



*Genotype **plus** Environment*
Integration for a more sustainable dairy production system

Random forests prediction of blood metabolic clusters of dairy cows: comparing three types of milk biomarkers

Leslie Foldager

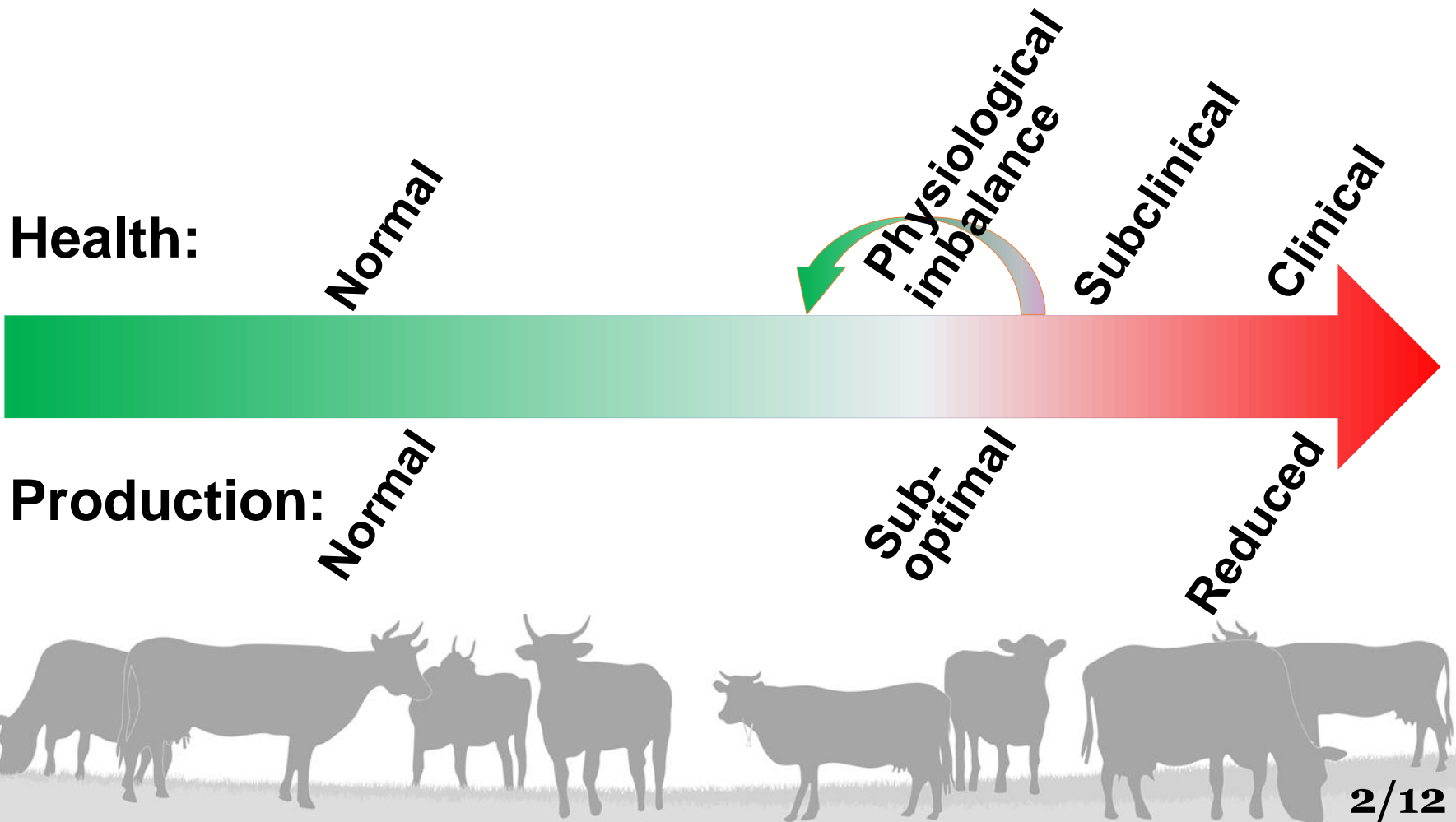
Department of Animal Science
Bioinformatics Research Centre
Aarhus University, Denmark



Fifth DairyCare Conference, Thessaloniki, Greece, March 19th and 20th 2018

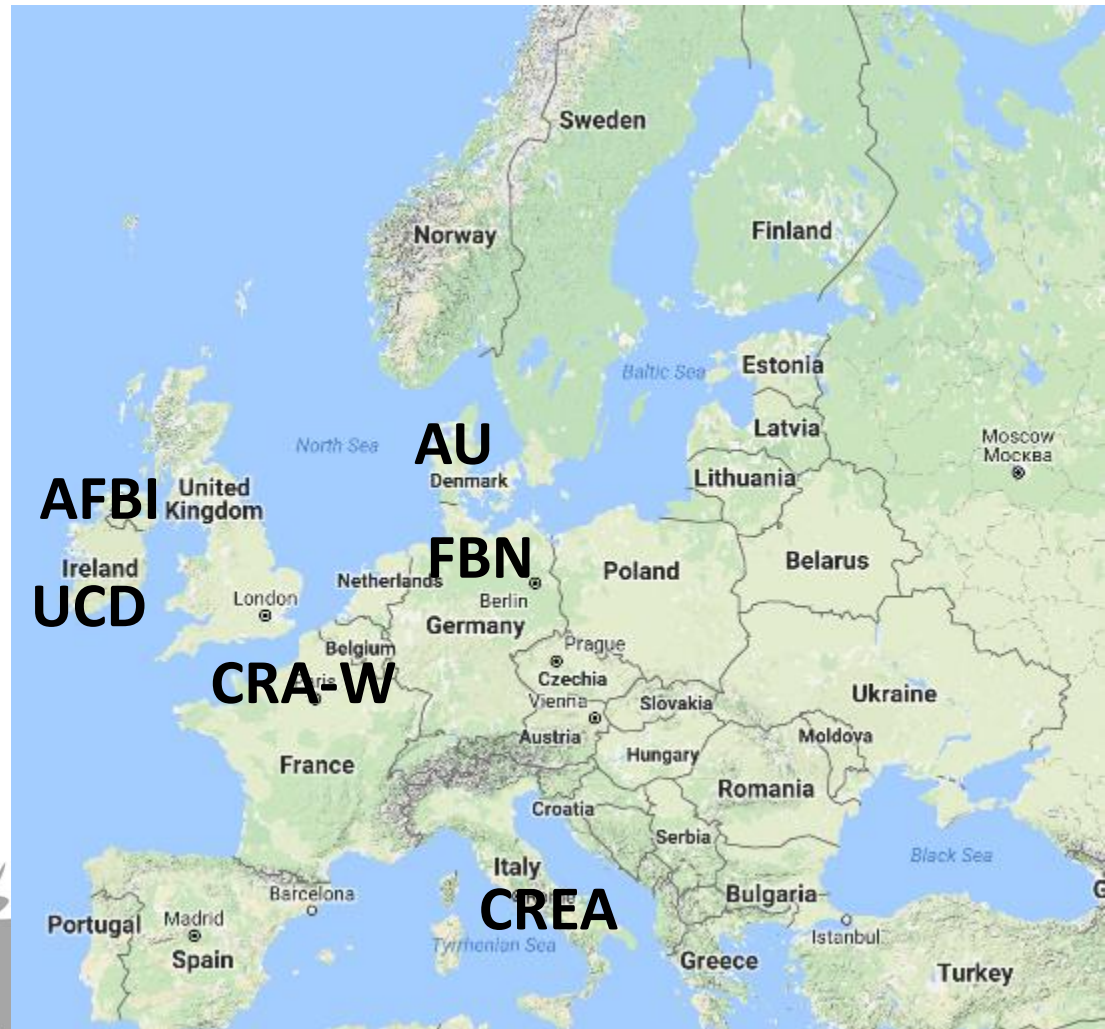


AIM: Can milk biomarkers be used for early prediction of physiological imbalance and disease?



Samples in GplusE WP3

- 234 Holstein cows
- 55 of 1st parity
- 66 of 2nd
- 113 of 3+ lactations
- observed 1-50 DIM
- in 6 research herds



Milk biomarkers/predictors

- Milk metabolites and enzymes (MM)
 - 6 metabolites: Glucose 6–Phosphate (**Glu6P**), Free Glucose (**GluFree**), β -hydroxybutyrate (**BHB**) [\log_{10}], Isocitric acid (**IsoC**), **Urea**, Uric acid (**UA**)
 - 2 enzymes: N-acetyl- β -D-glucosaminidase (**NAGase**) [\log_{10}], Lactate dehydrogenase (**LDH**) [\log_{10}]
- MIR spectra for 212 selected wavenumbers¹
- IgG glycan profiles (19 peaks)

1) J. Dairy Sci. 99(6):4816-4825,
DOI: 10.3168/jds.2015-10477

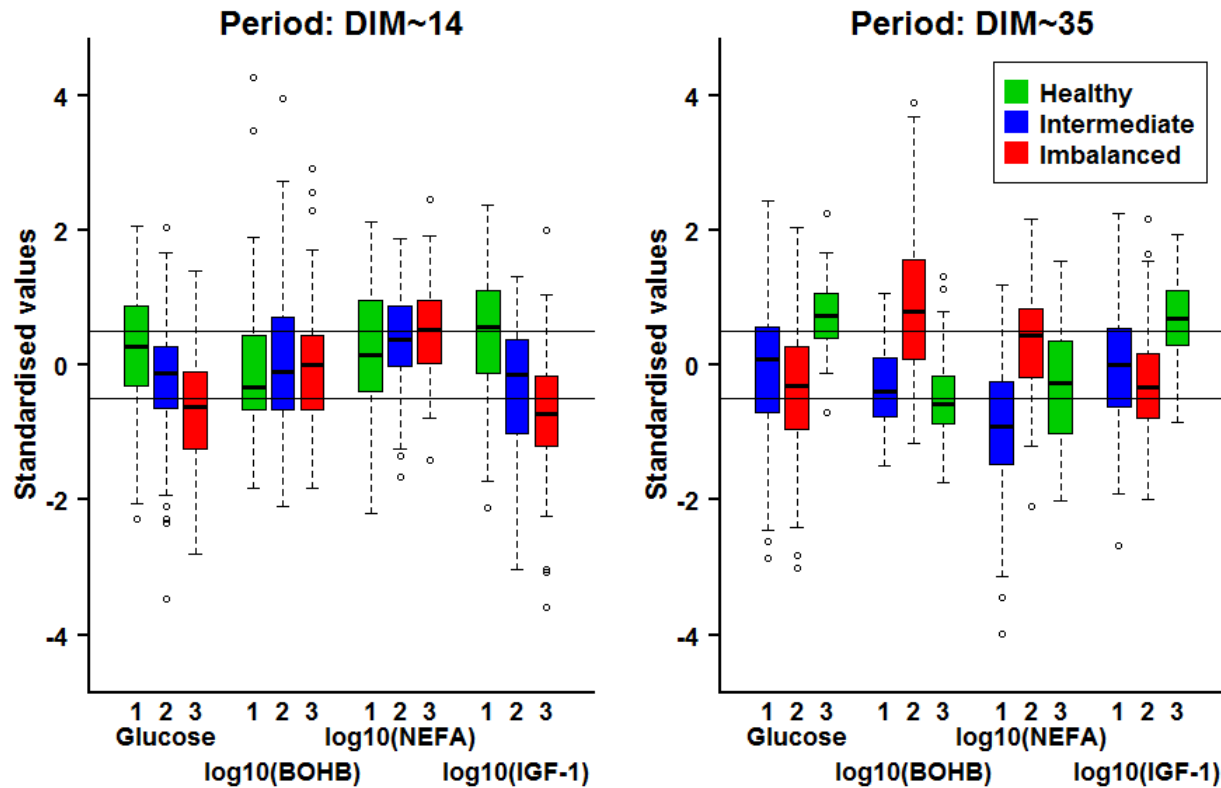


Metabolic clusters

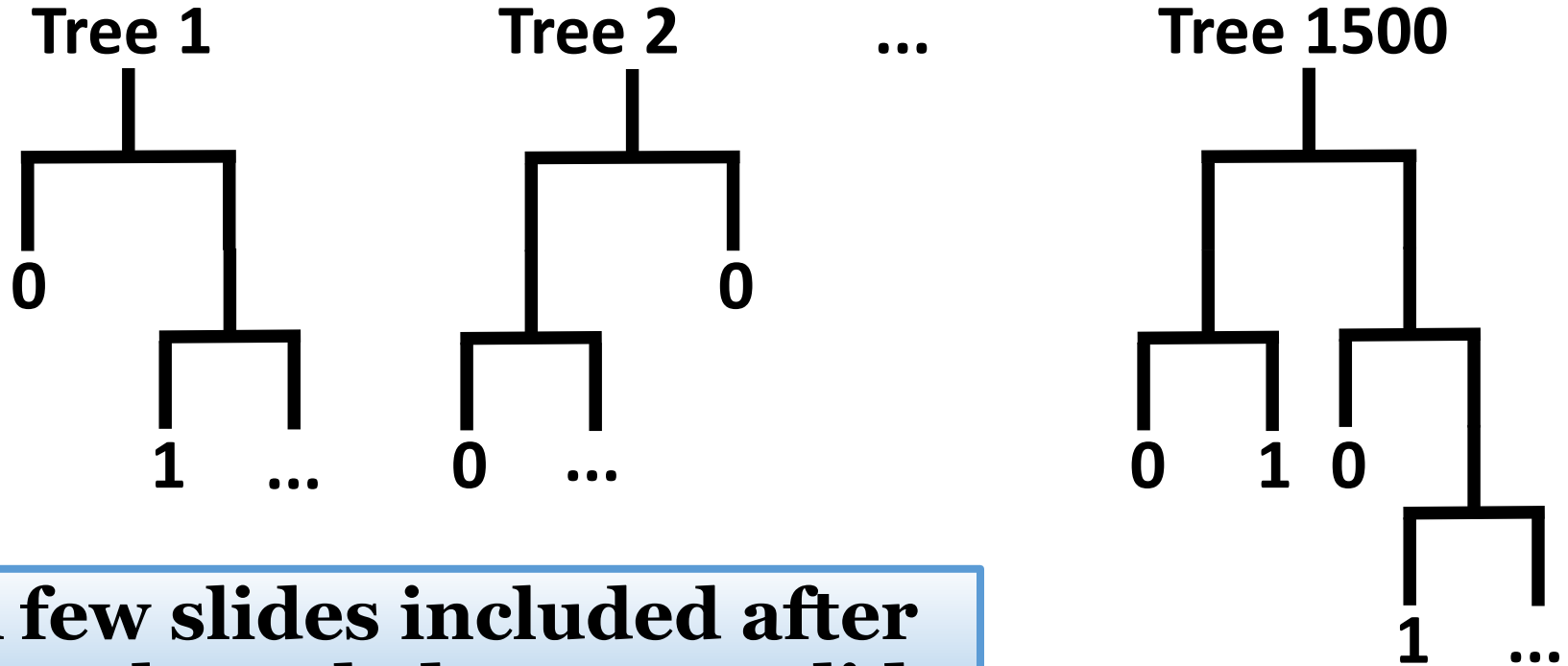
- By k-means (k=3) of four standardized blood biomarkers:
 - 3 plasma metabolites: **Glucose**, **BHB** [\log_{10}] and Non-Esterified Fatty Acids (**NEFA**) [\log_{10}]
 - 1 serum hormone: Insulin-like Growth Factor-1 (**IGF-1**) [\log_{10}]
- Within parity (1, 2, 3+) and period (~14, ~35 DIM)
- ... and within period across parity groups



Clusters as health groups



Random forests prediction model



A few slides included after the acknowledgements slide



Evaluation of performance by Leave-One-cow-Out-Cross-Validation

- **Accuracy – all 3 clusters**
 - Accuracy = $\text{sum}(\text{diagonal}) / \text{sum}(\text{all})$

- **Accuracy – imbalanced**

	1+2	3
1+2	$39+13+21+67=$ 140	$17+16=$ 33
3	$12+7=$ 19	19

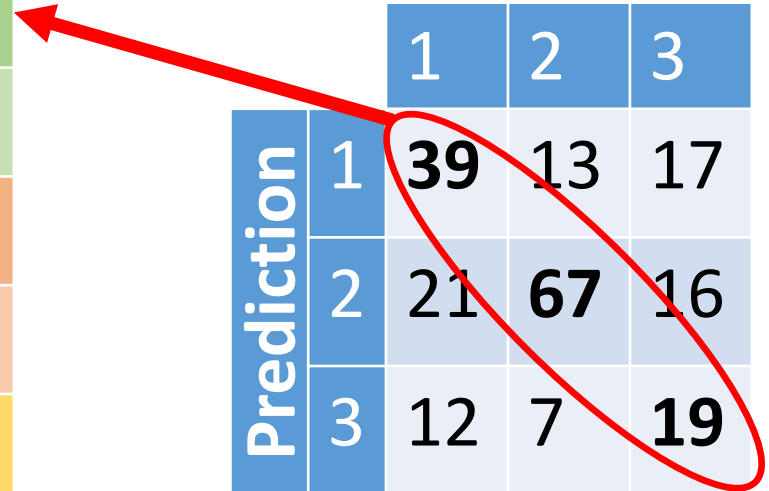
		Reference		
		1	2	3
Prediction	1	39	13	17
	2	21	67	16
	3	12	7	19



Accuracy of predicting clusters

Par.	DIM	Overall (95% CI) 3x3
MM	14	0.59 (0.52 - 0.66)
	35	0.59 (0.52 - 0.66)
MIR	14	0.54 (0.47 - 0.61)
	35	0.49 (0.41 - 0.56)
IgG	14	0.52 (0.43 - 0.61)
	35	0.35 (0.27 - 0.44)

		Reference		
		1	2	3
Prediction	1	39	13	17
	2	21	67	16
	3	12	7	19



Accuracy for imbalanced group

	DIM	Overall (95% CI) 3x3		Imbal grp vs rest 2x2	
MM	14	0.59 (0.52 - 0.66)		0.75 (0.69 - 0.81)	
	35	0.5	Reference		0.79 (0.73 - 0.85)
MIR	14	0.5	1+2	3	0.81 (0.75 - 0.86)
	35	Prediction	39+13+21	17+16	0.69 (0.62 - 0.76)
IgG	14		+67=140	= 33	0.72 (0.64 - 0.80)
	35		3	12+7=19	19



Agreement among biomarkers

DIM14	IgG glycans	MIR spectra
MM (N=116)	0.63 0.72 0.80	0.62 0.72 0.80
MIR (N=120)	0.67 0.76 0.83	
DIM35	IgG glycans	MIR spectra
MM (N=109)	0.56 0.66 0.75	0.63 0.72 0.80
MIR (N=110)	0.51 0.61 0.70	
N=192/186	MIR DIM14	MIR DIM35
MM (DIM14/35)	0.72 0.78 0.84	0.59 0.66 0.73



Acknowledgments and Disclaimer

- T. Larsen et al.: lab analyses of milk and blood samples
- M. Bjerring: information from databases
- M. Hostens, M. Salavati, C. Grelet: collaboration on statistical analysis and data
- K.L. Ingvarstsen, M.T. Sorensen: discussion and interpretation
- GpluseE Consortium: all the other contributors to the study



This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 613689

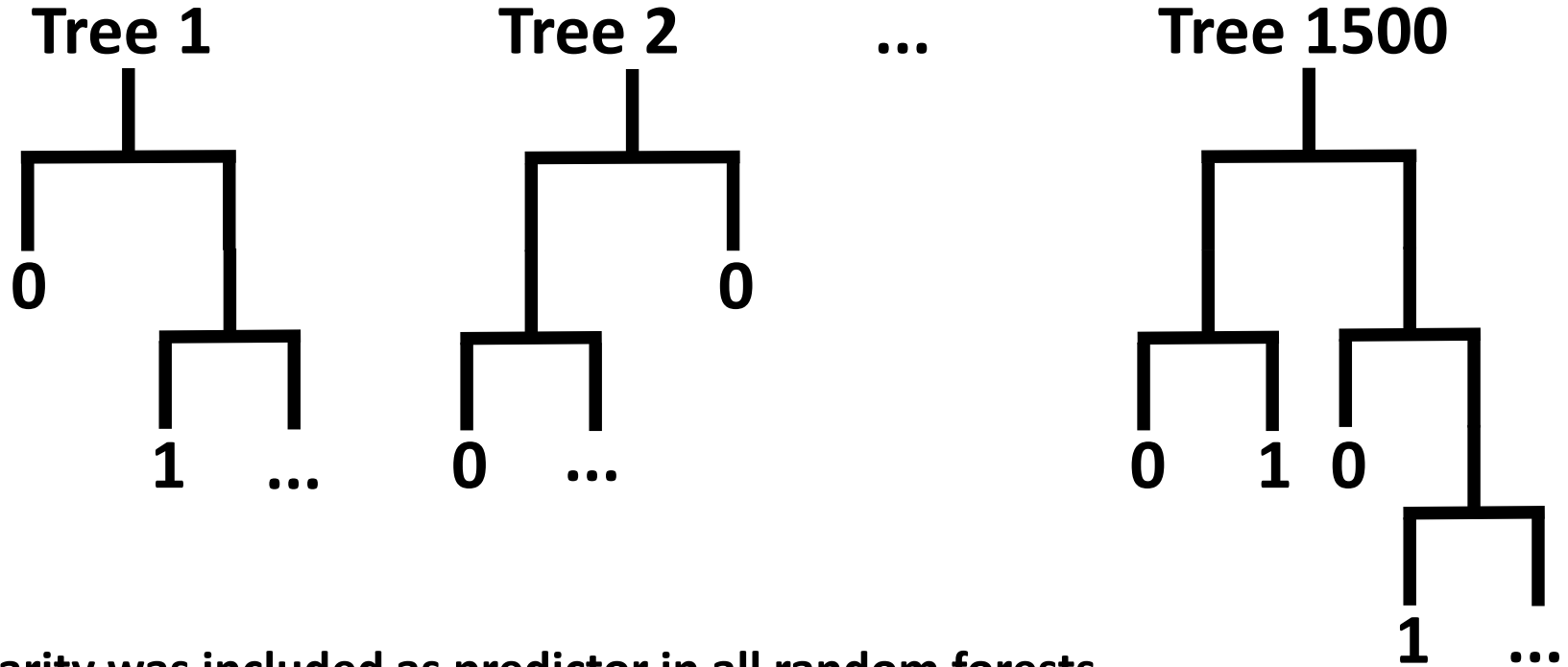
The views expressed in this publication are the sole responsibility of the author(s) and do not necessarily reflect the views of the European Commission.



The presenting author acknowledges the support (reimbursement) from the DairyCare network.



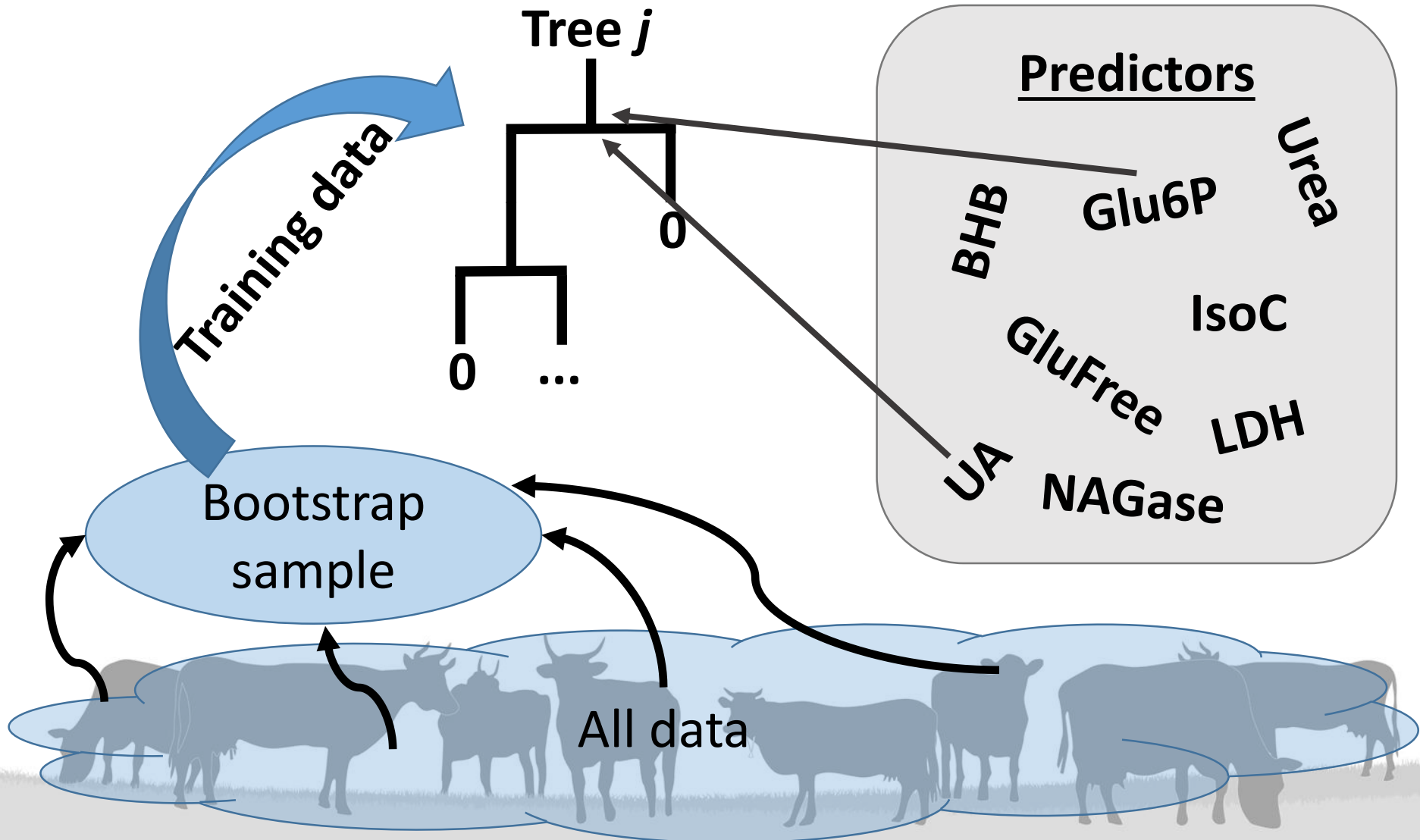
Random forests model¹



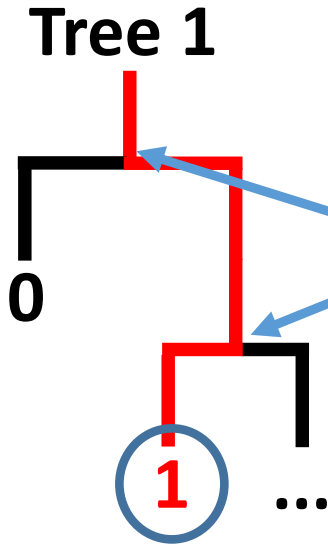
1) Parity was included as predictor in all random forests models unless restricting to a specific parity group



Random forests algorithm



Predicting for new data vector



- For each tree
 - apply decision rules on the new set of measures
 - using threshold at each split
 - obtain the predicted status



... by majority voting

