Early prediction of lactational milk, fat and protein using daily milk data

Oded Nir (Markusfeld), Gil Katz, Liron Reuveni

The primer:

J. I. Weller and E. Ezra, Genetic and phenotypic analysis of daily Israeli Holstein milk, fat, and protein production as determined by a real-time milk analyzer, JDC, Vol. 99 No. 12, 2016

- ✓ **Scope: >37,000 Holstein cows spanning over 2 years**
- ✓ **Finds agreement between Afimilk's inline milk lab real time analysis and between DHIA monthly tests.**
- ✓ **Selected for 'Editor's Choice‘ of JDSc**

Slides are by the courtesy of Ephraim Ezra

Heritabilities, genetic and environmental correlations among 7,866 first parity 305 d lactations computed from the ICBA and AfiLab records.

Trait	Heritabilities		Correlations	
	ICBA	AfiLab	genetic	environmental
Milk (kg)	**0.33**	0.35	**1.00**	0.96
Fat (kg)	**0.23**	0.31	**0.59**	0.70
Protein (kg)	**0.27**	0.32	**0.86**	0.87
% fat	**0.48**	0.57	**0.70**	0.66
% protein	**0.55**	0.46	**0.56**	0.52

Heritabilities were higher for the AfiLab records for all traits, except for % protein.

Phenotypic correlations among complete and extended 1st parity lactations the last ICBA test day and the last two weeks of AfiLab records.

FAT (kg)

Trait	Mean days in milk at truncation								
	30	60	90	120	150	180	210	240	270
ICBA	0.67	0.75	0.79	0.87	0.91	0.93	0.95	0.95	0.96
Afilab	0.77	0.84	0.89	0.92	0.94	0.95	0.96	0.96	0.97

PROTEIN (kg)

Trait	Mean days in milk at truncation								
	30	60	90	120	150	180	210	240	270
ICBA	0.70	0.76	0.78	0.87	0.90	0.92	0.94	0.94	0.95
Afilab	0.72	0.83	0.87	0.90	0.93	0.94	0.95	0.95	0.96

Prediction of complete lactations in Afifarm

- Our objective: To adapt the **large scale retrospective** study's method to a **prospective prediction** of complete (305_days) lactations in individual herds
  - ✓ For selection
  - ✓ For production planning (quota, summer/winter)
- The operational need: To enable farmers to get the decision as early as possible, **but before breeding**

Waiting Periods

Herds	Cows/herd	Voluntary waiting period (days)	Days to 1st AI
13,885	158.4 ± 325 SD	58.4 ± 5.6 SD	95.2 ± 26.9 SD

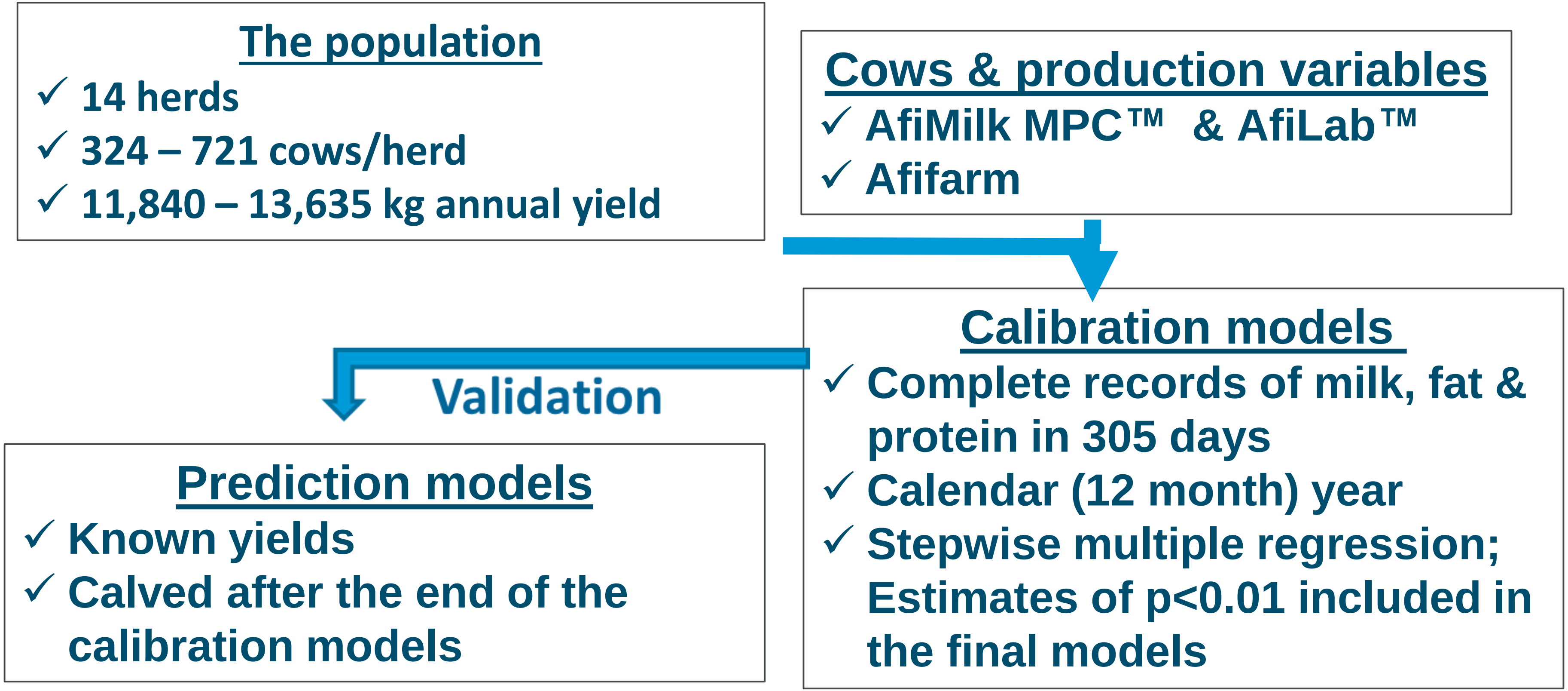
Ferguson J.D. & Skidmore A. (2013). JDS 96 (2) 1269 -1289

Days to 1st AI	50	51 - 80	81 - 110	111 - 150
1st lactation	0.4%	41.4%	45.2%	13.0%
2nd lactation	9.7%	58.4%	26.9%	5.1%

Ezra E. (2013). HerdBook Summary (Hebrew). ICBA

Our objective is to be able to make the decision at 60 DIM

Material & methods



The population

- ✓ 14 herds
- ✓ 324 – 721 cows/herd
- ✓ 11,840 – 13,635 kg annual yield

Cows & production variables

- ✓ AfiMilk MPC™ & AfiLab™
- ✓ Afifarm

Calibration models

- ✓ Complete records of milk, fat & protein in 305 days
- ✓ Calendar (12 month) year
- ✓ Stepwise multiple regression; Estimates of $p<0.01$ included in the final models

Validation

Prediction models

- ✓ Known yields
- ✓ Calved after the end of the calibration models

Criteria for Success

- R^2= RSquare of the summary of fit
- r = Correlations to actual production
- The gap (%) between the predicted & the actual production (APD)
- Rate of wrong positive cows (1 – specificity) for selection (cows culled by mistake)

Prediction after 54 DIM 14 herds

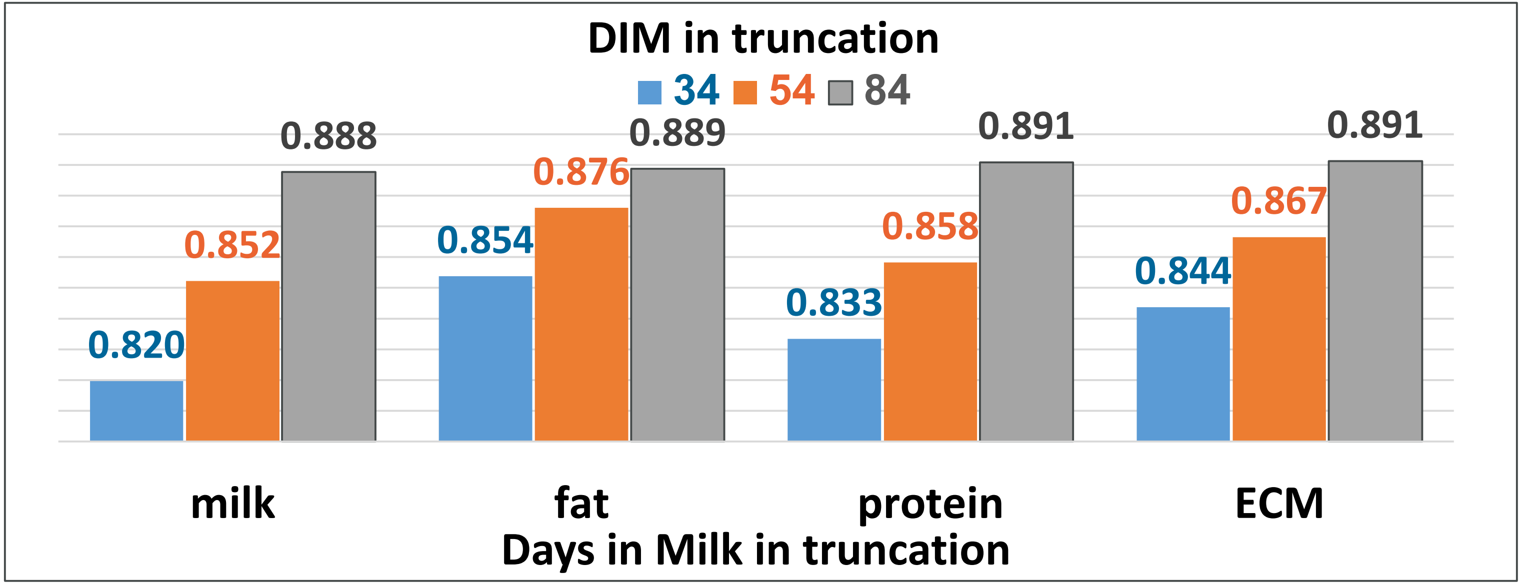
(Kg) Variable	The gap (%) to actual production	cows culled by mistake (%)
milk	-3.6%-7.0%	0.0%-18.5%
fat	-3.7%-7.8%	0.0%-23.8%
protein	-2.4%-6.5%	0.0%-13.2%
ECM	-1.8%-5.4%	0.0%-9.1%

✓The gaps between actual & predicted production allow safe planning

✓Selection for low production after 54 DIM could be unsatisfactory

Prediction improves with time (1) - Correlations

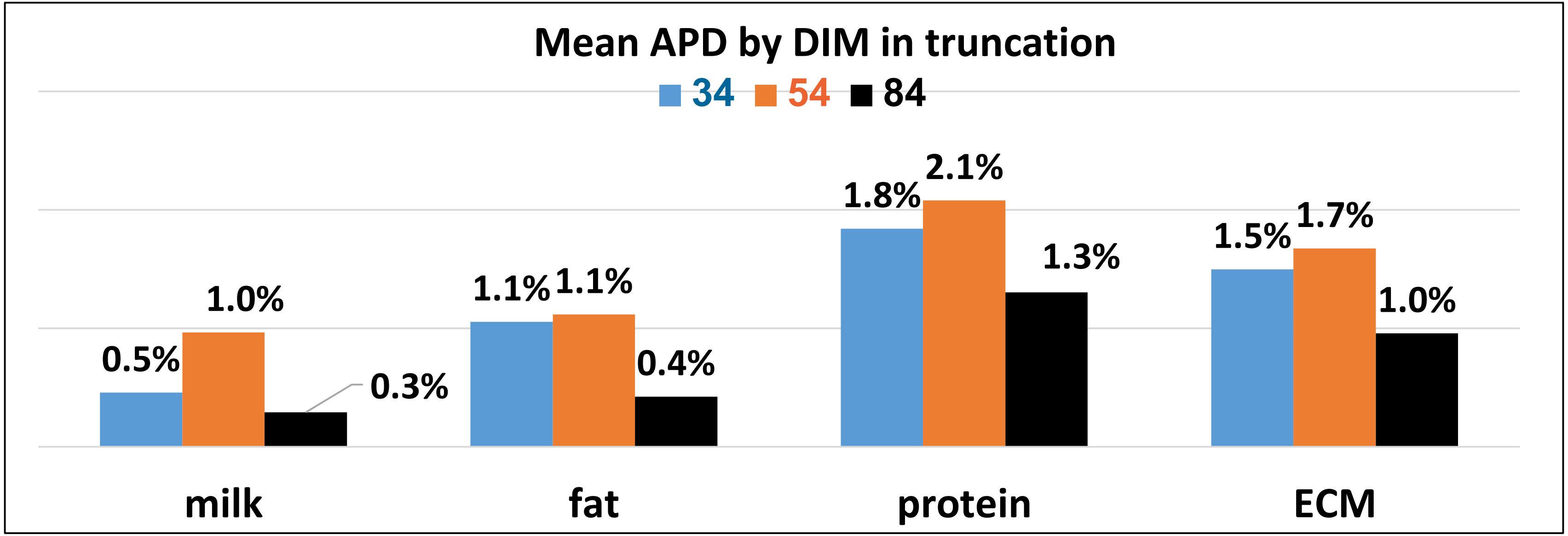
mean r between predicted & actual yields (14 herds)



- ✓Correlations improve with time in all production variables
- ✓Estimates, are estimates, even those carried out at 294 DIM could be erroneous!

Prediction improves with time (2) – Production (APD)

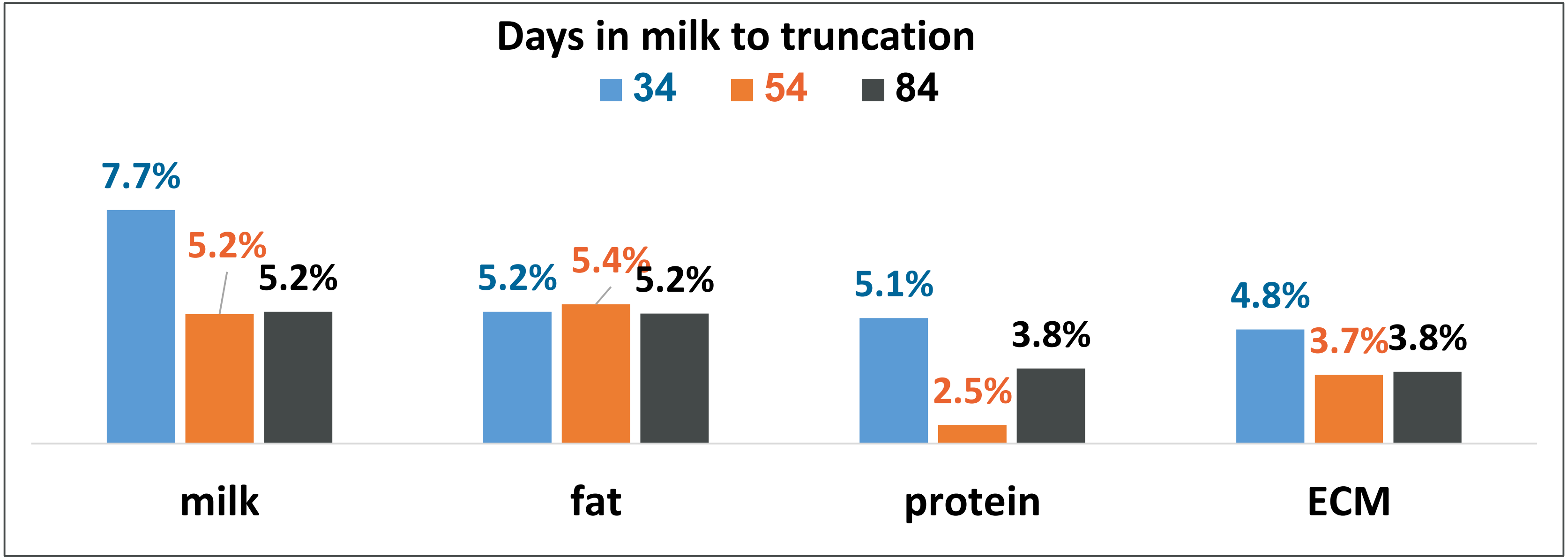
Gaps between predicted & actual 305 d yields (14 herds)



- ✓ APD (production gaps between actual & estimated yields) were satisfactory for all production variables
- ✓ Production planning could be corrected with time

Prediction improves with time (3) - Selection for yield

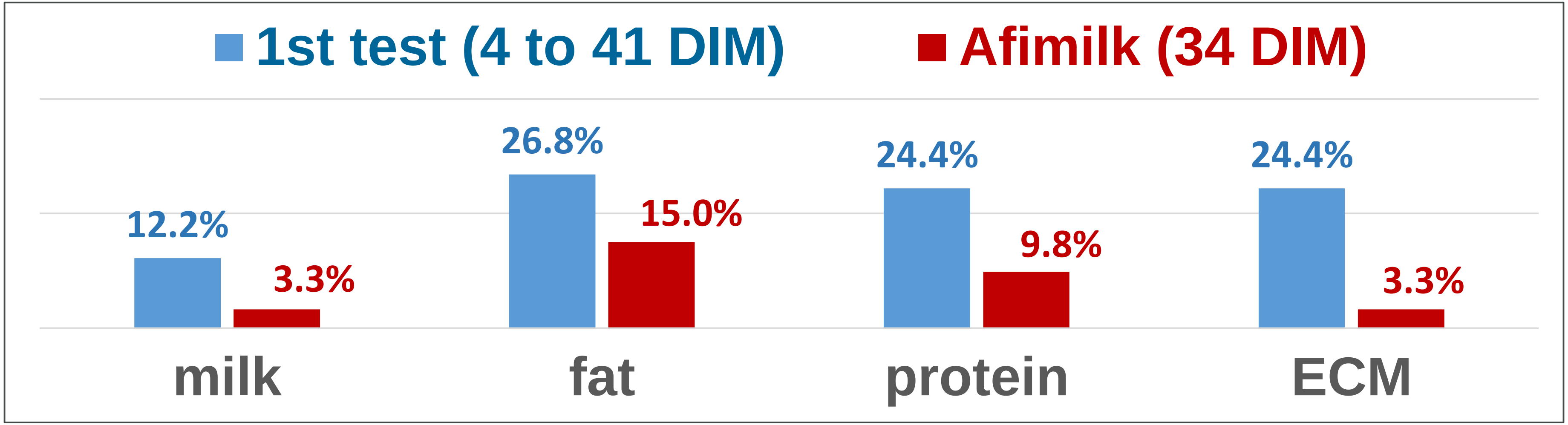
Rates (%) of cows erroneously culled for selection/yield (14 herds)



- ✓ 25% lowest yielders culled for selection
- ✓ Risk of erroneous culling is higher at 34 DIM
- ✓ Prediction improvement with time is not constant

Selection, erroneous culling (%), Afimilk 34 DIM versus 1st milk test

A Sample Herd, lowest 25% culled



✓Risk of erroneous culling after 1st test is higher than after 34 DIM in Afimilk

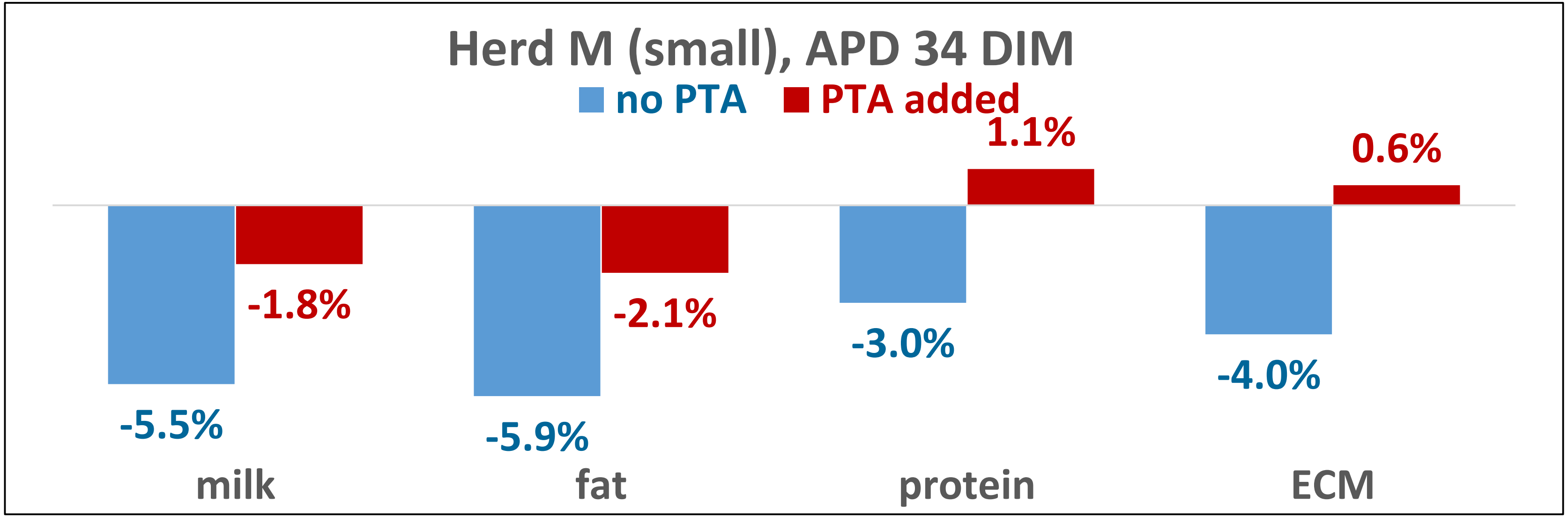
✓Correlations with actual production are higher in Afimilk

How to improve prediction?

- Add variables
  - ✓ Must have accurate individual data
  - ✓ Must have an effect on the outcome
- Improve models
  - ✓ Individual lactation numbers models vs. Common (all lactations combined models (allowing for lactation number)

Improve Prediction – Add Variables – Genomics (PTA), 1

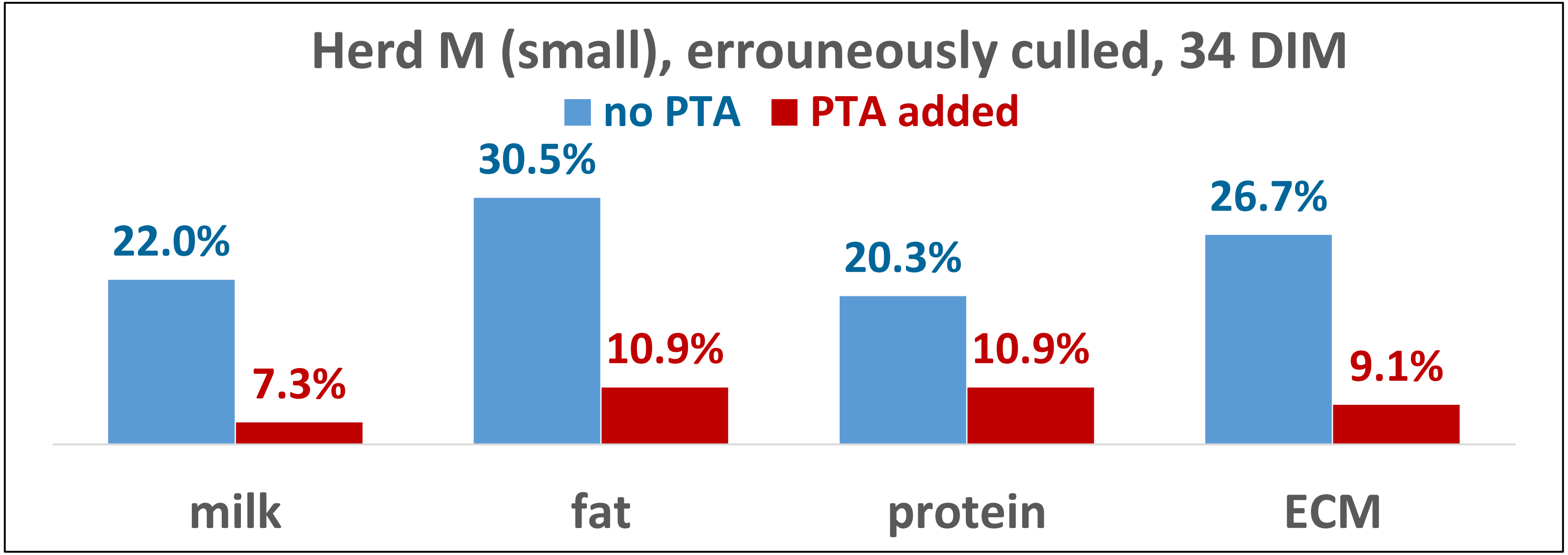
Parameters of APD after 34 DIM, Herd 15 (175 cows)



- ✓ APD (Predicted – Actual yields (%))
- ✓ APDs are reduced in all production variables after adding PTA to the models

Improve Prediction – Add Variables – Genomics (PTA) 2

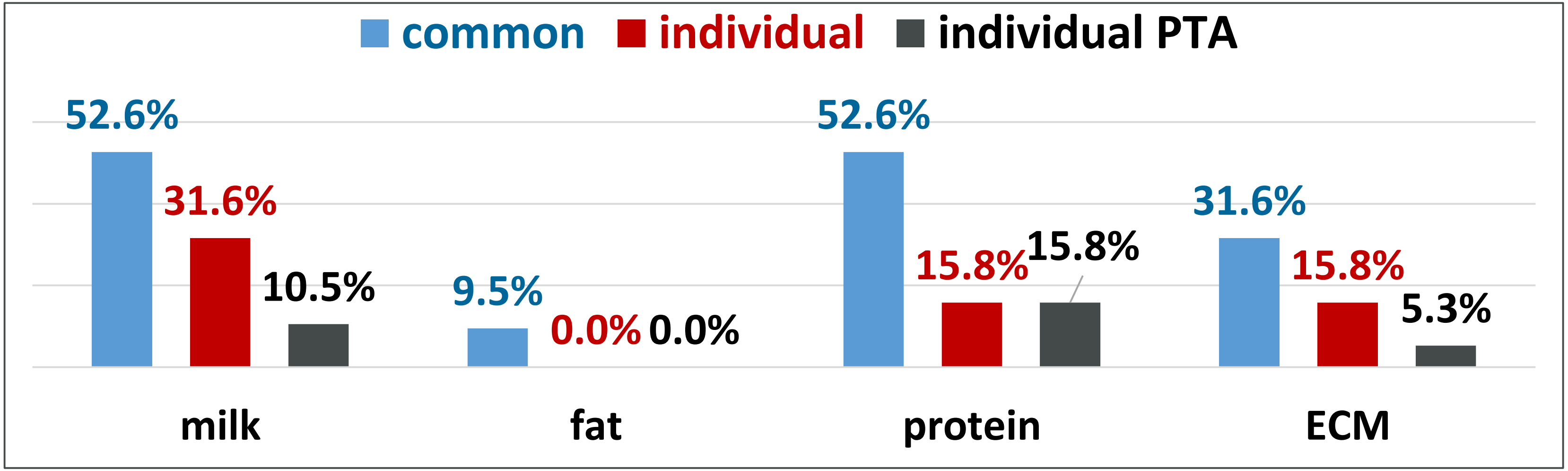
Rates (%) of cows erroneously culled after 34 DIM, Herd 15, 175 cows



- Rate cows culled by mistake = (1 – specificity)
- Erroneously culling is reduced in all production variables after adding PTA to the models

Individual vs. Common models (2)– Add PTA to the Models

1^{st} lactation cows erroneously culled **by 34 DIM**, Herd 10, n=721 cows



- ✓Rates of wrong positive cows (1 – specificity) for selection (cows culled by mistake)
- ✓ Improve when Individual replace the common (combined) models
- ✓Adding PTA to the individual cows models takes it even further

Planning : Annual predicted production (2018) (NOA)

Selected topics

#	topic	Jan.	Feb.	March		Oct.	Nov.	Dec.
5	n calved cows	62	65	63		76	100	67
6	n calved heifers	31	35	38		44	43	26
12	mean cows in milk	830	852	899.5		810	842	854.5
13	mean cows	965.5	970.5	979		992	1001.5	1002
14	mean dry cows	135.5	118.5	79.5		182	159.5	147.5
17	daily yield cows in milk (l)	32.95	33.91	34.4		30.04	31.01	32.95
18	daily yield all cows (l)	28.33	29.77	31.61		24.53	26.07	28.1
24	month marketing (l)	847,819	837,839	959,233		754,292	783,300	872,836
25	month quota (l)	828,891	757,367	869,061		788,720	783,821	842,608
28	above quota production (%)	2.3	10.6	10.4		-4.4	-0.1	3.6

Summary & Conclusions (1)

- ✓ Prediction is an on going process. The prospective prediction of all production variables improves with time from calving
- ✓ Accuracies of models based on AfiMilk MPC™ were higher than those established after the routine monthly milk tests
- ✓ Predictions improved by adding cow variables such as lengths of gestational & dry periods, twins & stillbirth, metritis, conductivity, & fat to protein ratio
- ✓ Rate of erroneous culling (1-specificity) is the most efficient criteria for an early selection

Summary & Conclusions (2)

- ✓ Extreme managemental changes might impair the quality of prediction, screening filters will minimize erroneous culling
- ✓ Separate calibration models applied to the various parities (1, 2, 3+) are superior to common ones
- ✓ Addition of genomic information may improve prediction
- ✓ Predicting in small herds is possible
- ✓ Accurate data is essential!
- ✓ **The present results allow planning of production, selection & culling in 54 DIM, and even earlier in lactation**

From Prediction to Genomics

Phenotype(t) = Genotype + Environment(t, x, y, z ….)

- **genotype** - the complete heritable genetic identity
- **phenotype** - description of the actual physical characteristics
- **Environment** - the circumstances, objects, or conditions of the surroundings (management, …, … , ….)

a) the dependent variables were production traits and diseases;
b) the examined factors were the PTA and the relevant managemental factors

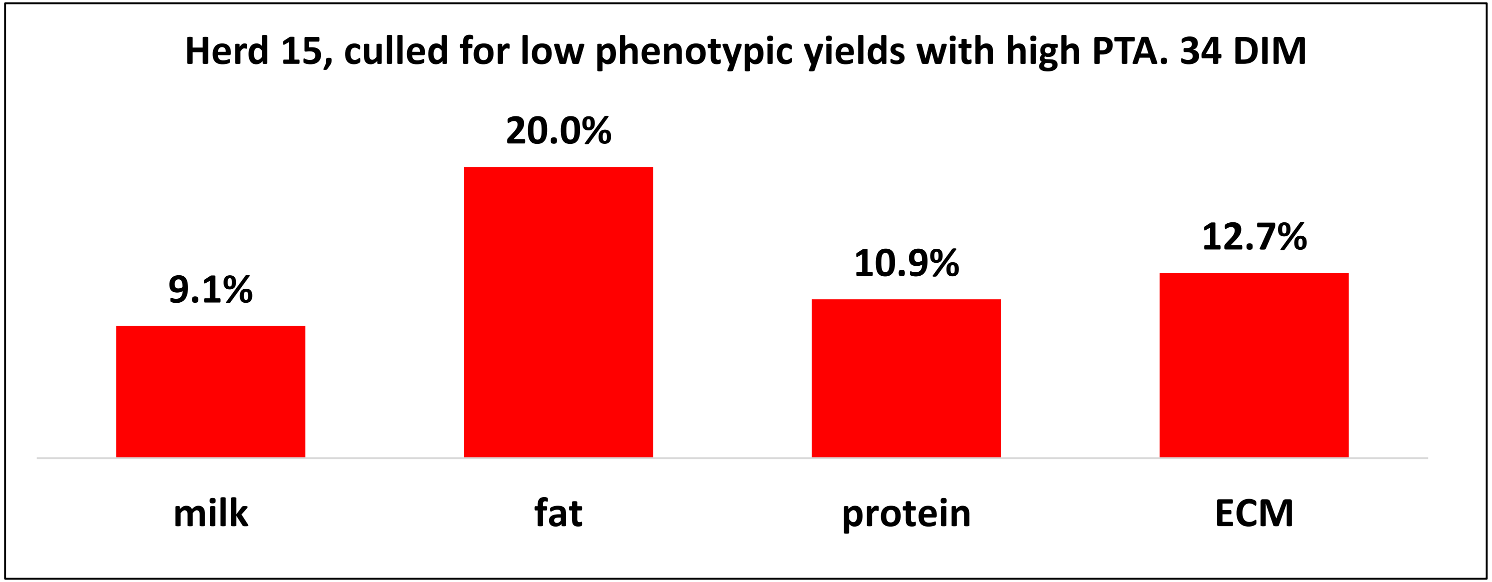
Can we make sure environmental factors are not mistaken for genomics in the selection process ?

- Phenotype(t)= Genotype + Environment(t, x, y, z ….)

Is there function F that for a given P_{t0}, P_{t1} can predict P_{tn}?

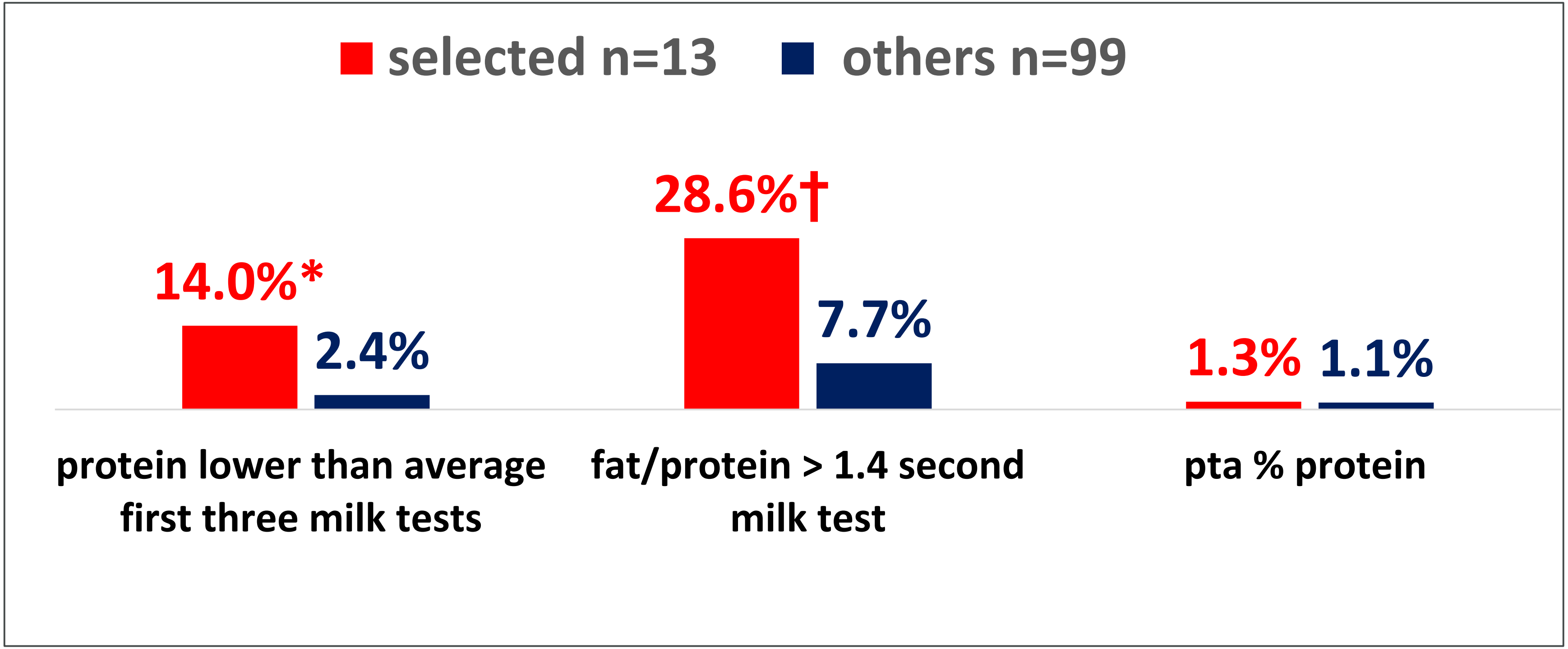
$$F(P\_{t0}, P\_{t1}, P\_{t2}, x\_{t0}, x\_{t1}, x\_{t2}, y, z\ldots) = P\_{tn}$$

Selection for low Production is not Genomic Selection!

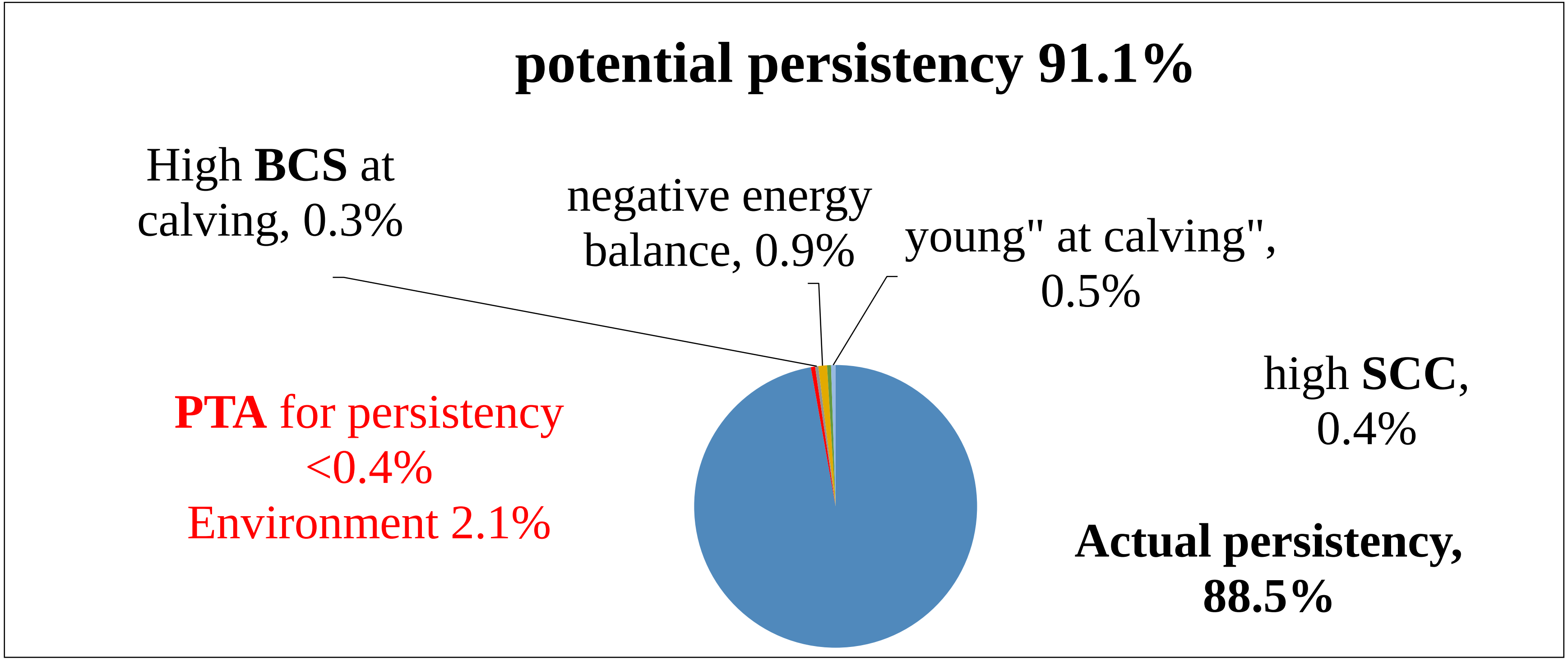


- ✓ Low phenotypic (actual) production = Lowest 25%
- ✓ High PTA = Highest 75%
- ✓ Don’t confuse!

Herd #1. **Selection for high milk protein percentage. Mistaking the phenotype for the genotype.**



Herd #2. Genetic & Environmental contributions to low milk persistency of first lactation cows.

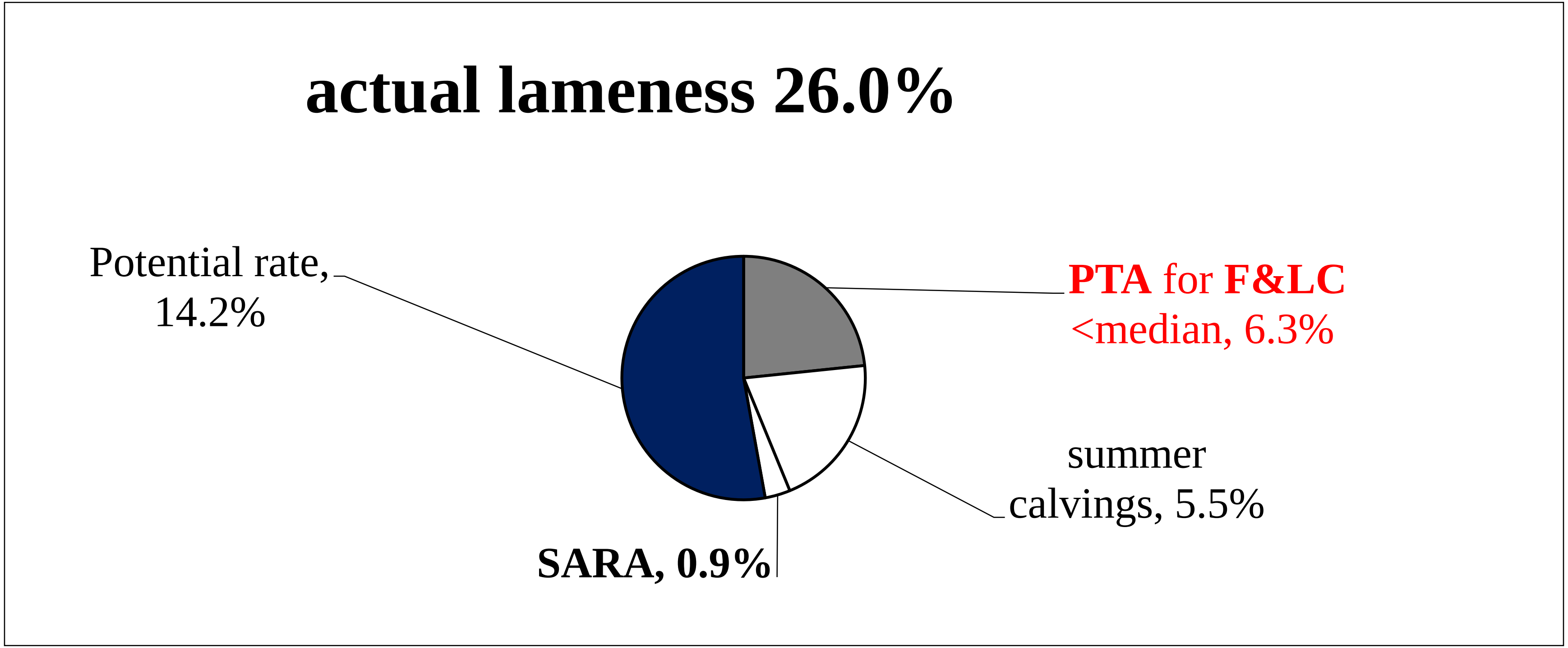


potential persistency 91.1%

PTA for persistency <0.4%
Environment 2.1%

BCS = body condition score; **SCC** = somatic cell count;

Herd #3. Genetic & Environmental contributions to lameness (up to 180 DIM) in first lactation cows.



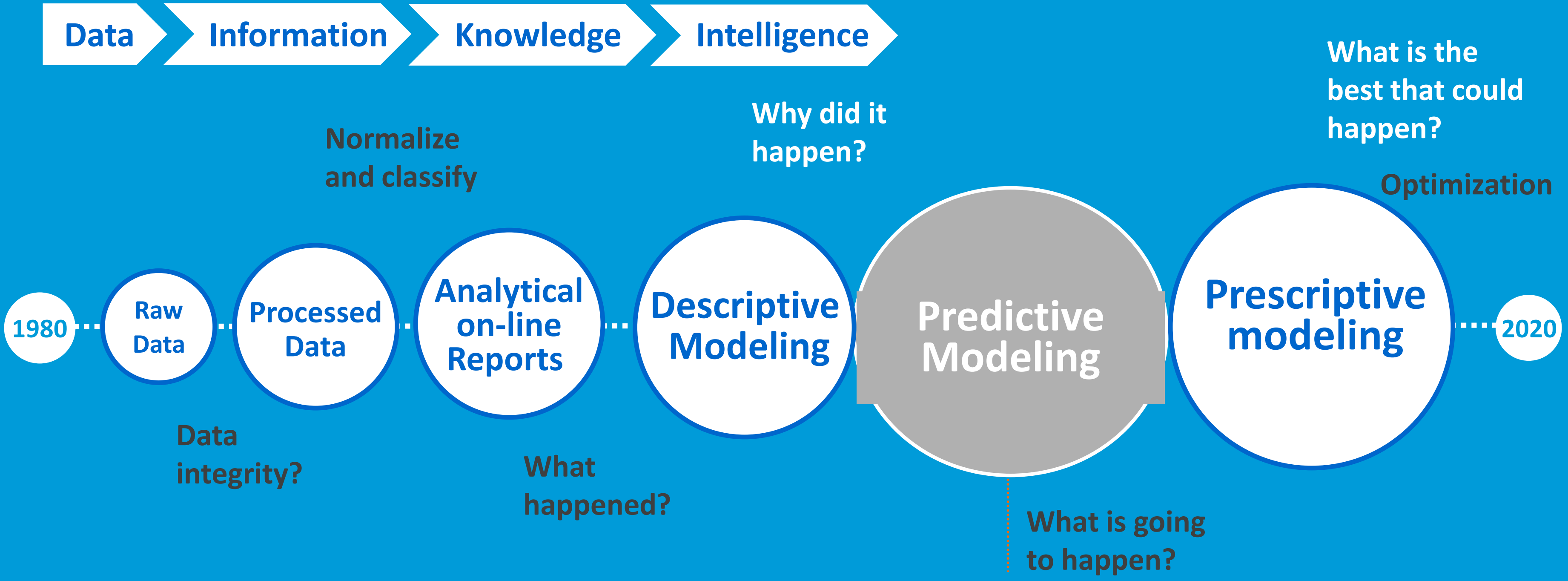
F&LC = Feet & Leg Composite; **SARA** = subacute rumenal acidosis

Conclusions (2):

- The key to select the secondary property is the interaction **the genomic*the economic importance of the property*the options of improving the phenotypic trait through managemental improvement.**
- **Improvement in availability and reliability of individual cow data (manual or automatic such as Afilab) will improve quantification of managemental factors to the phenotypic traits**.
- **In clinical trials, genomics should be taken into account.**

Phenotypes, Genotype, and Environment

Where do we go from here??



Data → Information → Knowledge → Intelligence

1980 ··· Raw Data ··· Processed Data ··· Analytical on-line Reports ··· Descriptive Modeling ··· Predictive Modeling ··· Prescriptive modeling ··· 2020

Data integrity?

Normalize and classify

What happened?

Why did it happen?

What is going to happen?

What is the best that could happen?

Optimization

Thank you!

Q&A

afimilk®
Vital know-how in every drop