Journal of Dairy Research (2016) **83** 326–333. © *Proprietors of Journal of Dairy Research 2016* 326
doi:10.1017/S0022029916000406

Shorter sampling periods and accurate estimates of milk volume and components are possible for pasture based dairy herds milked with automated milking systems

Claudia Kamphuis*†, Jennie K Burke, Sarah Taukiri, Susan-Fay Petch and Sally-Anne Turner

DairyNZ Ltd., Private Bag 3221, Hamilton 3240, New Zealand

Received 10 November 2015; accepted for publication 20 June 2016

Dairy cows grazing pasture and milked using automated milking systems (AMS) have lower milking frequencies than indoor fed cows milked using AMS. Therefore, milk recording intervals used for herd testing indoor fed cows may not be suitable for cows on pasture based farms. We hypothesised that accurate standardised 24 h estimates could be determined for AMS herds with milk recording intervals of less than the Gold Standard (48 hs), but that the optimum milk recording interval would depend on the herd average for milking frequency. The Gold Standard protocol was applied on five commercial dairy farms with AMS, between December 2011 and February 2013. From 12 milk recording test periods, involving 2211 cow-test days and 8049 cow milkings, standardised 24 h estimates for milk volume and milk composition were calculated for the Gold Standard protocol and compared with those collected during nine alternative sampling scenarios, including six shorter sampling periods and three in which a fixed number of milk samples per cow were collected. Results infer a 48 h milk recording protocol is unnecessarily long for collecting accurate estimates during milk recording on pasture based AMS farms. Collection of two milk samples only per cow was optimal in terms of high concordance correlation coefficients for milk volume and components and a low proportion of missed cow-test days. Further research is required to determine the effects of diurnal variations in milk composition on standardised 24 h estimates for milk volume and components, before a protocol based on a fixed number of samples could be considered. Based on the results of this study New Zealand have adopted a split protocol for herd testing based on the average milking frequency for the herd (NZ Herd Test Standard 8100:2015).

Keywords: Robotic milking, milk recording, protocol, sampling duration.

Milk recording data are used by breeding companies to identify sires and cows that will contribute to the genetic gain of future generations of dairy cattle. For farms with conventional milking systems, where cows are milked as a batch with more or less fixed milking intervals, protocols to collect and process milk recording data are clearly defined in standards. In contrast, standards describing protocols to collect and handle milk recording data from farms with automated milking systems (**AMS**), where cows are milked 24 h per day, with varying milking intervals between and within cows, often vary between countries (Miglior et al. 2002; Bucek et al. 2014), if available at all. Until recently, New Zealand (**NZ**) did not have a standard for collecting and handling milk recording data on AMS farms. As a consequence, NZ farmers using AMS were unable to submit milk recording data to the national database to be used for estimating genetic merit for milk traits. This resulted in a reduced accuracy of milk volume, fat, protein and somatic cell score genetic merit for their cows, as the estimates are based solely on ancestry data. This limited the opportunities of AMS farmers to make use of milk production and milk quality information for herd management decisions (e.g., culling decisions).

Sampling protocols for herd testing have previously been reported for AMS systems (Peeters & Galesloot, 2002; Hand et al. 2006; Leclerc et al. 2012). In fact, the ICAR protocols have been developed based on studies conducted in AMS systems where cows are housed (ICAR, 2016). While a 16 h sampling period (suggested by Hand et al. 2006) may be appropriate in systems with higher milking frequencies, it may not be suitable for pasture based systems since in a

*For correspondence; e-mail: c.kamphuis@gddiergezondheid.nl
†Current address: GD Animal Health, PO Box 9, 7400 AA Deventer, The Netherlands.

review on AMS systems, Lyons et al. (2014) reported that cows on pasture based AMS farms have lower milking frequencies than indoor fed cows that are milked by AMS. It is, thus, important to understand the impact of the lower milking frequencies on the accuracies of the sampling protocol for herd testing. Additionally, the ICAR protocols are based on studies where multiple exclusion criteria are applied to the datasets. Due to the lower milking frequencies within pasture based AMS systems, and the requirements to minimise the length of the sampling period for cost reasons for farmers, it is important that exclusion criteria are kept to a minimum as more exclusion criteria would result in fewer cows having the required number of samples during the milk recording test period. Further, where a countries herd test regulations require that the compositional analyses are linked to the milk volume at the time of sampling, such as in the NZ herd testing regulations, calculations using a rolling average for milk yield that are proposed by ICAR are not permitted.

In a pastoral based dairy system, Jago & Burke (2013) tested a protocol which involved the collection of recorded milk yields from the AMS management system and automated collection of milk samples from all cows milked by the AMS during a 48 h sampling period. They concluded that a 24 h sampling period provided satisfactorily high correlation with the 48 h protocol, while minimising the number of cows without herd test results. Their conclusion, however, was based on data from a single research farm, on which the milking frequency averaged just 1·3 times per day (Jago et al. 2004). It is anticipated that this milking frequency is lower than that on commercial pastoral based farms milking using AMS, and thus that the sampling period may reduce even further in practice.

In addition, consideration to the cost of analysis of the milk samples is required as a 48 h sampling period, coupled with a higher milking frequency, would generate more samples for analysis. Where herd test regulations require the matching of the milk sample composition to the milk yield at the time of collection, such as in NZ, the additional samples increase costs significantly as farmers pay per sample analysed in a laboratory. A composite sample may be considered as an alternative, but then farmers also pay for the subsampling required to generate a composite sample which is proportionate based on milk yield at the time of collection (i.e., the subsampling for the composite milk sample must be proportional to the milk volume from which the milk sample was collected relative to the total milk yield over the total milk recording interval). Both analysis approaches add considerable costs. Additionally, a 48 h sampling period is time consuming and error-prone (Jago & Burke, 2013), and significantly disrupts daily routines on AMS farms.

Thus, the current study was designed to determine optimum milk recording intervals for AMS within commercial pastoral based dairying systems, where lower herd average milking frequencies are expected than on farms where cows are fed indoors, and where higher milking frequencies are expected than those reported on a pastoral based research farm. It aimed at retaining acceptable accuracy for standardised 24 h estimates for milk volume and composition, while reducing the milk recording test period from a 48 h milk recording protocol thereby reducing costs and improving practicality.

Materials and methods

Data collection

A detailed description of the data collection process has been described previously (Kamphuis et al. 2015). Briefly, data were collected on five farms with AMS located across NZ between December 2011 and February 2013. The selected farms represented a range of pastoral based farm systems of varying herd sizes and breeds (Jersey, Friesian and Friesian- Jersey cross bred animals), and included the two main AMS suppliers in NZ (DeLaval International AB, Tumba, Sweden and Lely Industries NV, Maassluis, The Netherlands). Before the start of each test period, it was confirmed by either the AMS supplier or the farmer that the AMS units and milk sampling devices were installed and conformed to operational specifications.

Each test period was conducted according to the 48 h protocol previously used by Jago & Burke (2013). One test period, thus, involved continuous data collection from the AMS management system, including identification numbers of the AMS unit, the automated milk sampling device, and the cow, the date and time of each cow milking, date and time of the previous milking of that same cow, milking interval as the time difference between current and previous milking, milk yield as recorded by milk meters installed on each AMS unit, and whether or not the current milking had been completed according to AMS software. Milk samples from each cow milking during the 48 h were collected automatically by the milk sampling devices. A unique identification number, identifying the order in which the 25 ml milk samples were collected over the 48 h period, enabled the milk composition data to be matched with the milk yield data from the AMS management system. Milk samples of known milk composition (QA samples) were included in the sequence to enable confirmation of the milk composition data sequence provided by the laboratories. The milk samples were refrigerated at 4 °C until they were analysed for milk composition (fat, crude protein, and lactose yield and per cent, and SCC) by a certified laboratory (Testlink Laboratories, in Hamilton, NZ for milk samples collected in the North Island, and in Christchurch, NZ for milk samples collected in the South Island).

A total of 13 milk recording test periods were initiated, ranging between one to five test periods per farm (Table 2). One test period (on farm 5) was cancelled before the completion of the 48 h period, due to obvious malfunctioning of the milk sampling device during the first evening sampling. During the remaining 12 test periods, a total of 11 127 milk samples from 2879 cow-test days were collected and submitted for analyses. However, in the final data analysis, the milk composition data from three test periods were also excluded from the milk composition analysis because of milk analysis errors identified by the non-matching of the QA samples with the known milk composition. Milk yield data from these three test period were still included in the analysis for 24 yield estimates for milk volume.

Statistical analyses

The procedure for estimating standardised 24 h milk, fat, crude protein and lactose yields was based on the description provided by Jago & Burke (2013). From the start of a 48 h sampling period, for each milking for each cow, the interval from the previous milking was calculated and recorded as milking interval in days. For each cow, for each milking with test period data during this 48 h sampling period, yields for fat, crude protein and lactose were calculated using the milk composition results from the laboratory and the corresponding milk yields as recorded by the AMS management system. One of the two AMS suppliers recorded milk yield in litres and these were multiplied by 1·03 to convert to milk yield in kg. Total milk yield, fat yield, protein yield and lactose yield and milking interval were then calculated for this 48 sampling period from these milkings. The 24 h standardised yields were then calculated by dividing total milk, fat, crude protein, and lactose yields by the total milking interval. Standardised 24 h estimates for fat, crude protein and lactose per cent were calculated from these standardised 24 h yields. An example of the procedure followed is provided in Table 1, where standardised 24 h estimates for milk, fat and crude protein yield and fat and crude protein per cent are calculated for one cow with three complete milkings within a 48 h sampling period. A similar procedure was used for 24 h standardised estimates for SCC.

As many cow milkings per test period as possible were included in the estimation of the standardised 24 h yields and per cent. However, the first cow milking after the start of the 48 h sampling period was excluded if the previous milking of that cow (outside the sampling period) was labelled 'incomplete' by the AMS management software. Additionally the last cow milking within the sampling period was excluded if labelled as incomplete. Other incomplete cow milkings during the sampling period were only excluded when milk recording data (yield and/or milk composition data) were missing; otherwise they were included for further analyses. For the 48 h sampling period, this resulted in 10 048 'successful' milk samples (90·3% from 11 127 milk samples) from 2879 cow-test days (Table 2) that were used for further analyses.

Standardised 24 h estimates for milk, fat, crude protein and lactose yields and per cent and SCC from the 48 h sampling period (Jago & Burke, 2013) were used as reference values (Gold Standard), and compared with values generated from nine alternative sampling scenarios. These

Table 1. Example of estimating standardised 24 h yield for milk volume, fat and crude protein for one cow (cow ID 1001) milked by an automated milking system and where milk samples are automatically collected for a 48 h sampling period

Cow ID	Milking	Date and Time of milking	Milking interval recorded by management system in hours (*in days*)	Milk volume recorded by in-line milk meter (kg)	Laboratory results		Calculated 24 h estimates yields	
					Milk fat %	Milk protein %	Milk fat	Milk protein
1001	−1	21/02/2013 9:24						
Start of a 48 h sampling period at 12:00 noon								
1001	1	21/02/2013 22:56	13·53 (*0·5639*)	18·8	5·3	3·8	0·996 (18·8 × 5·3%)	0·71 (18·8 × 3·8%)
1001	2	22/02/2013 16:13	17·28 (*0·7201*)	29·1	5·1	3·7	1·48 (29·1 × 5·1%)	1·08 (29·1 × 3·7%)
1001	3	23/02/2013 11:08	18·92 (*0·7882*)	34·2	5·4	3·8	1·85 (34·2 × 5·4%)	1·30 (34·2 × 3·8%)
End of a 48 h sampling period at 12:00 noon								
Total			49·73 (*2·0722*)	82·1			4·3272	3·0907
Standardised 24 h yield and composition (calculated)				39·62 (82·1/2·0722)	5·27 (2·088/39·62)	3·76 (1·491/39·62)	2·088 (4·327/2·0722)	1·491 (3·091/2·0722)

Table 2. Number of cow-test days, number of successful cow milkings within each test period and milking interval when sampling for 48 h on five farms that milk automatically

Farm	Test period	Cow-test days (*n*)	Cow milkings (*n*)	Milking interval (48 h)
1	1	171	557	15·6
	2	106	322	16·9
	3	154	514	14·9
	4	156	536	14·4
	5	172	393	15·4
2	6	296	1069	13·6
	7	311	945	15·7
	8	333	981†	15·4
	9	318	1143	12·9
3	10	194	581†	10·8
4	11	527	2570	10·9
5	12	141	437†	16·4
	13‡			
	Total	2879	10 048	
	Average			14·4

The table includes all the successful cow-test days submitted for milk composition analysis, before 1999 cow milkings were identified as invalid due to milk sampling and analytical errors in test periods 8, 10 and 12
†Excluded from the final milk composition analysis. Data is included in the milk volume analysis
‡Test period cancelled due to an obvious malfunction of the automatic milk sampling device during the first evening of the 48 h sampling period

included six shorter sampling periods (8, 12, 16, 18, 24, and 36 h), each commencing from the start of the Gold Standard 48 h period, or three scenarios where a fixed number of visits of the cow to the AMS after the start of the 48 h sampling period (one, two or three visits) was used. Standardised 24 h estimates were determined for each scenario, as well as the proportion of cow-test days without any 24 h standardised estimates (missed cow-test days), expressed as a percentage of the total cow-test days/test period per farm.

In this study, Lin's concordance correlation coefficients, which account for both accuracy and precision of paired readings (Lin, 1989), was used to evaluate the level of agreement between the Gold Standard and the alternative sampling scenarios. As with the more familiar Pearson's correlation coefficient, concordance correlation coefficients ranged between −1 to 1, with a value of 1 representing perfect agreement. Lin's concordance correlation coefficients were calculated using GenStat (VSN International, 2013). The absolute bias and standard deviation, and bias and standard deviations as a percentage of the Gold Standard were also calculated to facilitate comparisons with ICAR (2016) accuracy limits.

Results

The average milking interval for the 12 completed test periods was 14·4 h, ranging between 10·8 and 16·9 h (Table 2). Milking intervals varied between farms, but also within farm depending on timing of the test period within

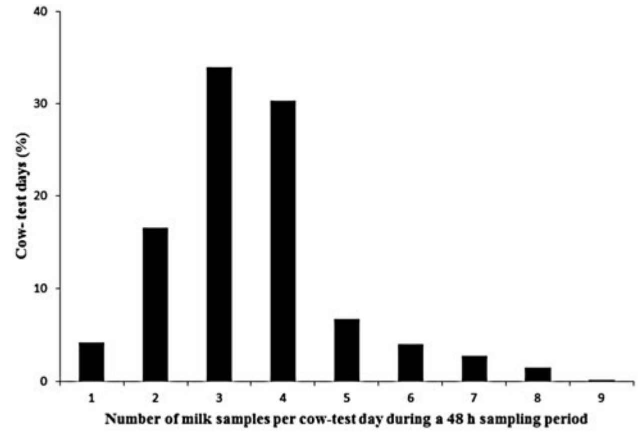

Fig. 1. Number of milk samples collected per cow-test day, using a 48 h sampling period, on five commercial farms milking cows with automated milking systems.

the milking season (data not shown). The number of successful milk samples per cow-test day varied between 1 to 9 for a 48 h sampling period (Fig. 1). The percentage of cows achieving two, three or between four and nine successful milk samples in the sampling period were 16·6%, 34·0% and 45·2%.

Significant issues with the automated collection of milk samples were identified with one test period (farm 5). Moreover, in three out of the 12 remaining test periods (test periods 8, 10 and 12) there were strong suggestions that results from the milk composition analyses returned from the laboratory didn't match to the correct milk sample. Results for milk components from these three test periods were deemed invalid and excluded from the analyses for standardised 24 h estimates for milk composition only. Thus, a total of 10 048 successful cow milkings with milk yield data from 2879 cow-test days were included for the analysis on milk volume (Table 3), and 8049 successful cow milkings from 2211 cow-test days with both milk yield and milk composition results were available for the analysis on milk composition (Table 4).

There was high agreement between standardised 24 h estimates for milk volume for the alternative sampling scenarios compared to the Gold Standard (concordance correlation coefficient on average ≥0·96, bias as a per cent of the Gold Standard on average <0·9), even when the sampling period was reduced to 8 h or when only one milk sample was used (Table 3). The standard deviation of the differences, as a per cent of the Gold Standard ranged from 7·28 (for the 8 h sampling period) to 1·95 when three milk samples were used, while the bias (as a per cent of the Gold Standard) was greatest for the 8 h sampling period (0·83) and lowest when three milk samples were included (0·02). Similar results were observed for standardised 24 h estimates of crude protein yield and per cent, and lactose yield and per cent, with average concordance correlation coefficients ≥0·93 (Table 4). The absolute bias and standard

https://doi.org/10.1017/S0022029916000406 Published online by Cambridge University Press

Table 3. Number of cow-test days, average proportion of missed cow-test days (%, calculated by averaging the proportion of the total cow-test days/test period per farm), average concordance correlation coefficient, bias (as a % of the Gold Standard), and standard deviation of the difference (as a % of the Gold Standard) between predicted standardised 24 h yield for milk volume estimated using the Gold Standard (48 h sampling period) and alternative test day scenarios (shorter sampling periods or fixed number of milkings/cow per test day) when milking cows with automated milking systems

Test day scenario	Cow-test days (n)	Average proportion of missed cow-test days (range)	Concordance correlation coefficient (range)	Bias %† (range)	SD Dif %‡ (range)
Sampling period (h)					
48	2879	ref	ref		ref
36	2854	0·6 (0–2·8)	0·99 (0·99–1·0)	−0·23 (−0·82 to 0·77)	2·86 (1·88–4·46)
24	2798	2·7 (0–9·9)	0·98 (0·97–0·99)	−0·14 (−2·87 to 1·66)	4·77 (3·46–6·48)
18	2670	7 (1·3–13·8)	0·97 (0·95–0·99)	−0·18 (−3·20 to 2·35)	6·13 (4·89–7·48)
16	2644	8 (3·8–4·7)	0·97 (0·95–0·99)	−0·16 (−3·24 to 2·59)	6·30 (4·96–7·54)
12	2398	16·3 (7·2–29·1)	0·96 (0·95–0·98)	−0·38 (−2·88 to 2·48)	6·72 (5·81–7·95)
8	1771	38·2 (21·7–54·1)	0·96 (0·94–0·98)	−0·83 (−3·40 to 2·90)	7·28 (6·01–8·61)
Fixed number of milk samples					
1	2718	4·8 (0·9–9·5)	0·97 (0·95–0·98)	−0·29 (−2·30 to 2·16)	6·77 (5·83–7·88)
2	2876	0·15 (0–0·7)	0·99 (0·98–0·99)	0·09 (−1·60 to 1·13)	4·13 (2·76–5·72)
3	2879	0 (0–0)	1·00 (0·99–1·0)	−0·02 (−0·43 to 0·36)	1·95 (0·94–3·47)

†Bias. Mean difference between the Gold Standard (48 h sampling scenario) and alternative sampling scenarios as a percentage of the Gold Standard.
‡SD Dif %. Standard deviation of the difference between the Gold Standard (48 h sampling scenario) and alternative sampling scenarios as a percentage of the Gold Standard

deviation of the differences for crude protein yield and per cent and for lactose yield and per cent were all on average below 0·05 (bias) and 0·10 (standard deviation). Concordance correlation coefficients were lower for SCC and averaged below 0·90 for the 8 h sampling period (Table 4), while the bias was greater than 0·05 for the 8 and 16 h sampling period, and when two milk samples were used. In addition, the standard deviation only fell below 0·10 for the 36 h sampling period, and when three milk samples were used. Lowest values were observed for fat yield and per cent. Concordance correlation coefficients averaged below 0·90 for an 18 h or shorter sampling period for fat yield, and below 0·90 for a 24 h period or shorter sampling period for fat per cent. When sampling was limited to one milk sample only, concordance correlation coefficients averaged below 0·9 for fat yield, and below 0·8 for fat per cent (Table 4). The standard deviations for all sampling period scenarios were all greater than 0·01 for fat per cent, while only the 36 and 24 h sampling period, and when either two or three milk samples were used were below 0·01 for fat yield. When sampling was limited to one milk sample only, or when the sampling periods were 12 or 8 h long, the bias was greater than 0·05 for fat per cent. This was in contrast to fat yield, where the bias for all sampling scenarios, were less than 0·05.

There was an increase in proportion of cow-test days without 24 h standardised estimates for milk volume (Table 3), as the sampling period decreased, or as the number of fixed milk samples reduced. In the alternative scenario of an 8 h sampling period, on average 38·2% of cow-test days had no 24 h standardised estimate of milk volume (Table 3). In this scenario, the proportion of missed cows could be as high as 54%, depending on the average milking frequency of the herd (Table 3). The proportion (and range) of missed cow-test days were similar between the 24 h standardised estimates for milk volume and milk components for the different alternative sampling periods. Comparable results were also found when collecting a fixed number of milk samples. Where the proportion of missed cow-test days were, on average, 4·8, 0·15 and 0% when collecting one, two, or three milk samples for 24 h standardised estimates for milk volume (Table 3), proportions were 4·6, 0·14, and 0%, respectively, for 24 h standardised estimates for milk components.

Discussion

The current study applied a 48 h milk recording protocol across 13 milk recording test periods on each of five commercial AMS farms to determine the number of milk samples collected per cow under field conditions, and to evaluate the effect of reducing the proposed sampling time on standardised 24 h yield estimates for milk volume and composition.

The average milking interval of 14·4 h in our current study was similar to the milking intervals for pasture based AMS systems, albeit higher than that on a pastoral based research farm (Jago & Burke, 2013), and higher than those reported for indoor-based AMS by Lyons et al. (2014). This implied that the AMS farms in the current study had a representative range for milking intervals occurring in commercial practice. Milking intervals, however, did vary considerably between farms in the current study and affected the number of successful milkings per cow-test day; the shorter the milking interval, the higher the proportion of cows with more than three successful milk samples per test period when sampling for 48 h.

Table 4. Number of cow-test days, and average concordance correlation coefficient between predicted standardised 24 h yield for milk component yields and concentrations, and somatic cell count, estimated using the Gold Standard (48 h sampling period) and alternative test day scenarios (shorter sampling periods or fixed number of milkings/cow per test period) when milking cows with automated milking systems

Test day scenario	Cow-test days (n)	Concordance correlation coefficient (range)						
		Fat yield	Fat per cent	Crude protein yield	Crude protein per cent	Lactose yield	Lactose per cent	Somatic Cell Count
Sampling period (h)								
48	2211	ref	ref	ref	ref	ref	ref	ref
36	2193	0·97 (0·96–0·99)	0·95 (0·93–0·96)	0·99 (0·99–1·0)	0·99 (0·98–1·00)	0·99 (0·99–1·00)	0·99 (0·98–1·0)	0·98 (0·96–0·99)
24	2159	0·93 (0·87–0·97)	0·86 (0·79–0·94)	0·98 (0·97–0·99)	0·98 (0·97–0·99)	0·98 (0·97–0·99)	0·97 (0·96–0·98)	0·94 (0·90–0·96)
18	2068	0·87 (0·80–0·93)	0·77 (0·69–0·90)	0·96 (0·95–0·98)	0·97 (0·94–0·99)	0·97 (0·95–0·99)	0·95 (0·92–0·98)	0·92 (0·88–0·95)
16	2051	0·86 (0·80–0·93)	0·76 (0·69–0·89)	0·96 (0·95–0·98)	0·97 (0·94–0·99)	0·97 (0·95–0·99)	0·94 (0·92–0·98)	0·92 (0·86–0·94)
12	1874	0·83 (0·74–0·91)	0·72 (0·65–0·87)	0·96 (0·94–0·98)	0·96 (0·93–0·98)	0·96 (0·95–0·98)	0·94 (0·92–0·97)	0·90 (0·86–0·94)
8	1385	0·81 (0·73–0·90)	0·70 (0·60–0·86)	0·95 (0·93–0·97)	0·96 (0·92–0·98)	0·96 (0·94–0·98)	0·93 (0·90–0·97)	0·89 (0·83–0·93)
Fixed number of milk samples								
1	2087	0·83 (0·76–0·90)	0·73 (0·65–0·87)	0·96 (0·95–0·97)	0·96 (0·93–0·98)	0·97 (0·95–0·98)	0·93 (0·91–0·97)	0·91 (0·86–0·94)
2	2209	0·95 (0·91–0·98)	0·92 (0·86–0·95)	0·98 (0·97–0·99)	0·99 (0·98–1·0)	0·99 (0·98–0·99)	0·98 (0·97–0·99)	0·96 (0·94–0·98)
3	2211	0·99 (0·98–1·0)	0·98 (0·97–1·0)	1·0 (0·99–1·0)	1·0 (0·99–1·0)	1·0 (1·0–1·0)	1·0 (0·99–1·00)	0·99 (0·98–1·0)

On average, ~45% of the cows had more than three successful milkings, but this proportion was much higher for farm 4 (74·6%; results not shown), which also had one of the shortest milking intervals.

Jago & Burke (2013) indicate that there is an increased risk of human errors with longer sampling periods on farms with AMS since milk samples have to be transferred manually to storage facilities up to three times a day due to the limited capacity of automated milk sampling devices. In our current study, an excess of 2000 cow milkings were lost due to milk sampling device malfunctions or sampling and analysis errors. Moreover, our current study found that ~45% of the cow-test days had more than three successful milkings during 48 h of sampling. These factors emphasise that 48 h of sampling for commercial AMS farms increases the risk of errors (mechanical and human errors) and the costs associated with herd testing, and that a shorter sampling period (hours) would be beneficial. Based on the correlation coefficients found by Jago & Burke (2013), they suggest reducing the sampling time to 24 h, although they find the 36 h sampling period to be optimal in terms of accuracy, they show only a small loss in accuracy results when using a 24 h sampling period. Results from the current study confirmed their suggestion with the average concordance correlation coefficients remaining >0·90 for standardised 24 h yield estimates for milk volume and all milk components. The only exception is fat per cent, where the current study reported an average concordance correlation coefficient of 0·86 when sampling for 24 h.

The current study applied specific exclusion rules to estimate standardised 24 h yields for milk volume and composition. These exclusion rules only concerned those milkings that were labelled as 'incomplete' by the AMS software. It was acknowledged that different AMS systems and (software) versions within AMS brand may apply different rules to define 'incomplete' milkings. The decision to exclude specific 'incomplete' milkings may have resulted in higher concordance correlation coefficients. However, no further exclusion criteria were applied and, as a consequence >90% of the data were used for further analyses. This is in contrast to previous studies, e.g. Peeters & Galesloot (2002) where several criteria are used to exclude entire cow-test days. As a result, their analyses are based on ~46% of the original dataset and this increases the likelihood that the accuracy of their estimated 24 h yields is overestimated. Moreover, in the current study we did not adjust for covariates nor apply requirements on minimum milking intervals. Studies by Hand et al. (2006) and Leclerc et al. (2012) report improved accuracies for milk recording protocols that adjust for covariates compared to protocols without these adjustments. Additionally, Peeters & Galesloot (2002) include the requirement that milking intervals should be >4 h. Leclerc et al. (2012) report reduced accuracies for 24 h estimates for milk components at lower milking intervals (<6 h). These criteria (including covariates and setting limits on milking interval) were not applied in our current study and may have resulted in lower concordance correlation coefficients. Future research should include the effect of applying these criteria on the accuracies of estimates and concordance correlation coefficients.

Concordance correlation coefficients reported in the current study inferred that the sampling period could be reduced even further to 8 h or be limited to the collection of one milk sample per cow without losing too much accuracy in the 24 h standardised milk yields for milk volume, crude protein yield, crude protein per cent, lactose yield, and lactose per cent. This is in line with general agreement across published studies that there is no real concern regarding accurate estimates for crude protein, regardless of the sampling protocol applied (e.g. Buenger et al. 2002; Hand et al. 2006; Leclerc et al. 2012). As with our current study, previous studies (Buenger et al. 2002; Hand et al. 2006; Leclerc et al. 2012) identified that the definition of a suitable (that is, practical and acceptable accuracy) milk recording protocol will largely depend on the accuracy of estimates of fat yield and/or fat per cent. Leclerc et al. (2012) suggest a milk recording protocol resulting in a correlation of ~0·85 for fat per cent to be sufficiently accurate. This is lower than the 0·938 concordance correlation coefficients for fat per cent recommended by Hand et al. (2006), using a 16 h sampling period protocol. With the 16 h sampling period used in the current study, average concordance correlation coefficients were above 0·92 for milk volume and the milk components, protein, lactose and SCC, but those for fat were limiting with average concordance correlation coefficients of 0·86 for fat yield and 0·72 for fat per cent. Moreover, for a 16 h test period, on average >8% of the cow-test days would not have milk recording results. In addition, although the bias for the 16 h sampling period met the ICAR (2016) bias limit of <0·05%, it did not meet the limits for the standard deviation of <0·1%. These results indicate that a 16 h protocol is unlikely to provide the required accuracies within a low milking frequency environment.

Using a milk recording protocol in which a fixed number of milk samples are collected could solve the issue of insufficient accuracies for 24 h estimates of fat yield and per cent, and the risk of missing cow-test days, or cows with missing composition data. Peeters & Galesloot (2002) conclude that a 24 h fat per cent can be estimated from one milk sample with adequate accuracy and precision. This contradicts with the results we found here, where concordance correlation coefficients dropped below 0·85 for fat yield and fat per cent in the alternative scenario where only one milk sample per cow is collected. However, results of our study demonstrated that average concordance correlation coefficients remained >0·90 for milk yield and all milk components, including fat yield and fat per cent, when a fixed number of two milk samples per cow were collected. Moreover, the average proportion of missed cow-test days was very low (<0·5%), with a maximum of missed cow-test days of 0·65% (result not shown). When applying this alternative sampling scenario (a fixed number of samples), future research should study whether the standardised 24 h scenario needs adjustment for diurnal variations in milk composition (Bouloc et al. 2002) in case the time interval between two milk samples is too short. Further research is required to determine the potential impacts of diurnal variations and milking frequency on milk composition and therefore the accuracy of the standardised 24 h estimates for milk volume and components, before a protocol based on a fixed number of samples could be considered.

The data generated in this study and conclusions made were considered by NZ Animal Evaluation Ltd. who have adopted a split protocol for herd testing on AMS farms within NZ (NZ Herd Test Standard 8100:2015). The split herd testing protocol, is based on the average milking frequency for the herd and uses a longer sampling period (36 h) for farmers with a farm system that generates lower herd average milking frequencies (<2 times per day) than for farmers with a farm system that generates higher herd average milking frequencies (16 h; ≥2 times per day). While the data inferred that a protocol using a fixed number of samples could also have been recommended, there is still uncertainty about the impact of diurnal variation on milk composition on the wider range of farm systems.

In conclusion, implementing a 48 h milk recording protocol was found to be inappropriate for commercial AMS farms, due to the greater number of milk samples collected, leading to increased analytical costs, increased likelihood of milk sampling device malfunctions or sampling and analytical errors. In addition, a 48 h milk recording causes considerable and, based on the results of this study, unnecessary disruption to milking routines. Alternative sampling periods were evaluated for the reduction in accuracy and increased proportion of missing cow-test days. Based on the results of this study a 48 h milk recording protocol is unnecessarily long for collecting accurate yield estimates during milk recording on AMS farms. Collecting two milk samples only per cow was optimal, however, future research is required to determine the impact of diurnal variations in milk composition and including covariates and setting limits on milking interval on standardised 24 h estimates for milk volume and components, to further improve the standardisation calculations, especially for herds milked using AMS with low milking frequencies.

The authors acknowledge Barbara Dow (DairyNZ Ltd, Hamilton, NZ) and Harold Henderson (AgResearch, Hamilton, NZ) for their statistical support and the participating farmers for contributing their data. This study was funded by NZ Government through the Primary Growth Partnership research programme and by NZ dairy farmers through DairyNZ Inc.

References

Bucek P, Zottl K, Onken F, Klopčic M, Radzio D, Mészarós G, Barac Z, Ryba Š, Dianová M & Kučera J 2014 Practical aspects in milk recording in Central and Eastern Europe and its effect on the guidelines. Manuscript available at: http://www.icar.org/index.php/icar-meetings-news/berlin-2014-presentations-pictures-and-inewsletter/ Accessed on June 27th 2016

Buenger A, Pasman T, Bohlsen E & Reinhardt F 2002 Transformation of AMS records to 24 hour equivalents. In *Performance Recording of Animals – State of the Art 2002*, pp. 87–96 (Eds J Crettenand, J Moll, C Mosconi & S. Wegmann). Interlaken, Switzerland. EAAP publication no. 107. Wageningen, the Netherlands: Wageningen Academic Publishers

Bouloc N, Dervishi V, & Delacroix J 2002 Milk recording and automatic milking systems: simplification by reducing the daily time test from 24 to 12 hours. In *Performance Recording of Animals – State of the Art 2002*, pp. 57–63. (Eds J Crettenand, J Moll, C Mosconi & S Wegmann). Interlaken, Switzerland. EAAP publication no. 107. Wageningen, the Netherlands: Wageningen Academic Publishers

Hand KJ, Lazenby D, Miglior F & Kelton DF 2006 Short communication: Comparison of protocols to estimate twenty-four-hour fat and protein percentages for herds with a robotic milking system. *Journal of Dairy Science* **89** 1723–1726

ICAR (International Committee for Animal Recording) 2016 Pages 400–401 in ICAR rules, standards, and recommondentations for testing. Section 11- Testing, approval, and checking of milk recording devices. Manuscript available at www.icar.org Accessed on June 27th 2016

Jago JG & Burke JK 2013 Estimating 24-hour milk, fat, and protein yields and somatic cell count for automatically milked cows in pastoral production systems. *Proceedings of the New Zealand Society of Animal Production* **73** 29–32

Jago JG, Bright K, Copeman P, Davis K, Jackson AK, Ohnstad I, Wieliczko R & Woolford MW 2004 Remote automatic selection of cows for milking in a pasture-based automatic milking system. *Proceedings of the New Zealand Society of Animal Production* **64** 241–245

Kamphuis C, Dela Rue B, Turner S-A & Petch S-F 2015 Devices used by automated milking systems are similarly accurate in estimating milk yield and in collecting a representative milk sample compared with devices used by farms with conventional milk recording. *Journal of Dairy Science* **98** 3541–3557

Leclerc H, Huquet B, Minery S, Bourrigan X, Thomas G & Saunier D 2012 Analysis of the accuracy of protocols in robotic milking herds for estimating 24-hour fat and protein (yields and percentage). Session T4 in Proceedings of the 38rd Annual ICAR Meeting, Cork, Ireland. Manuscript available on http://www.icar.org/Cork_2012/Manuscripts/Published/Bourrigan.pdf. Accessed on August 8, 2014

Lin LI 1989 A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **45** 255–268

Lyons NA, Kerrisk KL, Dhand NK, Scott VE & Garcia SC 2014 Animal behavior and pasture depletion in a pasture-based automatic milking system. *Animal* **8** 1506–1515

Miglior F, de Roos S, Liu Z, Mathevon M, Rosati A, Schaeffer LR & VanRaden P 2002 Report of the ICAR Working Group on lactation calculation methods. Review and update of guidelines for milk recording. In *Proceedings of the 33rd Biennial Session of ICAR*, pp. 227–231

New Zealand Herd Test Standard 2015. Dairy Herd Testing (NZS 8100:2015)

Peeters R & Galesloot PJB 2002 Estimating daily fat yield from a single milking on test day for herds with a robotic milking system. *Journal of Dairy Science* **85** 682–688

VSN International 2013 *GenStat for Windows* 16th edition. Hemel Hempstead, UK: VSN International. http://www.GenStat.co.uk