Using technology and machine learning to assess health and welfare of calves

Authors: Johanna M.C. Brans, Gemma A. Miller, Marie J. Haskell, Carol-Anne Duthie

November, 2026

Overview

Integrated technologies can be used to improve livestock productivity, health and welfare, making it easier for farmers to track individuals and individual responses. Using daily clinical health assessments, data from a 'smart' calf ear tag combined with environmental and automatic feeder data allowed for the development of an early-warning system for bovine respiratory disease in pre-weaned dairy-bred calves.

Main Findings

This brief examines the application of machine learning and new technology for bovine respiratory disease (BRD) detection in pre-weaned dairy-beef calves. Technologies used were an automatic milk feeder, a pedometer, and a newly developed calf-specific ear tag combined with environmental data.

- Using all data available, different machine learning (ML) techniques were explored to assess accuracy of alerts (alignment between generated alerts and actual disease) and accuracy of disease detection (number of disease events identified correctly): up to 68% of alerts were considered as "true alerts" and up to 62% of disease events were correctly identified.
- Combining outputs from different ML techniques improved the accuracy of disease detection to 79%. The accuracy of alerts remained similar (up to 60%).
- Different combinations of the different data sources were also explored to assess impact on model performance. Models based only on data generated from the automatic milk feeder, delivered the best model performance, with 70% of disease events correctly identified.
- This work demonstrates that early disease detection is possible using integrated data sources, supporting improved disease management, targeted and optimised treatments.

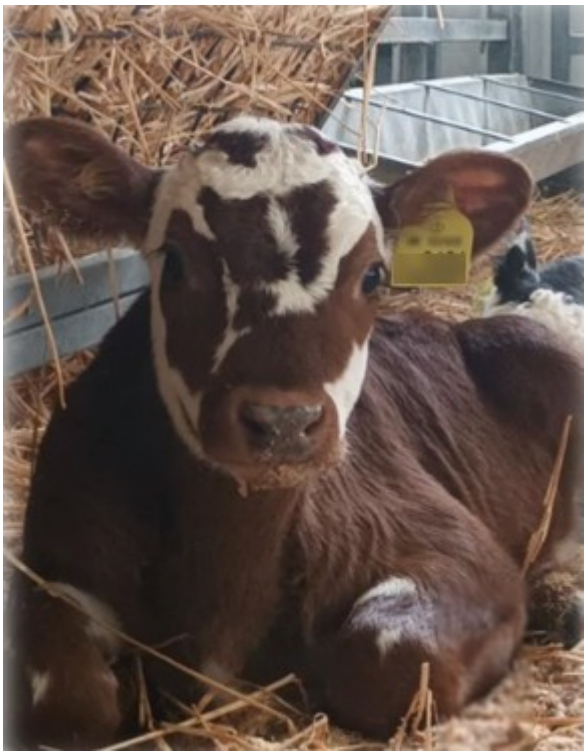


Image credit: SRUC

Introduction

Dairy-bred calves are important for both milk production and dairy-beef systems. One of the most significant health issues affecting these calves globally is bovine respiratory disease (BRD), which is estimated to affect between 30% and 40% of calves in the UK (Agriculture and Horticulture Development Board, 2024). There is currently no specific diagnostic tool for BRD, nor a standard description, as multiple pathogens can be responsible for the respiratory infection (Buczinzki and Pardon, 2020; Kamel et al., 2024; Churchill et al., 2025). The development of wearable sensors for pre-weaned calves (Brans et al., 2026) and routine access to wider on-farm data sources, alongside machine learning techniques could enable development of early health warning systems.

What was the aim of the study?

To assess the performance of different machine learning techniques and models that combined data from different technologies and wider on-farm data sources for detection of BRD in pre-weaned dairy-bred calves.

Methods

Data from a total of 100 calves was used. Individual animal data (i.e., ID, date of birth, birth weight, pen number, age, sex, breed type and daily scorer) alongside data from a calf-specific 'smart' ear-tags (86 variables), automatic milk feeder (28 variables), commercially available pedometer (5 variables) and weather station (24 variables). After removing the highly correlated variables, the dataset contained 100 variables in total (Figure 1).

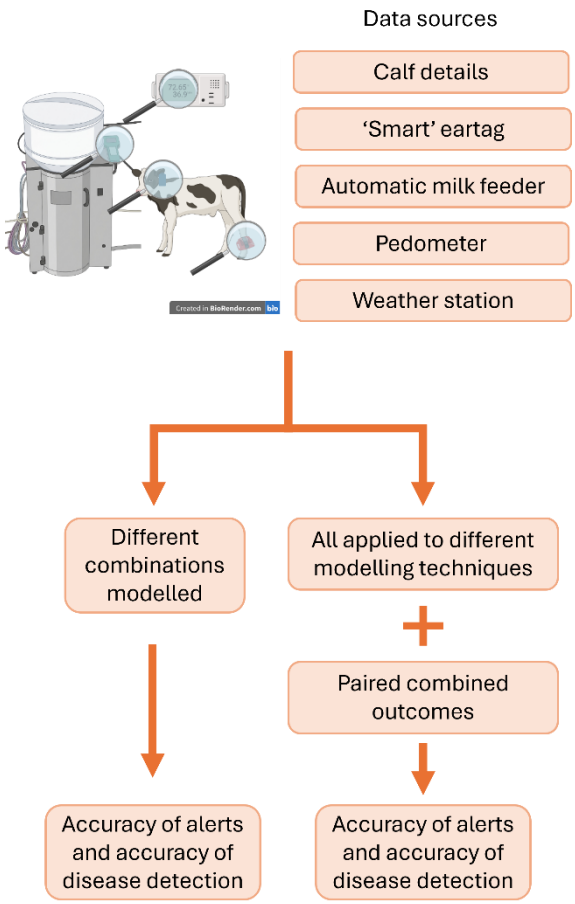


Figure 1: Graphical representation of the methods used in the project

Each calf was given a daily health status, based on its daily monitoring score (Wisconsin score; McGuirk and Peek, 2014) which was used to create two key metrics to assess model performance: the accuracy of alerts (alignment between generated alerts and actual disease – Figure 2) and the accuracy of disease detection (number of disease events identified correctly – Figure 3).

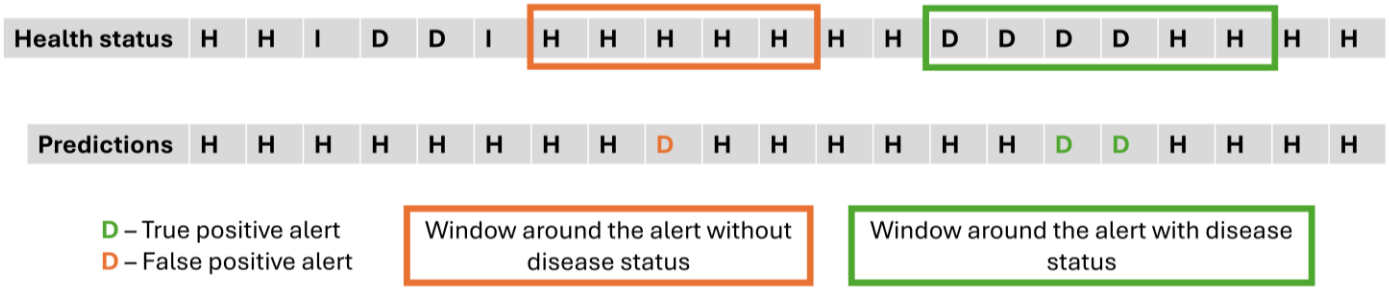


Figure 2: Visual representation of the accuracy of alerts, where each grey box is a day, H represents a 'Healthy' score, I represents the 'Intermediate' score and D represents a 'Disease' score.

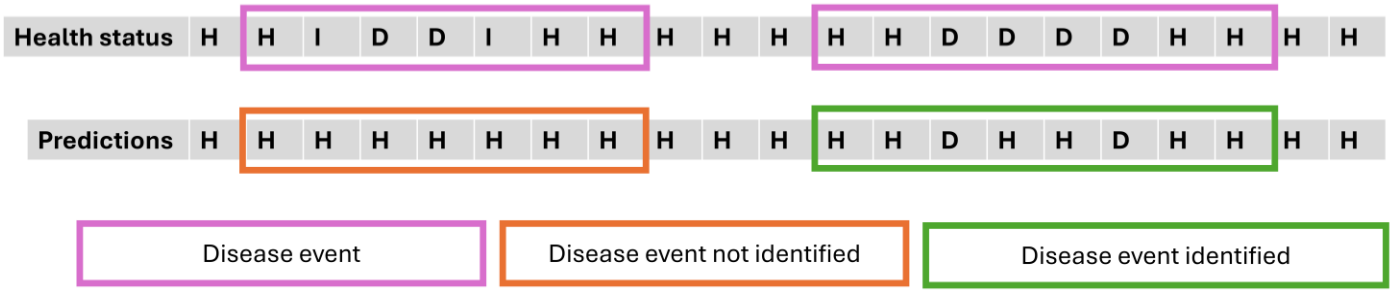


Figure 3: Visual representation of the accuracy of disease detection, where each grey box is a day, H represents a 'Healthy' score, I represents the 'Intermediate' score and D represents a 'Disease' score.

Two analyses were done, the first one looked at four different machine learning modelling techniques, using the 100 variables available. The outputs were paired and the accuracy of alerts and accuracy of disease detection were calculated on single and paired outputs. For the second analysis only one modelling technique was used and variables were combined per technology and type of data (daily data or data averaged over 1 or 2 or 3 preceding days). The models contained between 3 and 29 variables and in total for the second analysis, 40 different models were explored.

Results

Comparison of modelling techniques:

- Single modelling techniques performed better at the accuracy of alerts, whereas paired outputs from modelling techniques performed better at the accuracy of disease detection (Table 1).

Table 1: Accuracy of alerts and accuracy of disease detection ranges for the different modelling techniques

	Accuracy of alerts	Accuracy of disease detection
Single modelling technique	47% to 68%	32% to 62%
Paired outputs from modelling technique	51% to 60%	57% to 79%

Comparison of data sources and assessment of disease detection on data from days prior to disease:

- The best model for accurate alerts was based data from all technologies calculated on day 0 of disease, providing 68% of accurate predictions (Table 2).
- The model which identified the most diseased events (70%) was based on milk feeding behaviour traits calculated on day 2 prior to onset of disease (Table 2).
- The % of correct alerts dropped based on data collected prior to the disease event (d-1, -2 ,-3) but still remained high at 59% on day -3. However, for disease detection the accuracy remained similar at d-3 (Table 2).

Table 2: Accuracy of alerts and accuracy of disease detection ranges for the different combinations of technologies

	Accuracy of alerts	Accuracy of disease detection
Day 0	52% to 68%	31% to 66%
Day -1	47% to 59%	32% to 69%
Day -2	48% to 61%	32% to 70%
Day -3	49% to 59%	32% to 63%

Policy Implications

- The use of new technology combined with machine learning could improve cattle health, farm management and industry efficiency.
- Data integration and machine learning have the potential to quickly and passively identify sick calves, improving treatment response time and calf' outcomes.
- Improved access to on-farm data sources, and integration between different systems will support faster progress in this area.

For more information on this work please contact:
E: johanna.brans@sruc.ac.uk

W: https://sefari.scot/research/projects/data-driven-innovations-for-improved-sustainability-of-ruminant-productions

@SEFARIscot info@sefari.scot

Scottish Government
Riaghaltas na h-Alba
gov.scot

References: AHDB, 2024, https://ahdb.org.uk/knowledge-library/pneumonia-in-calves; Buczinski and Pardon, 2020, https://doi.org/10.1016/j.cvfa.2020.03.004; Kamel et al., 2024, https://doi.org/10.3390/ani14040627; Churchill et al., 2025, https://doi.org/10.3168/jds.2024-25827; Brans et al., 2026, https://doi.org/10.1016/j.livsci.2025.105859; McGuirk and Peek, 2014, https://doi.org/10.1017/S1466252314000267