Do-it-yourself milk recording as a viable alternative to supervised milk recording in Ireland

Martin Burke[‡], *Donagh Berry*[†] , *Mary O'Keeffee**, *& Pat O'Connor**

[‡] *The Irish Cattle Breeding Federation*

[†] *Moorepark Research Centre*

* *Dairygold AI & Farm Services*

EXCECUTIVE SUMMARY

- Evaluation of the data suggests that the do-it-yourself (DIY) method adopted on the participating farms measure milk production variables in accordance with expectations

- Accurate prediction of 24-hour milk, fat and protein yield is achievable using either AM or PM samples. Prediction of daily somatic cell count (SCC) is less accurate although the sensitivity and specificity of predicted daily SCC at identifying actual daily SCC >200,00 is high

- Accuracy of predicting 24-hour fat and protein yield can be augmented by recording two consecutive milk weights and only one milk composition. This also facilitates the accurate recording of individual cow-testday milking interval. No effect of including two milk weights in the prediction model for daily SCC was observed

- AM/PM SCC alone is as good if not a better indicator of daily SCC than predicted daily SCC using prediction equations

- Milking interval defined as individual cow-testday interval measured in minutes fitted the data better than individual cow-testday interval rounded to the nearest half-hour, which was in turn superior to average herd-testday interval and average herd interval.

- Results from this study suggest DIY milk recording is a viable alternative to supervised milk recording

- The effect of using predicted daily yields from DIY milk recording on the accuracy of prediction of 305-day yields should be investigated. Research should also be undertaken on possible differential weights applied to predicted daily yield in a test-day model.

1. Introduction

The level of milk recording in Ireland is low relative to most other major milk producing countries (ICAR, 2002), thereby hindering genetic progress within the national dairy herd. Possible reasons for this low participation in milk recording in Ireland include the cost and inconvenience of the current official milk recording service.

Alternate AM-PM milk recording schemes (Porzio, 1953) are officially approved by the International Committee for Animal Recording (ICAR) and are implemented in a number of countries including Germany, Austria, France, The Czech Republic, Croatia and Italy (ICAR, 2002). For a twice a day milking regime, the differences between the morning (AM) milk yield and the evening (PM) milk yield is primarily a function of milking interval, stage of lactation and also perhaps a significant interaction between milking interval and stage of lactation (Everett and Wadell, 1970). Therefore, any prediction equation used to estimate 24-hour yield from either AM samples or PM samples should incorporate milking interval and lactation stage within the analysis.

Schaeffer ***et al.***, (2000) observed that 24-hour fat yield may be predicted with an accuracy of 0.89 from an AM-milking and 0.88 from a PM-milking. Their model adjusted for the herd average time interval between AM and PM milkings and estimated the prediction equations within subclasses of days in milk, parity and season of calving. Berry et al. (2004a) using data from research herds in Southern Ireland concluded that 24-hour yield and subsequent 305-day yield can be accurately predicted from alternate AM-PM milk recording schemes.

Preliminary analysis (Berry et al., 2004b) revealed that the previously derived prediction equations from research data (Berry et al., 2004a) unsatisfactorily predicted 24-hour yields in a data set of DIY milk records when the time interval between AM and PM milking was >10 hours; the range in milking interval used to predict 24-hour yield in the research data was between 6 and 9 hours. The unsatisfactory fit of the equations to the DIY data were due mainly to the uncertainty of extrapolation in milking interval beyond which data was originally available to derive the equations. Preliminary analyses (Berry et al., 2004b) along with other international studies (Schaeffer et al., 2001) revealed a poorer fit of equations when applied to data independent to that used in deriving the prediction equations.

2. Objective

The objective of the present study was to derive prediction equations using a subset of data collected to-date from the pilot DIY study co-ordinated by ICBF and Dairygold. The prediction equations will be validated in the remaining subset of the data from the DIY pilot study.

3. Materials and Methods

3.1 Raw data

Milk weights and milk samples were collected across 23 DIY herds in Southern Ireland for both AM and PM at roughly four-weekly intervals. Milk composition and somatic cell count (SCC) was determined in Dairygold laboratories and the resulting data, merged to the respective milk weights were uploaded onto the ICBF database. A data set of participating DIY herds was extracted on the 12th August 2004. The initial, unedited data set consisted 7,945 part-day observations from 1,581 cows across 23 herds. In total, 68 herd-testdays were included in the data set; the number of part-day observations per herd-testday varied from 10 to 576. Part-day milk yield varied from 0.8 kg to 46.7 kg. Part-day fat, protein and lactose percentage varied from 1.5-7.58%, 2.19-5.12%, and 3.31-5.44%, respectively. Somatic cell count varied from 7 to 9928 cells/ml; the SCC of nine records was 9999 which is the maximum cell count possible in IRIS. In total 118 composition results were missing while 111 SCC results were missing. Missing analyses were observed for 35 herd-testdays across 19 herds; the number of samples missing per herd-testday varied from 1 to 9 with one herd-testday having 38 samples missing.

Days in milk varied from 1 to 535. Parity varied from 1 to 23. Records were available for cows calving across all months of the year.

3.2 Data editing

Only consecutive non-zero PM-AM (or AM-PM) samples per cow-testday were retained. Parities greater than two were grouped together. Days in milk were grouped into seven classes, each of 50 days interval from 0 to 300 and a final class for >300 days. Following editing 3,850 records (includes both AM and PM records as one record) from 1,565 cows across 23 herds on 68 different herd-testdays were available for inclusion in the analysis. Milking interval was defined as the difference, in minutes, between morning and evening milking for each cow-testday. Alternative definitions were identified whereby milking interval was defined as individual cow testday milking interval rounded to the nearest half hour, average herd-testday milking interval, and average herd milking interval.

3.3 Data analysis

The procedures adopted in the present study were similar to those reported by Berry et al., (2004a). A linear multiple regression model, fitted in SAS (SAS, 2004) to predict actual 24-hour yield from AM and/or PM samples was as follows:

$$Y_{ijk} = [b_0 + b_1(MI) + b_2(Milk)_i + b_3\,(Fat)_i + b_4\,(Protein)_i\,]_k + e_{ijk}$$

Where:

Y_{ijk} = 24-hour yield (milk, fat or protein yield)

MI = herd milking interval from AM to PM (milking interval from PM-AM is directly related so was not included in the model)

$(Milk)_i$ = milk yield on the i^{th} milking of the day

$(Fat)_i$ = fat yield on the i^{th} milking of the day

$(Protein)_i$ = protein yield on the i^{th} milking of the day

e_{ijk} = random residual effect

An additional independent predictor variable of SCC on the i^{th} milking was included in the prediction model for SCC only. Regression analyses were carried out within subclasses *k* to account for the heterogeneous means and variances of the different subclasses. In total 21 subclasses were created based on stage of lactation (0-49, 50-99, 100-149, 150-199, 200-249, 259-299, ≥300) by parity (1, 2, ≥3). The prediction equations were initially derived from 75% of the data randomly chosen from the data set using PROC SURVEYSELECT (SAS, 2004). The number of records per subclass varied from 18 to 411 in the sub-dataset; the average number of records per subclass was 138.

3.4 Tests for comparing alternative milk recording schemes

Prediction equations derived from 75% of the data were applied to the remaining 25% of the data. The comparison between predicted and actual 24-hour yield involved estimating the bias between the different measures (i.e., the average difference between the actual yield and predicted yield) and the variance of the difference between the measures (mean square error). The average bias was computed as the mean of the difference following subtraction of actual 24-hour yield from predicted 24-hour yield.

Correlations between actual and predicted yields were estimated and the 25% and 75% quartiles were also estimated. Correlation analyses were also used to evaluate the independence of the residuals whereby a correlation of zero indicates total randomness of the error.

The accuracy of predicting 24-hour yield was also investigated as:

$$\text{Accuracy} = (\sigma^2\text{actual} / (\sigma^2\text{actual} + \sigma^2\text{difference}))$$

where:

σ^2actual = variance of the actual yield,

σ^2difference = variance of the difference between the actual yield and the predicted yield

Prediction of actual SCC *per se* may not be as important as identifying samples of high (i.e., >200,000 SCC/ml) somatic cell count. Therefore, the sensitivity and specificity of predicted 24-hour SCC at identifying samples with a true daily SCC >200,000 cells/ml was investigated within the data. Sensitivity was calculated as the proportion of daily SCC >200,000 that had an predicted SCC >200,000, and specificity was calculated as the proportion of daily SCC <200,000 that had an predicted SCC of <200,000 cells/ml. Precision was defined as the proportion of correct (positives or negatives) diagnoses made from the AM/PM samples from all samples tested.

4. Results and Discussion

4.1 Data

The mean and standard deviation for milk yield, and composition in the edited data set are summarised in Table 1. In total, 1,031 records (27% of the data set) had a daily SCC >200,000. The ratio of AM SCC to PM SCC varied from 0.012 to 181.25 thereby indicating considerable diurnal variation. The ratios of AM to PM milk yield varied from 0.21 to 10.7; the average was 1.4. The large ratio (10.7) was an outlier and occurred when one cow produced 16 kg milk in the morning and 1.5 kg milk in the evening; compositions were similar for the AM and PM samples. It appears that the 16-kg milk weight was manually inputted. Correlations between AM yield, PM yield and true daily yield were similar to those previously reported from research data (Berry *et al.,* 2004a). Thus, the results are in line with expectations suggesting accurate recording and milk sampling using DIY procedures.

Table 1. *Average and standard deviation (SD) or milk production variables in the edited data set.*

		Milk (kg)	Fat (%)	Protein (%)	SCC (cells/ml)
AM	Average	15.5	3.25	3.32	258
	SD	4.97	0.65	0.31	699
PM	Average	11.5	4.05	3.37	354
	SD	4.13	0.75	0.32	841
Daily	Average	26.9	3.59	3.34	296
	SD	8.59	0.57	0.30	714

Correlations between AM SCC, PM SCC and true daily SCC are summarised in Table 2. The correlations in Table 2 indicate that AM SCC or PM SCC closely resemble daily SCC. The Spearman correlations reflect the similarly in how cows rank for SCC based on either AM, PM or daily SCC; this criteria will be most influential in culling decisions rather than SCC level *per se*. The results demonstrate, that even without prediction equations either AM or PM SCC ranks cows very similar to if both AM and PM samples were analysed separately and weighted by their respective milk yield.

The sensitivity, specificity, and precision of the AM sample as an indicator of daily SCC >200,000 was 82%, 99% and 95% respectively; the corresponding variables for the PM sample were 95%, 93% and 95%. The accuracy of predicting daily SCC from only using AM SCC or PM SCC (i.e., no prediction equations used) was 0.92 and 0.85, respectively. This substantiates previous remarks that even without prediction equations either AM or PM SCC are a good indicator of daily SCC.

Table 2. *Pearson (above diagonal) and Spearman (below diagonal) correlations between AM SCC, PM SCC and true daily SCC.*

	AM	PM	DAY
AM		0.80	0.96
PM	0.86		0.94
DAY	0.96	0.95	

4.2 Prediction of daily yields with either an AM or PM sample

The ability of the prediction equations to estimate daily milk, fat, and protein yield from either AM or PM samples is summarised in Table 3. There was a tendency for daily yield predicted from the AM sample to be an underestimate of true daily yield; the opposite was true for daily yield predicted from PM samples. This may have implications for prediction of 305-day yield and may suggest the superiority of an alternate AM-PM scheme. Berry et al. (2004a) reported that an alternate AM-PM milk record schemes predicted actual 305-day yield more accurately than a recording scheme based on all AM or all PM records.

Table 3. *Accuracy, correlation between predicted daily yields and actual daily yields, mean square error (MSE), mean, 25% percentile (Q1), and 75% percentile (Q3) of the residuals from predicting 24-hour milk, fat and protein yield from either an AM or PM sample in an independent data set.*

		Accuracy	Correlation	MSE	Mean	Q1	Q3
AM	Milk	0.93	0.96	5.88	-0.16	1.20	-1.48
	Fat	0.84	0.90	0.018	-0.010	0.069	-0.077
	Protein	0.92	0.95	0.006	-0.005	0.038	-0.047
	SCC	0.91	0.95	84747	-5.2	46.5	-19.2
PM	Milk	0.88	0.93	9.98	0.056	1.69	-1.52
	Fat	0.83	0.87	0.019	0.005	0.085	-0.071
	Protein	0.87	0.92	0.011	0.002	0.055	-0.050
	SCC	0.88	0.93	115257	-3.14	41.3	-22.1

Examination of the quartiles reveals that 50% of the predicted milk yields were within ±1.7 kg of the actual milk yield; this represents an error of 6% of the mean.

The sensitivity, specificity, and precision of daily SCC predicted from AM samples was 96%, 92% and 94% respectively; the corresponding variables for the PM sample were 93%, 92% and 93%. The high mean square error associated with SCC was attributable mainly to a few larger individual SCC. When AM SCC was restricted to be less than 9,999,999 cells/ml the mean square error was reduced by 10,000; the mean square error halved when AM SCC >5 million were removed from the analysis.

Accuracy of predicting 24-hour SCC was reduced through the use of the prediction equations compared to using the AM sample itself; however, the sensitivity was increased through the use of the prediction equations. Accuracy of predicting 24-hour SCC was increased through the use of the prediction equations compared to using the PM sample alone; however, the sensitivity, specificity and precision was reduced through the use of the prediction equations. Similarly, the correlations between predicted daily SCC and actual daily SCC were lower than correlations between AM/PM SCC and daily SCC.

Absolute correlations between residuals and predicted yields were less than 0.10 for milk, fat and protein yield; the majority were not significantly ($P>0.05$) different from zero. This suggests randomness of the error. The residual correlation for SCC varied from 0.10 to 0.14 and were significantly different from zero; however, such correlations were strongly influenced by testday records with exceptionally high SCC.

Graphical examination of the relationship between the residual and month of calving revealed no trend. Correlations between the residuals with month of calving were generally not significantly different from zero indicating no systematic linear bias across months of calving.

4.3 Prediction of daily yield using two consecutive milk weights but only one composition

A supplementary scenario was investigated whereby two milk weights (both AM and PM) were available but with either an AM or PM milk composition. The accuracy of prediction of daily milk, fat, and protein yield from two consecutive milk weights and either an AM or PM composition is summarised in Table 4. Daily milk yield was not predicted in this scenario as the true measure would be physically recorded.

In agreement with Schaeffer *et al.* (2000) the accuracy of predicting 24-hour fat and protein yield increased when both milk weights were available for inclusion in the prediction equation; the mean square error of the variance also decreased across both studies. The accuracy of predicting 24-hour protein yield was 1.00 and 0.99 from AM or PM samples, respectively; Schaeffer *et al.* (2000) also reported an accuracy of 0.995 and 0.99, respectively. The accuracy of predicting 24-hour fat yield from AM (0.93) or PM (0.90) samples was in agreement with Schaeffer *et al.* (2000) who reported accuracies of 0.93 and 0.92, respectively.

Table 4. *Accuracy, correlation between predicted yields and actual yields, mean square error (MSE), mean, 25% percentile (Q1), and 75% percentile (Q3) of the residuals from predicting 24-hour fat and protein yield from both AM and PM milk weights but either an AM or PM composition sample.*

		Accuracy	Correlation	MSE	Mean	Q1	Q3
AM	Fat	0.93	0.96	0.007	-0.004	0.044	-0.046
	Protein	1.00	1.00	0.000	0.001	0.010	-0.009
	SCC	0.91	0.95	83181	-5.76	48	-19
PM	Fat	0.90	0.95	0.010	0.003	0.058	-0.054
	Protein	0.99	1.00	0.001	0.001	0.013	-0.012
	SCC	0.88	0.93	115180	-3	46	-26

From the derived daily fat and protein yields, fat and protein percentages were calculated and compared to actual daily fat and protein percentages. For daily fat percentage 50% of the predicted records were within ±0.2% of actual records; the corresponding figure was ±0.04% for protein percentage.

The rank correlation within test-day (with a minimum of 20 cows per test-day) between fat yield predicted from AM or PM samples and true fat yield varied from 0.83 to 0.97 and from 0.83 to 0.96, respectively; 14 herd-test days were included in the analysis. The rank correlation within test-day (with a minimum of 20 cows per test-day) between daily protein yield predicted from AM or PM samples and true protein yield varied from 0.98 to 1.00 and from 0.97 to 0.99, respectively. Therefore, the ranking of cows for predicted daily fat and protein yield on any given test day was very similar to cow rankings based on actual 24-hour fat and protein yields.

The sensitivity, specificity, and precision of identifying daily SCC >200,000 using an AM sample incorporated within a prediction equation (with two milk weights) was 95%, 91% and 93% respectively; the corresponding variables for the PM sample were 94%, 92% and 93%. This was very similar to previous results where only one milk weight was included in the prediction equation.

However, farmers may only be interested in identifying the cows with the highest SCC. The rank correlation within the 14 herd-testdays between predicted 24-hour SCC from AM samples and actual 24-hour SCC varied from 0.83 to 0.97. The rank correlation within herd-testdays between predicted 24-hour SCC from PM samples and actual 24-hour SCC varied from 0.76 to 0.97.

Absolute correlations between residuals and predicted fat and protein yield were not significantly (P>0.05) different from zero indicating total randomness of the error. Residual correlations for SCC were 0.10 and 0.14 but were strongly influenced by test day records with exceptionally high SCC.

Graphical examination of the relationship between the residuals and month of calving revealed no trend. Correlations between the residuals with month of calving were generally not significantly different from zero indicating no systematic linear bias across months of calving.

4.4 Definition of milking interval

Four alternative definitions of milking interval were created: cow-testday milking interval in minutes; cow-testday milking interval rounded to the nearest half hour; average herd-testday milking interval, and; average herd milking interval. The effect of the definition of milking interval on the mean square error of prediction for daily yield using either an AM or PM sample, or with two milk weights and one composition sample is illustrated in Figure 1 and 2, respectively.

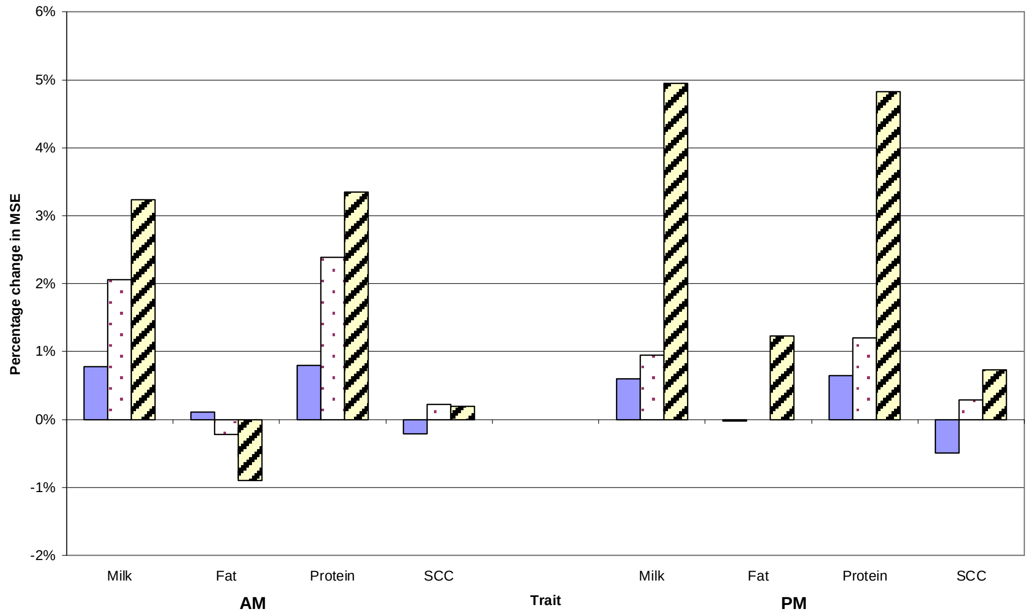

Figure 1. *Percentage change in mean square error (MSE) when milking interval is defined as cow-testday interval rounded to the nearest half hour (Full bars); average herd-testday (spotted bars), or average herd (striped bars) compared to cow-testday interval measured in minutes using either an AM or a PM sample*.

Indications are that the most accurate (i.e., in terms of lowest mean square error) definition of milking interval is individual cow-testday interval measured in minutes. The mean square error is increased when cow-testday interval is rounded to the nearest half-hour unit, and is further increased when milking interval is defined as average herd-testday and average herd milking interval. The effect of alternative milking interval definitions diminishes when two milk weights are included in the prediction model; this is partly attributable to the reduced importance of milking interval when two milk weights are included in the model.

Nevertheless, the impact of alternative definitions of milking interval thereby suggesting that if one sample was taken the milking interval could be supplied by the farmer. The milking interval on the testday when the record was taken would be superior.

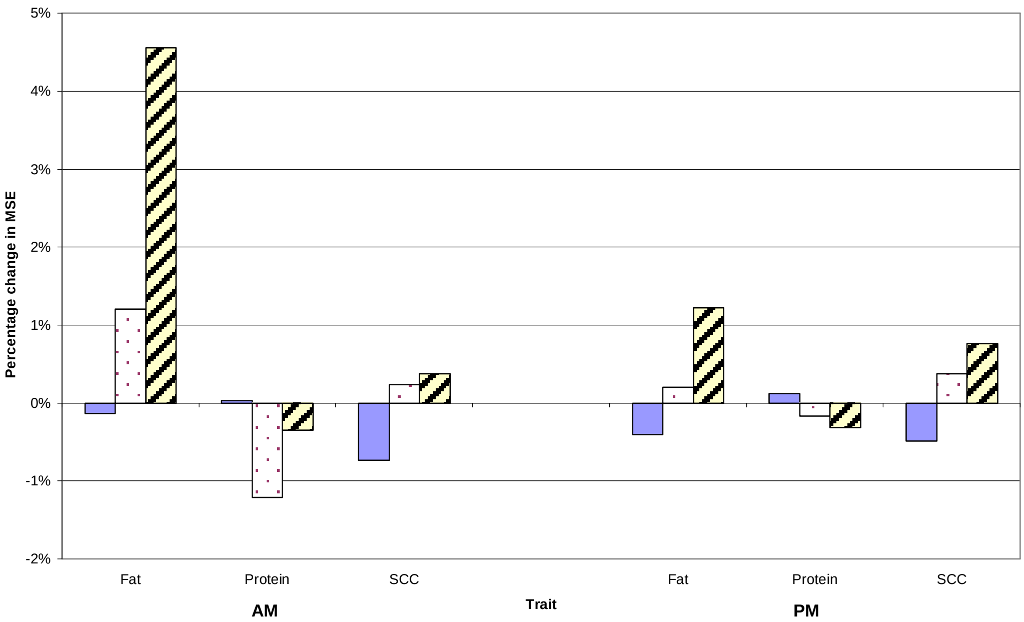

Figure 2. *Percentage change in mean square error (MSE) when milking interval is defined as cow-testday interval rounded to the nearest half hour (Full bars); average herd-testday (spotted bars), or average herd (striped bars) compared to cow-testday interval measured in minutes using consecutive milk weights and either an AM or a PM composition.*

4.5 Prediction equations for national use

Given these favourable results reported herein, prediction equations were re-derived using the whole (100%) data set. Correlations between previously derived solutions (75% of the data) and solutions from the whole data set were all greater than 0.98 for fat and protein yield. The correlation between previous and current solutions for AM SCC was 0.62; the corresponding correlation for PM SCC was 0.49. The correlations became stronger when the data was restricted to include only SCC records below a pre-defined threshold. This suggests that the prediction equations for SCC are not robust and question their usefulness in predicting daily SCC since AM/PM compositions themselves are a good indicator of daily SCC. However, the reduced accuracy of predicting SCC is not surprising given the considerable variation in SCC between two consecutive records.

To investigate this further average lactation SCC was calculated for true daily SCC, AM SCC alone (i.e., no use of prediction equations), PM SCC alone, daily SCC predicted using the prediction equations incorporating an AM sample, and daily SCC predicted using the prediction equations incorporating a PM sample. Pearson correlations between average lactation SCC (cows with at least 4 tests) derived from AM SCC alone, PM SCC alone, daily SCC predicted using the prediction equations incorporating an AM sample, and daily SCC predicted using the prediction equations incorporating a PM sample with true daily SCC were 0.99, 0.98, 0.99 and 0.98 respectively; the corresponding Spearman rank correlations were 0.94, 0.92, 0.96 and 0.94. Pearson correlations between average lactation SCC (cows with at least 1 test) derived from AM SCC alone, PM SCC alone, daily SCC predicted using the prediction equations incorporating an AM sample, and daily SCC predicted using the prediction equations incorporating a PM sample with daily SCC were 0.98, 0.97, 0.98 and 0.97 respectively; the corresponding Spearman rank correlations were 0.96, 0.94, 0.96 and 0.95. Thus, little benefit exists in predicting daily SCC using prediction equations.

The use of prediction equations for fat and protein yield is simple and can easily be incorporated within the ICBF database. The procedures utilised to predict daily yield from part-day samples is summarised in Appendix 1.

5. Conclusions

The results clearly show that daily yield can be accurately predicted from either an AM or PM sample. Because of the relatively larger diurnal variation expected in SCC compared to the other variables the accuracy of prediction of daily SCC was lower. Nevertheless, the accuracy of predicting daily yields can be augmented by the inclusion of two milk weights in the prediction model; little benefit was observed with two milk weights included to predict daily SCC. It should be borne in mind that no study has evaluated the current “dipper method” used in supervised milk recording in Ireland which may not necessarily be 100% accurate.

6. Acknowledgements

The co-operation of all farmers participating in the study is gratefully acknowledged as is assistance of Dairygold in the recording and analysis of the milk samples.

7. References

Berry, D. P., V. E. Olori, A.R. Cromie, R. F. Veerkamp, M. Rath, and P. Dillon. 2004a. Accuracy of predicting milk yield from alternative milk recording schemes. Animal Science (In Press).

Berry, D.P., M. Burke, M. O'Keeffe, and P. O'Connor. 2004b. Validation of equations for predicting daily milk yield and composition from part day samples. Report to ICBF. May 2004.

Everett, R. W. and Wadell, L. H. 1970. Sources of variation affecting ratio factors for estimating total daily milk yield from individual milkings. Journal of Dairy Science 53:1430-1435.

International Committee for Animal Recording. 2002. Position of milk recording; methods and organisation: 2000. www.icar.org/yearly.html. Visited 9th December 2002.

Porzio, G. 1953. A new method of milk recording. Animal Breeding Abstracts 21:1659.

SAS® User's Guide. Version 8.0.0 Edition 2002. SAS Inst., Inc., Cary, Nc.

Schaeffer, L. R., Jamrozik, J., Van Dorp, R., Kelton, D. F., and Lazenby, D. W. (2000). Estimating daily yields of cows from different milking schemes. Livestock Production Science 65:219-227.

Appendix 1.

Prediction equations were derived for each subclass of parity (1, 2, ≥3) and days in milk (0-49, 50-99, 100-149, 150-199, 200-249, 259-299, ≥300). For example, the prediction equation of a first lactation cow with a milk test day at 83 days in milk is:

$$FAT_{DAY} = 0.226686148 + (-0.019860586*INT_{PM-AM}) + (-0.02133069*MILK_{AM}) + (0.037295198*MILK_{PM}) + (1.199585129*FAT_{AM}) + (0.40690312*PROTEIN_{AM})$$

$$FAT_{DAY} = -0.085085796 + (0.013486432*INT_{PM-AM}) + (0.0294658*MILK_{AM}) + (-0.014821905*MILK_{PM}) + (1.324716931*FAT_{PM}) + (0.035625646*PROTEIN_{PM})$$

A cow was present on a farm with a milking interval of 8.28 hours between AM and PM milking. The cow had an AM milk yield, fat percent and protein percent of 8.7 kg, 3.13% fat, 3.23% protein, respectively, and a PM milk yield, fat percent and protein percent of 4.4 kg, 4.44% and 3.12%, respectively. Therefore this cow produced a fat yield of 0.468 kg. Predicted fat yield from the AM sample and PM sample were 0.482 kg and 0.481 kg, respectively.