MANNA

European Joint Doctorate in Molecular Animal Nutrition

Getting prepared for Metabolomics studies: deliberate selection of samples from large animal studies

R. Riosa[1], M.H. Ghaffari[1], M. Iwersen[2], D. Suess[2], M. Drillich[2], C. Parys[3] and H. Sauerwein[1]

rriosa@uni-bonn.de

[1]Institute of Animal Science, Physiology and Hygiene Unit, University of Bonn, 53115 Bonn, Germany
[2]Clinical Unit for Herd Health Management in Ruminants, University of Veterinary Medicine Vienna, 1210 Vienna, Austria
[3]Evonik Nutrition & Care GmbH, 63457 Hanau, Germany

UNIVERSITÄT BONN | University of Glasgow | EVONIK POWER TO CREATE | vetmeduni vienna

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD: Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

R. Riosa / 70th EAAP Annual Meeting, Ghent 2019
rriosa@uni-bonn.de

Context

Large sample sizes allow for studying complex traits and rare incidents (e.g. fertility and diseases)

BUT: Sample numbers for OMICs studies **have to be limited to fewer samples**

Need for selecting meaningful subsets of animals/samples representing the desired traits

How?

<u>OPTION 1</u>

You can select animals based on specific and fixed labels

→ ***"human selection"***

<u>OPTION 2</u>

You can select animals based on Artificial Intelligence techniques

→ ***"Machine Learning selection"***

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

Aims

THE AIM OF THIS PROJECT WAS TO USE MACHINE LEARNING AS A TOOL TO SELECT MEANINGFUL SAMPLES THAT ARE REPRESENTATIVE OF A *VARIATION IN THE POPULATION* BY TESTING DIFFERENT HYPOTHESIS

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

MANNA

European Joint Doctorate in Molecular Animal Nutrition

Starting Database

UNIVERSITÄT BONN

University of Glasgow

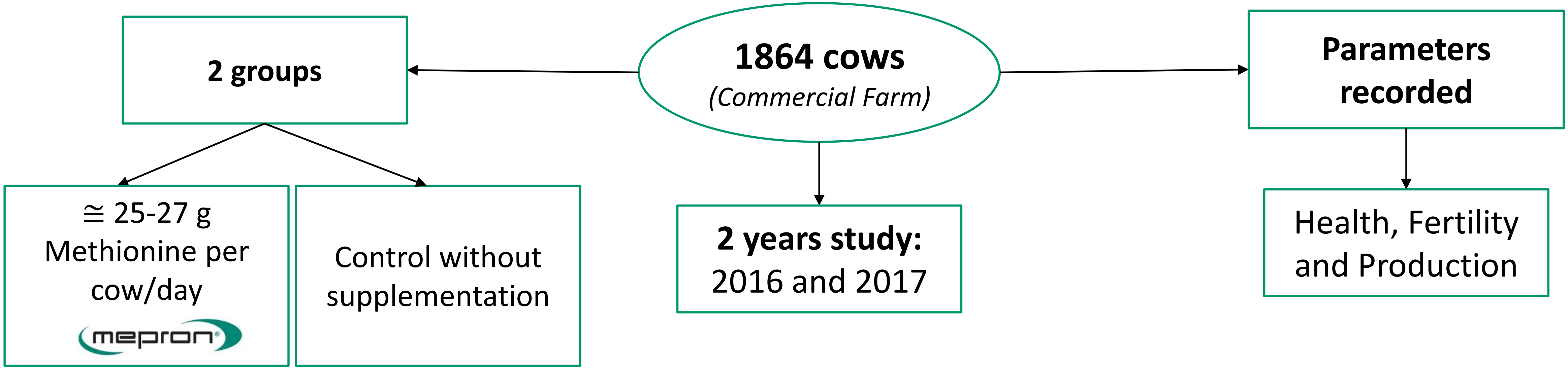
EVONIK POWER TO CREATE

vetmeduni vienna

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

R. Riosa / 70th EAAP Annual Meeting, Ghent 2019
rriosa@uni-bonn.de

Animal Trial



All these data were collected in Excel files + all the records relative to the cows were taken with the management software **DairyComp DC 305** (Valley Agricultural Software, USA) (e.g. cows changing groups, mastitis, sick pen…)

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

MANNA
European Joint Doctorate in Molecular Animal Nutrition

Screening & Animal Selection

UNIVERSITÄT BONN
University of Glasgow
EVONIK POWER TO CREATE
vetmeduni vienna

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

EAAP
European Federation of Animal Science

R. Riosa / 70th EAAP Annual Meeting, Ghent 2019
rriosa@uni-bonn.de

Screening of the Database

Starting from 1864 cows

Checked: **blood samples** and **availability of results** (*BCS, BFT, Production results…*)

Checked Met supplementation:
20-70 DIM + **no more 5 DIM** out of a group

Cows who entered the **trial twice** receiving Met in the 1st one were excluded

From **1864** cows to **1038** (*Masterfile*)

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

MANNA

European Joint Doctorate in Molecular Animal Nutrition

Using Machine Learning Techniques

UNIVERSITÄT BONN

University of Glasgow

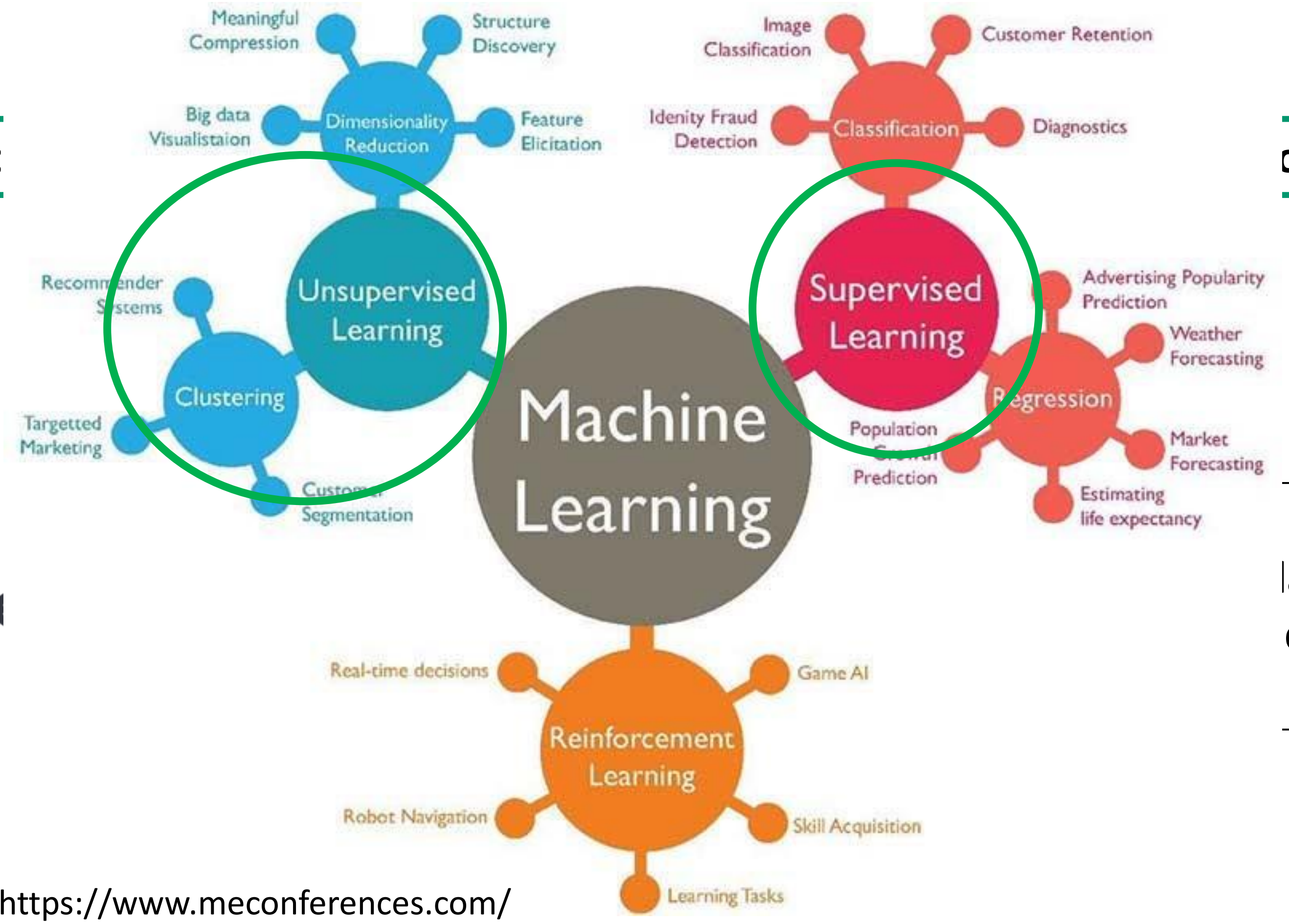
EVONIK POWER TO CREATE

vetmeduni vienna

This project has received funding from the European Union’s Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

What is Machine Learning (ML) ?

Is a subfield of **art** ... on and **inference**



https://www.meconferences.com/

rapidmin...

... company of the
... ata preparation,
... dictive analytics

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

R. Riosa / 70th EAAP Annual Meeting, Ghent 2019
rriosa@uni-bonn.de

Why using ML techniques?

Having a **large Database**, in terms of **number of animals** and **availability of results**

Permits to validly use **ML** techniques
→ high number of data means higher statistical power

Permits to analyse different variables and have a complete picture of the situation

RESULTING IN A BETTER OVERVIEW OF THE POSSIBLE *VARIATION IN THE POPULATION*

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

rriosa@uni-bonn.de

How did we use ML?

SUPERVISED ML – PREDICTION

- Permits to **identify patterns** in data
- Uses different models (*Native Bayes, Generalized Linear Model,* ***Deep Learning****…*)
- Importance of **precision**, **accuracy** and **sensitivity**

→We aimed to find interactions between treatments (Control vs. Met) and disease by this method

→ → we identified **Endometritis x Met being most significant**

UNSUPERVISED ML - CLUSTERING

- Permits to **identify groups** of similar data
- **K-means** and **X-means** are the models that were used
- We used both of them → trying to identify the best possible grouping

→ Via this method we identified different grouping based on the **BFT-BHB** interaction & **Pattern of Body Condition Loss** during transition period

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

R. Riosa / 70th EAAP Annual Meeting, Ghent 2019
rriosa@uni-bonn.de

Results from ML

We were able to select cows based on 4 different biological aspect

Supervised Machine Learning

- **Endom or Healthy vs Methionine or Control**
- **Success or not in the 1st AI considering health status**

Unsupervised Machine Learning

- **Clustering based on BHB and BFT**
- **Body Condition Loss around Parturition**

GROUPINGS WERE MADE DEPENDING ON WHAT WE WERE INTERESTED IN!!
EACH GROUPING IS DIFFERENT FROM THE OTHER → THE POWER OF ML

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

MANNA

European Joint Doctorate in Molecular Animal Nutrition

Validation of Results

UNIVERSITÄT BONN

University of Glasgow

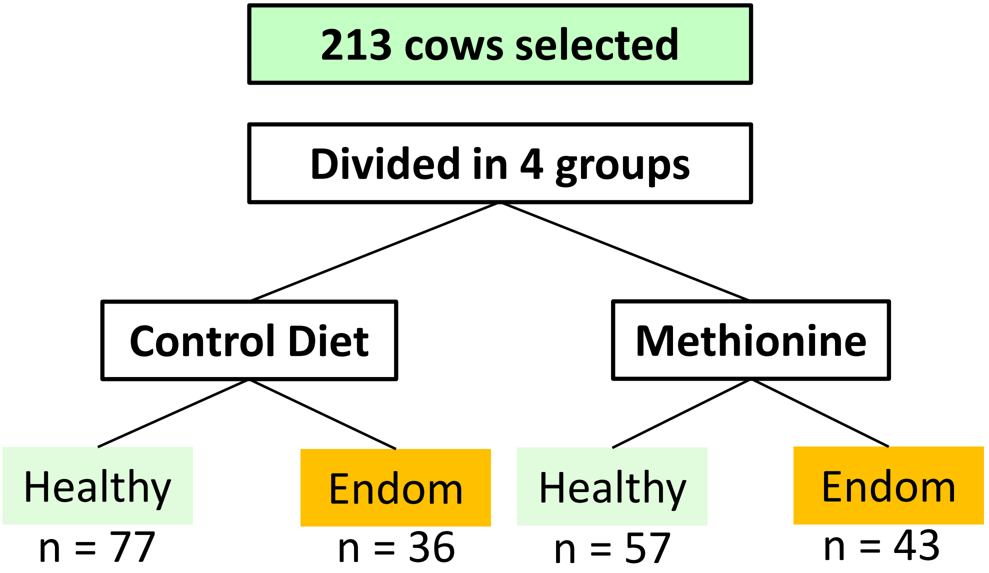
EVONIK POWER TO CREATE

vetmeduni vienna

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

R. Riosa / 70th EAAP Annual Meeting, Ghent 2019
rriosa@uni-bonn.de

Diet Treatment vs Health Condition



Data analyzed using the *MIXED procedure of SAS* (SAS Institute Inc., Cary, NC)
Model included: **random effect of cow and the fixed effect of the main factors M, CE, and their interaction (M × CE)**
Days in milk was considered a covariate for the dependent variables

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

EAAP European Federation of Animal Science

R. Riosa / 70th EAAP Annual Meeting, Ghent 2019
rriosa@uni-bonn.de

Diet Treatment vs Health Condition

Disease Effect

	P-value
Calving to conception interval (d)	0.0053
Milk at 30 DIM (Kg/d)	0.013
ECM at 30 DIM	0.079
Urea at 30 DIM (mg/100 mL)	0.003
Fat at 60 DIM (kg)	0.05
Protein at 30 DIM (kg)	0.04

Disease x Diet Effect

	P-value
BFT at 50 DIM (mm)	0.065
BCS at 50 DIM (score)	**0.04**
SCC at 70 DIM (10^3/ml)	0.097
ECM at 30 DIM	0.065
F:P Ratio at 30 DIM	**0.008**
Fat at 30 DIM (%)	**0.015**

Diet Effect

	P-value
BFT at 30 DIM (mm)	0.06
BCS at 30 DIM (score)	0.04
SCC at 70 DIM (10^3/ml)	0.039
Milk at 60 DIM (kg/d)	0.055
Protein at 30 DIM (kg)	0.094
Protein at 60 DIM (kg)	0.025

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

MANNA

European Joint Doctorate in Molecular Animal Nutrition

Conclusion & Future Objectives

UNIVERSITÄT BONN

University of Glasgow

EVONIK POWER TO CREATE

vetmeduni vienna

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

R. Riosa / 70th EAAP Annual Meeting, Ghent 2019
rriosa@uni-bonn.de

Conclusion & Future Objectives

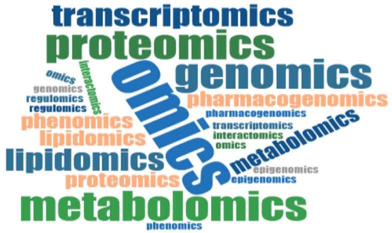
Having large datasets is desirable, but a considerable amount of time is necessary to prepare the data for Machine Learning

Machine Learning is an important tool that enables to find patterns and thus select meaningful samples

Results from the *Treatment vs Disease* study showed that the grouping was effective as statistical differences and interactions were found between the groups

Different phenotypes were detected
→ future **OMICs analyses** will permit us to have further insights and obtain a complete picture of the mechanisms involved

Machine Learning will be used again to further analyse the OMICs results with the **aim to identify markers** which can then be **targeted on all cows from the starting database**



This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

"Data is the new science. Big Data holds the answers. Are you asking the right questions?"
Patrick P. Gelsinger

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

MANNA
European Joint Doctorate in Molecular Animal Nutrition

THANK YOU FOR YOUR ATTENTION

UNIVERSITÄT BONN

University of Glasgow

EVONIK POWER TO CREATE

vetmeduni vienna

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

MANNA

European Joint Doctorate in Molecular Animal Nutrition

Supplementary Material

UNIVERSITÄT BONN

University of Glasgow

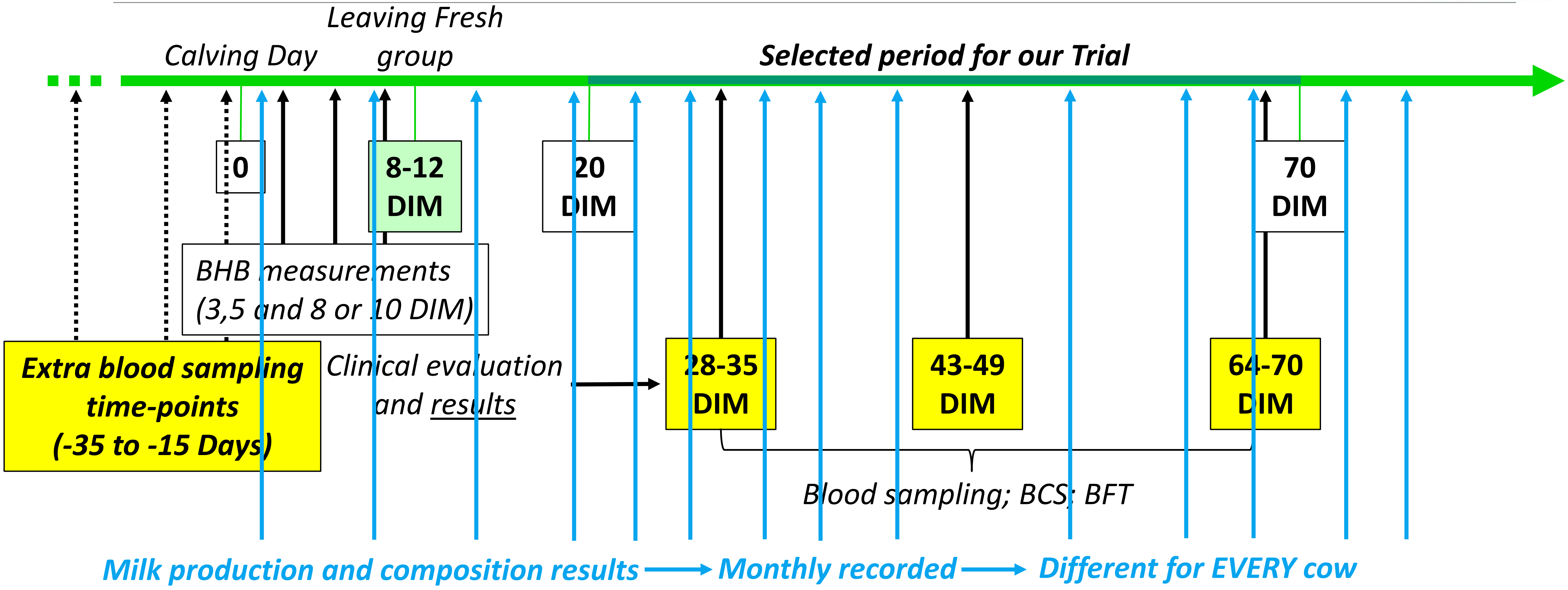
EVONIK POWER TO CREATE

vetmeduni vienna

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

R. Riosa / 70th EAAP Annual Meeting, Ghent 2019
rriosa@uni-bonn.de

Overview - Animal Trial



This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

Endometritis vs Healthy Selection

BASED ON ENDOMETRITIS

Evaluation of mucus content of vagina

Williams et al. (2005):
(*E0*) clear and transparent mucus
(*E1*) ≤ 50% off-white or white mucopurulent material and
(*E2*) ≥ 50% off-white or white mucopurulent material

Cow UN-HEALTHY if: E2 + ≥5%

Cow HEALTHY if: E0 + <5%

+

We selected the cows that were diagnosed with a **clinical** or a **puerperal metritis**

Proportion of polymorphonuclear neutrophils (PMN)

Melcher et al. (2014):
Determined by counting a total of 300 cells from a cytological vaginal sample

Cut-off value for diagnosis
→ **5%**

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD:
Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

R. Riosa / 70th EAAP Annual Meeting, Ghent 2019
rriosa@uni-bonn.de

Statistical Analysis

	Control		Methionine		SEM	P-Value		
	CE	HEL	CE	HEL		TREAT	DISEASE	T x D
BFT30DIM (mm)	10.17	11.04	10.14	9.24	0.238	**0.06**	0.79	0.11
BFT50DIM	8.81	10.25	9.23	8.69	0.249	0.16	0.40	**0.065**
BFT70DIM	8.61	9.91	9.07	8.82	0.265	0.86	0.23	0.14
BCS30DIM (score)	2.69	2.75	2.64	2.58	0.027	**0.04**	0.91	0.26
BCS50DIM	2.56	2.74	2.63	2.56	0.028	0.29	0.33	**0.04**
BCS70DIM	2.58	2.70	2.56	2.57	0.031	0.33	0.31	0.43

	Control		Methionine		SEM	P-Value		
	CE	HEL	CE	HEL		TREAT	DISEASE	T x D
Calving to Conception Interval (d)	114.43	95.47	121.67	101.56	3.299	0.47	**0.0053**	0.92

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD: Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

		Control		Methionine		SEM	P-Value				
		CE	HEL	CE	HEL		TREAT	DISEASE	T x D	PARITY	DIM
1st Time Point	Milk (kg/d)	35.32	38.07	34.70	38.49	0.594	0.50	**0.004**	0.52	0.10	<.0001
	ECM	33.55	36.55	32.94	36.94	0.585	0.62	**0.004**	0.67	0.006	<.0001
	SCC (10^3/ml)	157.59	115.72	148.52	128.93	15.307	0.79	0.27	0.39	0.76	<.0001
	Prot (%)	3.17	3.26	3.24	3.23	0.026	0.72	0.26	0.25	0.25	<.0001
	Fat (%)	3.73	3.75	3.58	3.81	0.065	0.95	0.42	0.43	0.002	<.0001
	Urea mg/100ml	21.78	24.12	21.13	25.24	0.549	0.63	**0.003**	0.51	0.55	0.099
2nd Time Point	Milk (kg/d)	42.13	42.45	43.14	43.90	0.383	0.12	0.50	0.78	0.009	0.74
	ECM	37.03	38.72	38.10	39.77	0.449	0.24	**0.07**	0.99	<.0001	0.14
	SCC (10^3/ml)	205.87	73.00	76.14	100.73	18.293	0.12	0.27	0.63	0.25	0.86
	Prot (%)	3.01	3.07	3.04	3.03	0.017	0.91	0.57	0.24	0.53	0.20
	Fat (%)	3.30	3.44	3.17	3.39	0.051	0.43	**0.05**	0.65	0.0003	0.014
	Urea mg/100ml	24.05	26.03	25.22	25.55	0.564	0.70	0.41	0.76	0.057	0.62
3rd Time Point	Milk (kg/d)	40.59	40.70	41.66	40.28	0.404	0.69	0.45	0.37	0.017	0.054
	ECM	36.57	37.82	36.73	36.78	0.420	0.60	0.44	0.47	<0.0001	0.054
	SCC (10^3/ml)	246.62	108.72	85.42	135.13	18.815	**0.039**	0.71	**0.097**	0.40	0.75
	Prot (%)	3.09	3.14	3.13	3.12	0.018	0.84	0.49	0.40	0.90	0.75
	Fat (%)	3.35	3.56	3.18	3.42	0.057	0.17	0.54	0.87	0.0002	0.84
	Urea mg/100ml	24.79	25.43	25.84	25.83	0.438	0.44	0.74	0.73	0.33	0.76

Statistical Analysis

Too many variations in DIM for Milk production → *Further selection based on DIM* **(Analyses with 54 cows)**

		Control		Methionine		SEM	P-Value			
		CE	HEL	CE	HEL		TREAT	DISEASE	T x D	PARITY
Around 30 DIM	Milk (kg/d)	40.13	44.51	41.99	46.27	0.695	0.30	**0.013**	0.94	0.78
	ECM	40.22	40.23	38.10	44.67	0.853	0.63	**0.079**	**0.065**	0.30
Around 60 DIM	Milk (kg/d)	40.05	43.19	42.92	47.24	0.810	**0.055**	**0.046**	0.73	0.35
	ECM	37.08	39.19	37.65	43.78	0.930	0.16	**0.034**	0.26	**0.038**

	Control		Methionine		SEM	P-Value			
	CE	HEL	CE	HEL		TREAT	DISEASE	T x D	PARITY
F:P 30DIM	1.36	1.10	1.08	1.23	0.037	0.32	0.49	**0.008**	**0.075**
F:P 60DIM	1.12	1.13	1.01	1.13	0.025	0.28	0.27	0.31	**0.06**

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD: Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

R. Riosa / 70th EAAP Annual Meeting, Ghent 2019
rriosa@uni-bonn.de

Statistical Analysis

		Control		Methionine		SEM	P-Value			
		CE	HEL	CE	HEL		TREAT	DISEASE	T x D	PARITY
Around 30 DIM	**Prot %**	3.01	3.06	3.11	3.11	0.025	0.19	0.65	0.75	0.30
	Prot kg	1.23	1.36	1.31	1.44	0.021	**0.094**	**0.005**	0.93	0.66
Around 60 DIM	**Prot %**	3.12	3.01	3.10	3.14	0.034	0.42	0.78	0.30	**0.046**
	Prot kg	1.25	1.30	1.32	1.48	0.029	**0.025**	**0.075**	0.31	**0.014**

		Control		Methionine		SEM	P-Value			
		CE	HEL	CE	HEL		TREAT	DISEASE	T x D	PARITY
Around 30 DIM	**Fat %**	4.09	3.36	3.36	3.80	0.112	0.53	0.60	**0.015**	**0.051**
	Fat kg	1.66	1.50	1.41	1.77	0.054	0.92	0.40	**0.016**	**0.09**
Around 60 DIM	**Fat %**	3.49	3.42	3.11	3.53	0.085	0.44	0.34	0.16	**0.009**
	Fat kg	1.40	1.48	1.34	1.66	0.048	0.58	**0.04**	0.24	**0.003**

This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD: Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA