J. Dairy Sci. 99:1–13
http://dx.doi.org/10.3168/jds.2015-10253

Field validation of protocols developed to evaluate in-line mastitis detection systems

C. Kamphuis,[1,2] B. T. Dela Rue, and C. R. Eastwood
DairyNZ Ltd., Hamilton 3240, New Zealand

ABSTRACT

This paper reports on a field validation of previously developed protocols for evaluating the performance of in-line mastitis-detection systems. The protocols outlined 2 requirements of these systems: (1) to detect cows with clinical mastitis (CM) promptly and accurately to enable timely and appropriate treatment and (2) to identify cows with high somatic cell count (SCC) to manage bulk milk SCC levels. Gold standard measures, evaluation tests, performance measures, and performance targets were proposed. The current study validated the protocols on commercial dairy farms with automated in-line mastitis-detection systems using both electrical conductivity (EC) and SCC sensor systems that both monitor at whole-udder level. The protocol for requirement 1 was applied on 3 commercial farms. For requirement 2, the protocol was applied on 6 farms; 3 of them had low bulk milk SCC (128×10^3cells/mL) and were the same farms as used for field evaluation of requirement 1. Three farms with high bulk milk SCC (270×10^3 cells/mL) were additionally enrolled. The field evaluation methodology and results were presented at a workshop including representation from 7 international suppliers of in-line mastitis-detection systems. Feedback was sought on the acceptance of standardized performance evaluation protocols and recommended refinements to the protocols. Although the methodology for requirement 1 was relatively labor intensive and required organizational skills over an extended period, no major issues were encountered during the field validation of both protocols. The validation, thus, proved the protocols to be practical. Also, no changes to the data collection process were recommended by the technology supplier representatives. However, 4 recommendations were made to refine the protocols: inclusion of an additional analysis that ignores small (low-density) clot observations in the definition of CM, extension of the time window from 4 to 5 milkings for timely alerts for CM, setting a maximum number of 10 milkings for the time window to detect a CM episode, and presentation of sensitivity for a larger range of false alerts per 1,000 milkings replacing minimum performance targets. The recommended refinements are discussed with suggested changes to the original protocols. The information presented is intended to inform further debate toward achieving international agreement on standard protocols to evaluate performance of in-line mastitis-detection systems.
Key words: automated mastitis detection, in-line sensors, performance evaluation, standard protocols

INTRODUCTION

Increasing herd sizes in major dairying countries is driving a need for automation and technologies that help farmers with their daily management decisions (Jago et al., 2013). In-line mastitis-detection systems have been developed to automate some tasks in management of mastitis and bulk milk SCC (**BMSCC**). Kamphuis et al. (2013a) reported the development of protocols to evaluate the performance of these systems against 2 criteria of importance to farmers. The systems need to identify cows with clinical mastitis (**CM**) and identify cows with a high SCC to manage BMSCC. Adoption rates of in-line mastitis-detection systems are reported to be ~26% among surveyed dairy farmers internationally who have also invested in other sensor technologies (Borchers and Bewley, 2015). In the Netherlands, adoption rates of 35% with conventional milking systems and 93% with automatic milking systems have been reported (Steeneveld and Hogeveen, 2015) for mastitis-detection systems based on electrical conductivity (**EC**). Edwards et al. (2015) reported a much lower adoption rate of 6% for in-line mastitis-detection systems in a representative survey of New Zealand dairy farmers. Farmers without an in-line mastitis-detection system ranked such systems in the top 3 of potentially useful technologies for their farm (Borchers and Bewley, 2015) or expressed an interest in investing in them (Edwards et al., 2015). Reasons for dairy farmers not to invest in decision support technolo-

Received August 14, 2015.
Accepted October 19, 2015.
[1]Corresponding author: claudia.kamphuis@wur.nl
[2]Current address: Chair Group Business Economics, Wageningen University, Hollandseweg 1, 6706KN, Wageningen, the Netherlands.

gies include the perception that current commercially available technologies are unproven, unreliable, and have an uncertain return on investment (Russell and Bewley, 2013; Borchers and Bewley, 2015; Steeneveld and Hogeveen, 2015).

Several scientific publications have reported on in-line mastitis-detection technologies, or new detection algorithms, which has been reviewed by Rutten et al. (2013). Most studies reported traditional performance indicators, such as sensitivity (**Se**; the proportion of gold standard-positive cases that are correctly detected as being positive for mastitis) and specificity (**Sp**; the proportion of gold standard-negative cases that are correctly detected as negative for mastitis). However, these performance indicators are difficult to interpret for nonscientific readers (Sherlock et al., 2008). Comparing performance indicators between studies, and thus between technologies and algorithms, is also extremely difficult because studies have used a variety of gold standards to define mastitis (Hogeveen et al., 2010). Claycomb et al. (2009) demonstrated that even small differences in gold standard definitions had a large effect on the number of true cases used for evaluation purposes. Additionally, whereas the gold standards between studies are similar, comparison of results remains difficult due to differences in the inclusion or exclusion of criteria used to create data sets to develop and validate detection models. For example, Kamphuis et al. (2010a) evaluated the performance of a mastitis-detection algorithm using 2 validation sets with differing certainty of mastitis status and found a difference in Se of ~50%. Last, studies use different time windows in which mastitis alerts are considered true positive or false negative. Kamphuis et al. (2010b) demonstrated that increasing the length of the time window significantly affects performance indicators: using a 24-h time window preceding the occurrence of a CM episode resulted in an Se of 40% at an Sp of 99%. Increasing the time window to 96 h preceding the occurrence until 72 h after the occurrence of a CM episode increased Se to 75% at the same Sp level of 99%.

To overcome the difficulties of interpreting performance indicators and the lack of uniform performance information, Kamphuis et al. (2013a) proposed a methodology to field evaluate the performance of in-line mastitis-detection systems, referred to as the protocols in the current paper, with respect to the 2 requirements (finding cows with CM, and finding cows with high SCC to manage BMSCC levels). The protocols are aimed at providing (1) robust and uniform information on performance of current in-line mastitis-detection systems against criteria of importance to farmers to support more informed investment decisions, and (2) a performance evaluation framework to help technology suppliers develop or improve their technologies. The protocols were developed using literature, expert knowledge, and data from previous trials, but were not validated in their final form on commercial farms. This paper reports on the application of the protocols on commercial farms to evaluate their practicality and to refine the methodology using analysis of the field validation data and feedback from technology supplier representatives.

MATERIALS AND METHODS

In our study, the protocols for identifying cows with CM and identifying cows with high SCC to manage BMSCC were validated. As the protocols involve different data collection steps and statistical analysis, the materials and methods are described for each requirement separately.

Requirement 1: Identifying Cows with CM Promptly and Accurately

Data were collected as per the protocol from 3 commercial dairy farms, located in South Canterbury, New Zealand. The farms were selected from a list provided by a supplier of mastitis-detection systems. All farms milked large (>500 cows), spring-calving herds through rotary parlors that had 2 in-line mastitis-detection systems installed. The systems recorded the maximum value of EC (**EC_{max}**) at every stall and SCC at every fourth stall where the SCC sensors were installed. Both sensor systems monitored whole-udder milk of individual cows. All 3 farms had mastitis-detection systems from the same supplier. Farm and data collection details are provided in Table 1. Before the trial started, a supplier representative conducted an inspection to confirm proper functioning of the sensor equipment.

The protocol recommends collection of a minimum of 20 CM episodes per farm. As the incidence of CM is higher during early lactation, data were collected early in the milking season of 2012 and 2013 (Table 1). Data on clot observations for determination of CM episodes were collected as per the protocol for requirement 1. In-line filters with removable stainless steel screens (Vision Mastitis Detector; Ambic Equipment Ltd., Witney, UK) were fitted to the long milk tubes at every stall. These screens were removed and visually inspected by trained staff for the presence of clots (visually assessed as >2 mm^2) after each cow completed milking. On farm 1 clot observations were made by farm staff; external personnel were used on the other 2 farms. Both farm staff and external personnel were trained in the procedure of observing clots and recording data. When clots were identified, the filter screen was removed and replaced with a new one. The removed filter was placed