in antibody levels as an early warning of Salmonella Dublin in bovine dairy herds. *J. Dairy Sci.* **2013**, *96*, 7558–7564. [[CrossRef](#)]
45. Mortier, R.A.; Barkema, H.W.; De, B.J. Susceptibility to and diagnosis of *Mycobacterium avium* subspecies *paratuberculosis* infection in dairy calves: A review. *Prev. Vet. Med.* **2015**, *121*, 189–198. [[CrossRef](#)]
46. Weber, M.F.; Kogut, J.; de Bree, J.; Van, S.G.; Nielen, M. Age at which dairy cattle become *Mycobacterium avium* subsp. *paratuberculosis* faecal culture positive. *Prev. Vet. Med.* **2010**, *97*, 29–36. [[CrossRef](#)]
47. Wolf, R.; Orsel, K.; De, B.J.; Barkema, H.W. Calves shedding *Mycobacterium avium* subspecies *paratuberculosis* are common on infected dairy farms. *Vet. Res.* **2015**, *46*, 71. [[CrossRef](#)] [[PubMed](#)]
48. Nielsen, S.S.; Toft, N. Effect of management practices on paratuberculosis prevalence in Danish dairy herds. *J. Dairy Sci.* **2011**, *94*, 1849–1857. [[CrossRef](#)] [[PubMed](#)]
49. Pillars, R.B.; Grooms, D.L.; Woltanski, J.A.; Blair, E. Prevalence of Michigan dairy herds infected with *Mycobacterium avium* subspecies *paratuberculosis* as determined by environmental sampling. *Prev. Vet. Med.* **2009**, *89*, 191–196. [[CrossRef](#)] [[PubMed](#)]
50. Raizman, E.A.; Wells, S.J.; Godden, S.M.; Fetrow, J.; Friendshuh, K.; Michael, O.J. Characterization of Minnesota dairy herds participating in a Johne’s disease control program and evaluation of the program risk assessment tool. *Prev. Vet. Med.* **2006**, *75*, 22–33. [[CrossRef](#)] [[PubMed](#)]
51. Corbett, C.S.; Naqvi, S.A.; De, B.J.; Kanevets, U.; Kastelic, J.P.; Barkema, H.W. Environmental sample characteristics and herd size associated with decreased herd-level prevalence of *Mycobacterium avium* ssp. *paratuberculosis*. *J. Dairy Sci.* **2018**, *101*, 8092–8099. [[CrossRef](#)]
52. Weber, M.F.; van Maanen, C.; von Bannisseht-Wijsmuller, T.; Lam, T.J.G.M. Sensitivity of environmental sampling for paratuberculosis. In Proceedings of the 10th International Colloquium on Paratuberculosis, Minnesota, MN, USA, 9–14 August 2009; pp. 23–24.
53. Lombard, J.E.; Wagner, B.A.; Smith, R.L.; McCluskey, B.J.; Harris, B.N.; Payeur, J.B.; Garry, F.B.; Salman, M.D. Evaluation of environmental sampling and culture to determine *Mycobacterium avium* subspecies *paratuberculosis* distribution and herd infection status on US dairy operations. *J. Dairy Sci.* **2006**, *89*, 4163–4171. [[CrossRef](#)]
54. Lavers, C.J.; McKenna, S.L.; Dohoo, I.R.; Barkema, H.W.; Keefe, G.P. Evaluation of environmental fecal culture for *Mycobacterium avium* subspecies *paratuberculosis* detection in dairy herds and association with apparent within-herd prevalence. *Can. Vet. J.* **2013**, *54*, 1053–1060.
55. Arango-Sabogal, J.C.; Fecteau, G.; Dore, E.; Cote, G.; Roy, J.P.; Wellemans, V.; Buczinski, S. Bayesian accuracy estimates of environmental sampling for determining herd paratuberculosis infection status and its association with the within-herd individual fecal culture prevalence in Quebec dairies. *Prev. Vet. Med.* **2021**, *197*, 105510. [[CrossRef](#)]