



Aspects of data validation and data quality based on veterinary diagnoses

C. Egger-Danner[1], W. Obritzhauser[2], C. Fuerst[1],
B. Grassauer[3], K. Zottl[4] , K. Fuchs[5] and B. Fuerst-Waltl[6]

egger-danner@zuchtdata.at

ICAR 2013 – Health Data Conference

Aarhus, 31st of May, 2013

Overview

- General aspects
- Example – Health monitoring AUSTRIA
  - use of data
  - standardisation
  - logistics
  - plausibility check
  - validation
- Aspects of validation
- Measures to improve data quality
- Conclusions

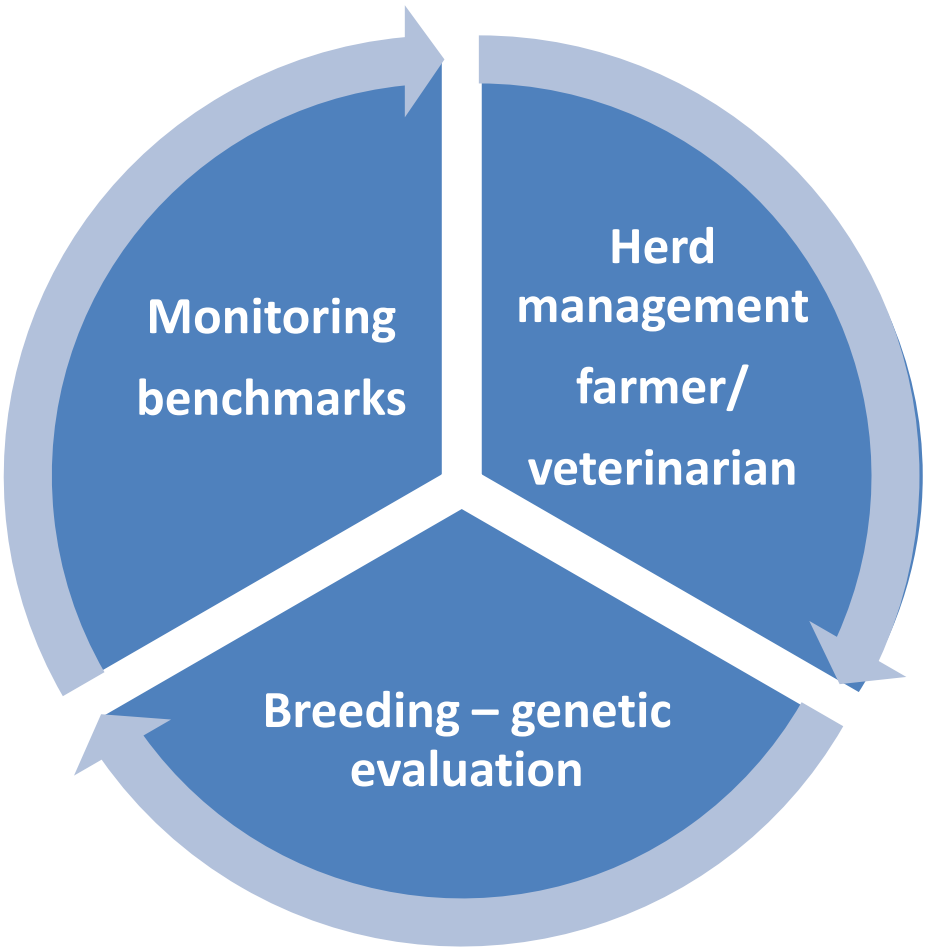
General

- Establishing a system of registration of veterinarian diagnoses **envolves participation of different stakeholders**
- System has to be adjusted to the existing circumstances with a **minimum of work / effort** for farmers and veterinarians
- **Benefit / use** of data for stakeholders is essential!
- Field data: envolve high emphasis on validation!

Differentiate farms with low frequency versus farms with incomplete health data recording

Use of direct health data

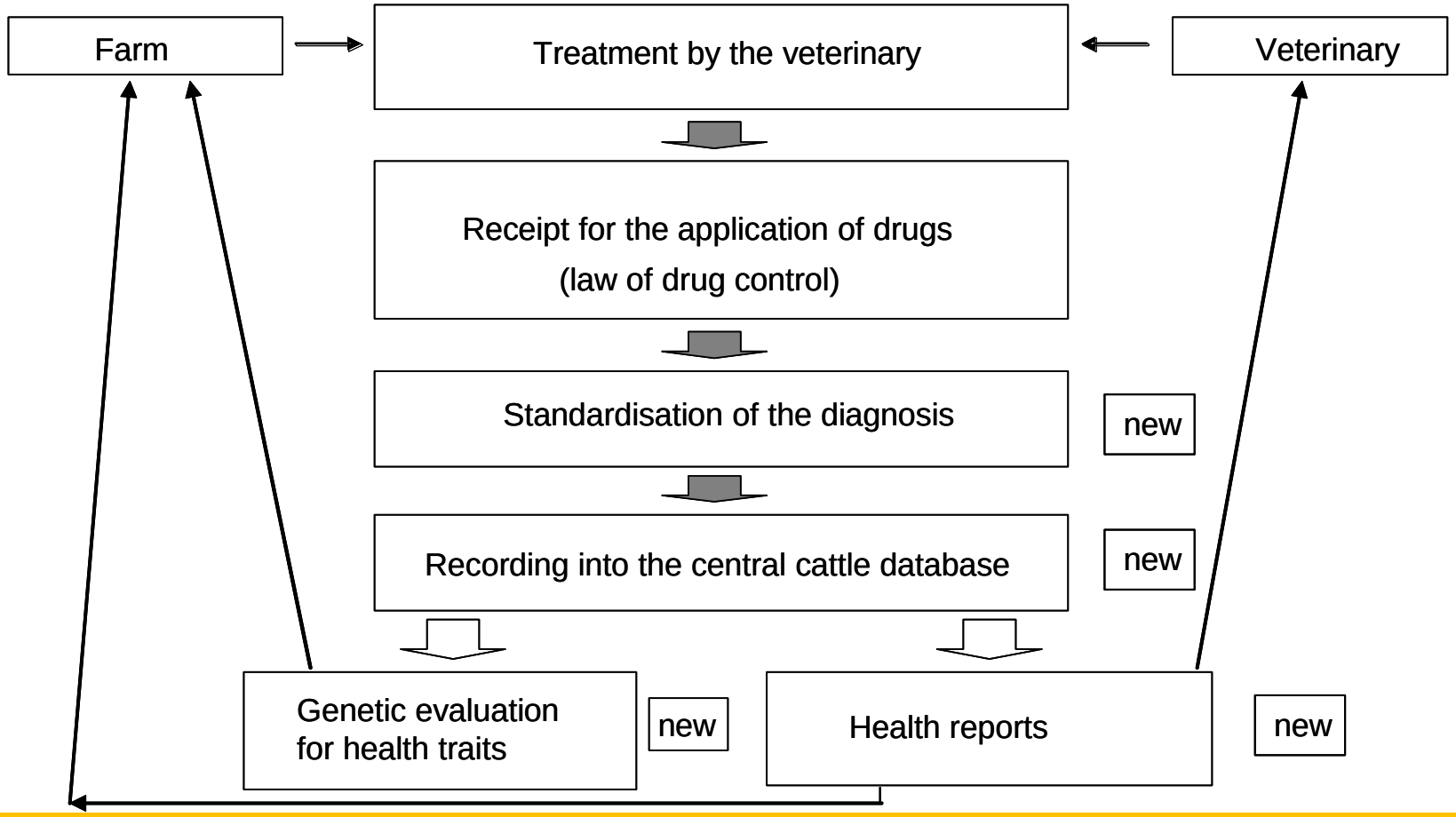
(Example Austria)



Different requirements on quality and quantity of diagnoses – impact on validation!

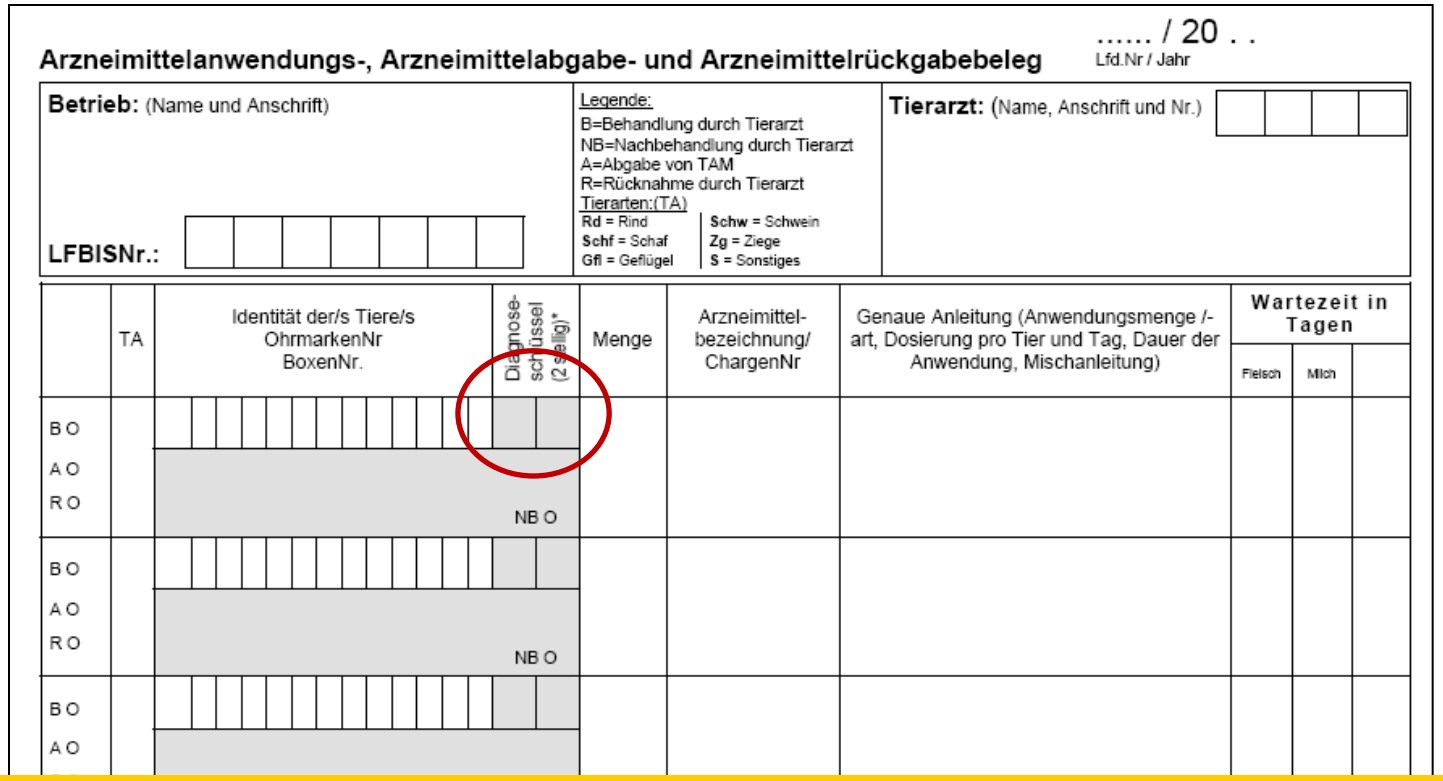
Health monitoring in Austria

Recording of data and backflow of information



2006 -2010 project, 2011 implementation into routine

New Receipt for the Application of Drugs



Arzneimittelanwendungs-, Arzneimittelabgabe- und Arzneimittelrückgabebeleg ..... / 20 . .
Lfd.Nr / Jahr

Betrieb: (Name und Anschrift)

LFBISNr.:

Legende:
B=Behandlung durch Tierarzt
NB=Nachbehandlung durch Tierarzt
A=Abgabe von TAM
R=Rücknahme durch Tierarzt
Tierarten:(TA)
Rd = Rind | Schw = Schwein
Schf = Schaf | Zg = Ziege
Gfl = Geflügel | S = Sonstiges

Tierarzt: (Name, Anschrift und Nr.)

	TA	Identität der/s Tiere/s OhrmarkenNr BoxenNr.	Diagnose-schlüssel (2-stellig)*	Menge	Arzneimittel-bezeichnung/ ChargenNr	Genaue Anleitung (Anwendungsmenge /-art, Dosierung pro Tier und Tag, Dauer der Anwendung, Mischanleitung)	Wartezeit in Tagen		
							Fleisch	Milch	
B O A O R O									
		NB O							
B O A O R O									
		NB O							
B O A O									

Recorded:

- Identity of the animal
- Identity of the farm
- Number of the veterinarian
- Code of diagnoses and date of diagnoses of first treatments only

Standardisation of Diagnoses

Austrian-wide code – published by the Ministry of Health in April 2006

Diagnoseschlüssel

Spezifische Kälberkrankheiten
11 Nabelentzündung
12 Nabelbruch
13 Sehnenkontraktur
14 Missbildungen
15 Ikterus haemolyticus neonatorum
16 Kälberdurchfall
17 andere Krankheiten des Kalbes

Erkrankungen des Verdauungstraktes
21 Durchfall
22 Tympanie
23 Pansenübersäuerung
24 Fremdkörpererkrankung
25 Labmagenverlagerung
26 Darmverschluss
27 andere Erkrankungen der Bauchhöhle
28 Erkrankungen der Maulhöhle
29 Erkrankungen der Speiseröhre

Stoffwechselkrankheiten
31 Gebärparese, Hypocalcämie
32 Tetanie
33 Azetonämie
34 andere Stoffwechselkrankheiten
35 Vergiftungen

Fruchtbarkeits-u. Abkalbest.
41 Gebärmutterentzündung
42 Stillbrunst, Azyklie
43 Ovarialzysten
44 Scheidenvorfall
45 Abortus und andere Störungen der Gravidität
46 Schwergeburt
47 Geburtsverletzungen
48 Nachgeburtsverhaltung
49 puerperale Erkrankungen

Eutererkrankungen
51 akute Euterentzündung
52 chronische Euterentzündung
53 Erkrankungen der Euter- und Zitzenhaut
54 Euterödem
55 Andere Eutererkrankungen
56 Prophylaktisches Trockenstellen

Klauen- und Gliedmaßenerkrank.
61 Panaritium, Mortellaro
62 Klauengeschwür, Krankheiten der Gelenke an den Klauen
63 Klauenrehe
64 Frakturen, Luxationen, andere Gliedmaßenverletzungen
65 Krankheiten von Muskeln und Sehnen
66 spastische Parese, Paralyse
67 Peritarsitis
68 Festliegen infolge Erkrankung des Bewegungsapparates
69 Krankheiten des Schwanzes

Erkrankungen der Atemwege
71 Erkrankungen der oberen Luftwege
72 Lungenentzündung
73 andere Lungenerkrankungen

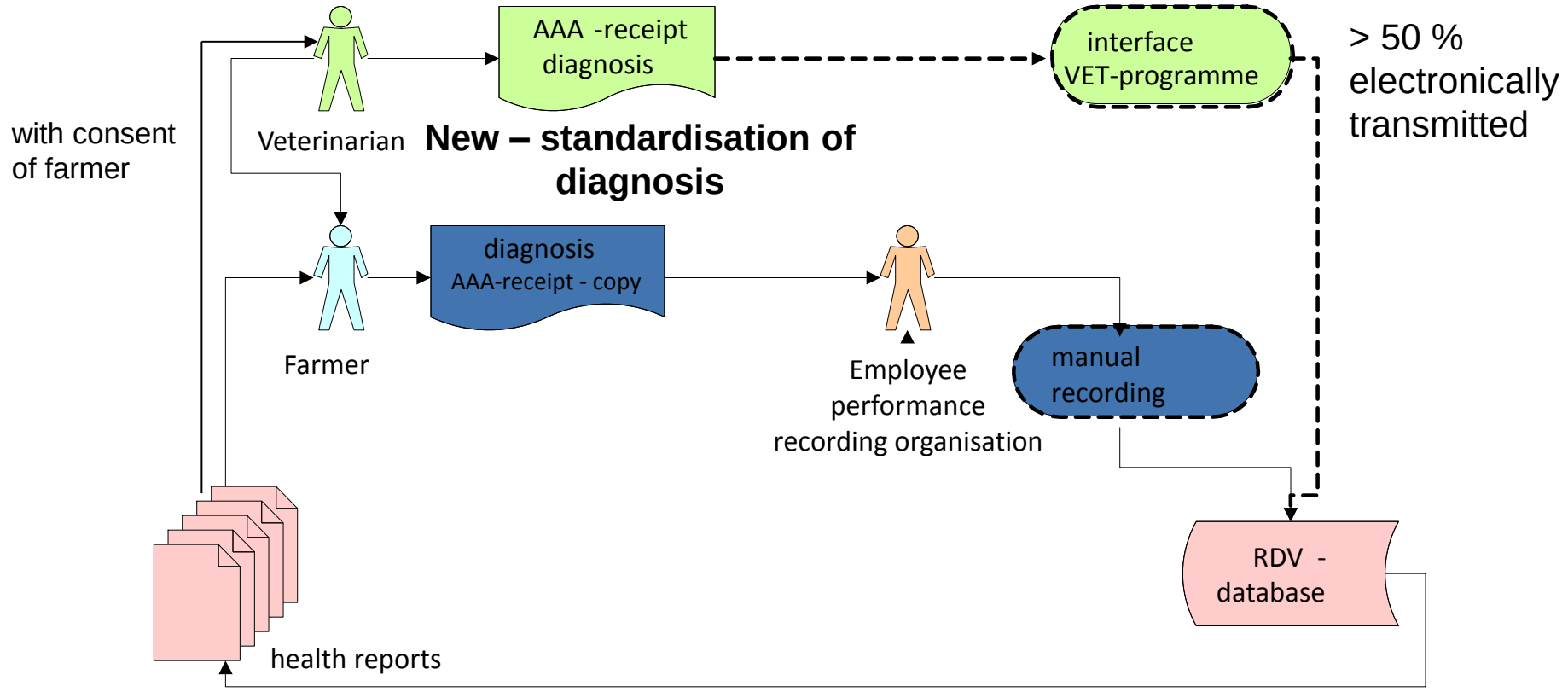
Herz-, Kreislauf- und Bluterkrank., Erkrankungen des Harntraktes
81 Herzerkrankungen
82 Septikämie, Anämie
83 Piroplasmose und andere Parasitosen des Blutes
84 Leukose
85 Erkrankungen der Gefäße und der Milz
86 Pyelonephritis
87 Erkrankungen der Harnblase

ZNS-Erkrankungen, Hauterkrankungen, Infektionen
91 ZNS-Erkrankungen
92 Erkrankungen der Sinnesorgane
93 Parasitosen und Infektionen der Haut
94 Erkrankung der Hörner
95 andere Hauterkrankungen
96 Allgemeininfektionen

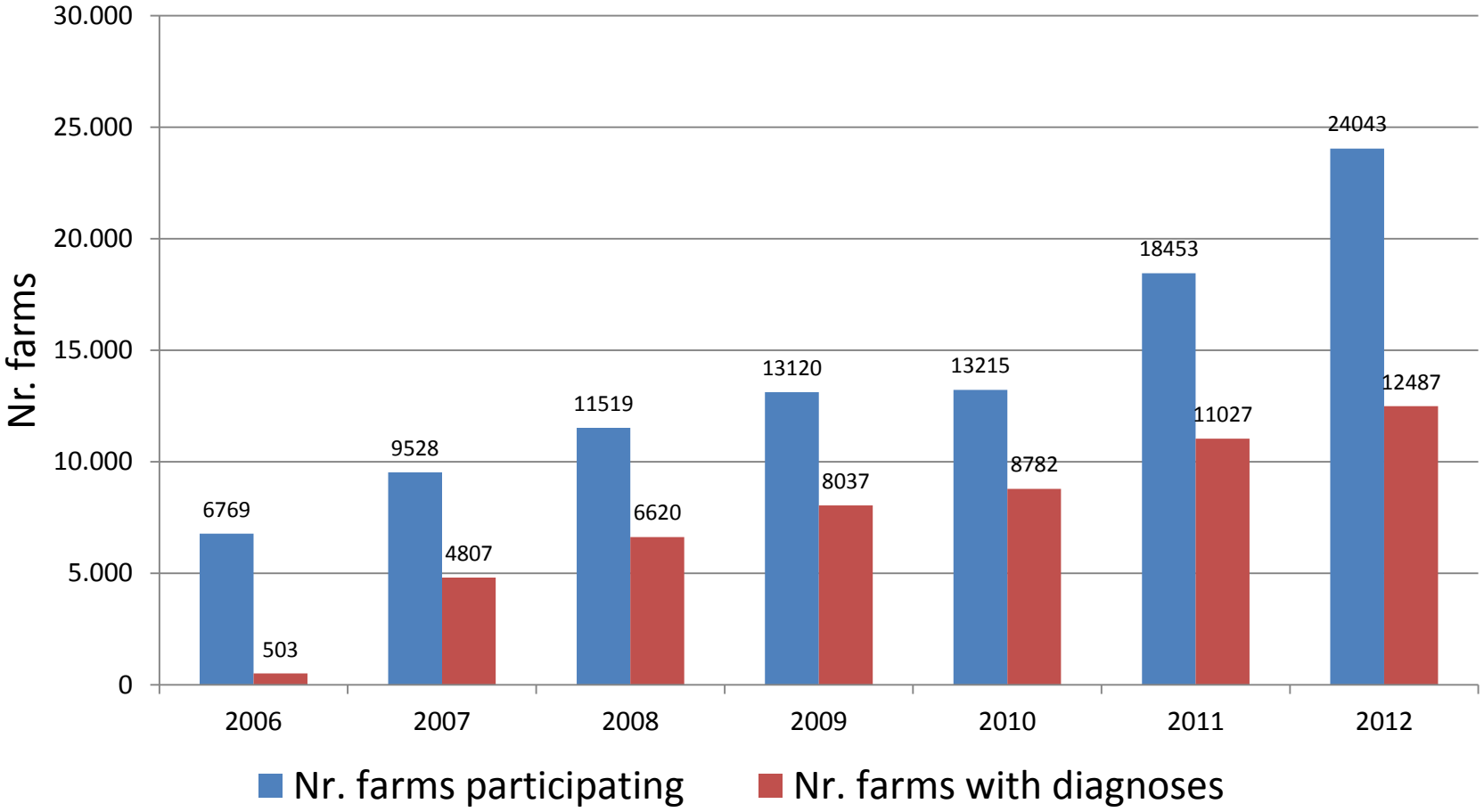
Sonstige Erkrankungen
01 Abmagerung, Kachexie
02 verminderte Fresslust, Inappetenz
03 Fieber, fieberhafte Allgemeinerkrank.
00 ohne Diagnose

... on-site diagnoses by veterinarians only
...currently no laboratory results.

Diagnoses Data - from Receipt to Data Base



Participation and percentage of farms with veterinary diagnoses



From around 55% of the farms participating in health monitoring veterinarian diagnoses are recorded presently. Percentage is increasing slowly.

Impact of recording on incidence rate (across breeds 2012)

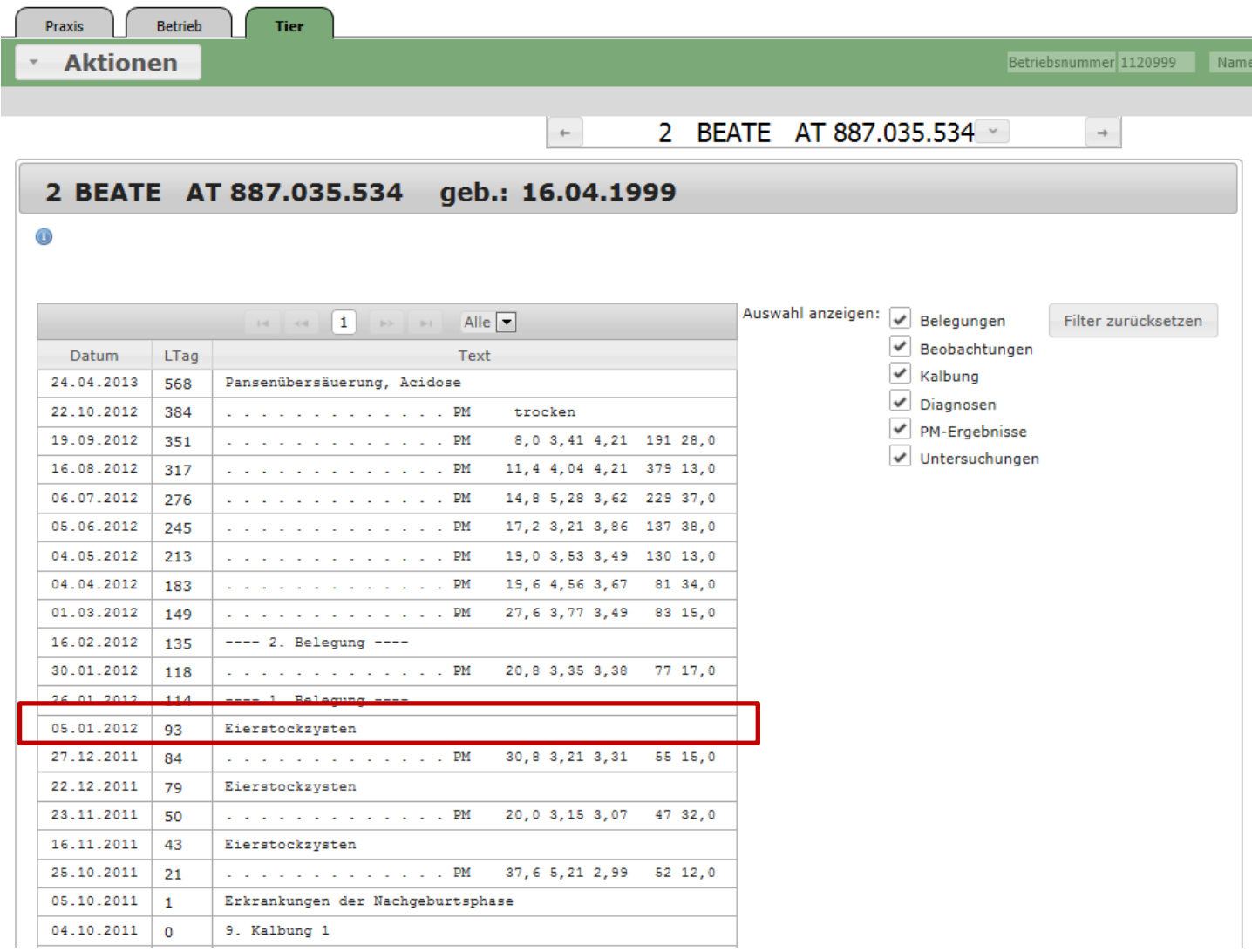
Traits	VET	PRO	Diff
Validated dairy cows	**36,756**	**110,597**	
Metabolic disorders	**5.65**	**4.08**	**-1.57**
Milkfever	4.36	3.10	-1.26
Ketosis	1.06	0.78	-0.29
Reproductive disorders	**26.14**	**18.54**	**-7.60**
Metritis	5.34	3.36	-1.97
Anoestrus	9.59	6.50	-3.09
Cystic ovaries	9.90	6.57	-3.33
Prolapse of vagina	0.13	0.08	-0.04
Retained placenta	3.34	3.49	0.15
Puerperal disorders	2.05	1.10	-0.95
Udder disorders	**18.16**	**15.73**	**-2.43**
Acute mastitis	12.72	11.37	-1.35
Chronic mastitis	5.84	4.65	-1.19
Hoof and claw disorders	**4.12**	**3.48**	**-0.64**
Panaritium, DD	2.30	1.78	-0.52
Hoof ulcer	0.92	0.96	0.04

VET: > 75% of data electronically transmitted to cattle data base directly by veterinarian
PRO: recording by employee of performance recording organisation

Plausibility checks

- **Plausibility checks before storage in data base** (e.g.: http://www.bmg.gv.at/cms/home/attachments/9/7/3/CH1141/CMS1271936439807/tgdkundm74200_46-ii-b-10-10gesundheitsprogrammrindprogramm.pdf)
  - correct ID of animal
  - farm has to participate in health monitoring
  - plausibility of date of diagnoses
  - plausibility of code of diagnoses
  - the same diagnoses per animal can only occur once a day,...
- **Plausibility check of first diagnoses**
  - e.g. retainced placenta can occur only once per lactation
  - possible periods between two consecutive first diagnoses
- **Plausibility checks by farmer and veterinarian**
  - provision of health reports und use within animal health programmes (farmers/veterinarians)

Plausibility check by farmers and veterinarians



Datum	LTag	Text
24.04.2013	568	Pansenübersäuerung, Acidose
22.10.2012	384	. . . . . PM trocken
19.09.2012	351	. . . . . PM 8,0 3,41 4,21 191 28,0
16.08.2012	317	. . . . . PM 11,4 4,04 4,21 379 13,0
06.07.2012	276	. . . . . PM 14,8 5,28 3,62 229 37,0
05.06.2012	245	. . . . . PM 17,2 3,21 3,86 137 38,0
04.05.2012	213	. . . . . PM 19,0 3,53 3,49 130 13,0
04.04.2012	183	. . . . . PM 19,6 4,56 3,67 81 34,0
01.03.2012	149	. . . . . PM 27,6 3,77 3,49 83 15,0
16.02.2012	135	---- 2. Belegung ----
30.01.2012	118	. . . . . PM 20,8 3,35 3,38 77 17,0
26.01.2012	114	---- 1. Belegung ----
05.01.2012	93	Eierstockzysten
27.12.2011	84	. . . . . PM 30,8 3,21 3,31 55 15,0
22.12.2011	79	Eierstockzysten
23.11.2011	50	. . . . . PM 20,0 3,15 3,07 47 32,0
16.11.2011	43	Eierstockzysten
25.10.2011	21	. . . . . PM 37,6 5,21 2,99 52 12,0
05.10.2011	1	Erkrankungen der Nachgeburtsphase
04.10.2011	0	9. Kalbung 1

Zeiler et al. 2013 / www.progesund.de

Diagnoseübersicht der letzten 3 Monate

Kühe

Nr.	Name	Lebensnummer	L.	Kalbung	Tg.	Diagnose	
6	KATHI	AT 152.954.616	3	05.12.12	69	12.02.13	Silent heat
17	SILWANA	AT 152.960.416	3	22.11.12	79	09.02.13	Ovarian cysts
27	FRADI	AT 516.220.847	8	22.02.13	13	07.03.13	Sole ulcer
48	FRANZA	AT 688.062.214	4	25.11.12	71	04.02.13	Clinical mastitis
53	ELEA	AT 495.663.418	1	01.11.12	76	16.01.13	Endometritis
60	FREIHEIT	AT 230.732.272	8	03.02.13	30	05.03.13	Clinical ketosis
					0	03.02.13	Hypocalcemia (milk fever)

Frischlaktierende Kühe (bis 100. Melktag) mit Eiweißgehalt <= 3 und/oder FEQ < 1,0 oder > 1,5

Nr.	Name	Lebensnummer	L.	Tg.	13.04.13 Eiw%	13.04.13 FEQ		03.03.13 Eiw%	03.03.13 FEQ	
20	FALKE	AT 563.115.317	2	18	3,40	**1,81**				
27	FRADI	AT 516.220.847	8	50	**2,83**	**1,75**		3,06	**2,13**	
31	FRIESE	AT 688.080.414	3	49	3,09	1,39		3,66	**0,99**	
60	FREIHEIT	AT 230.732.272	8	69	3,29	1,34	D	**2,83**	**2,14**	D

Kühe mit Zellzahl über 200.000 oder mit Euterdiagnosen (Schalmtest empfohlen)

Nr.	Name	Lebensnummer	L.	Tg.	13.04.13 Zellzahl	03.03.13 Zellzahl		24.01.13 Zellzahl
24	FROSCHI	AT 839.126.216	2	222	210	128		127
48	FRANZA	AT 688.062.214	4	139	191	140	D	38
	LILO	AT 295.048.914	4	287		232	D	324

http://www.bmg.gv.at/cms/home/attachments/9/7/3/CH1141/CMS1271936439807/tgdkundm74200_46-ii-b-10-10gesundheitsprogrammrindprogramm.pdf

Criteria for validation

Differentiate farms with low frequency versus farms with incomplete health data recording

- **Definition of valid observation period**
  - per farm
  - per cow - beginn / end of period per cow (time on farm with reliable registration of direct health data)
  - continuity of recording
- **Incidence rates of farms**
- Coding of diagnoses
- Use of data

Continuity of diagnoses by farms

avg. no. diagnoses per cow and year

	2006	2007	2008	2009	2010	2011	2012	
farm8557	0,65	1,28	0,81	0,69	0,81	1,12	1,27	ok
farm8558	0,00	0,00	0,00	0,00	0,00	0,06	0,07	
farm8559	1,24	1,94	1,11	1,35	1,18	0,77	0,52	ok
farm8560	0,15	0,07	0,17	0,18	0,00	0,07		?
farm8561	0,28	1,40	1,34	1,53	1,67	2,13	2,11	
farm8562						0,46	0,08	
farm8563	0,11	0,05	0,24	0,00	0,09	0,09		
farm8564	1,00	1,15	1,14	1,46	1,75	1,15	0,56	
farm8565	0,39	0,59	0,63	0,55	0,30	0,13		
farm8566	0,50	0,44	0,67	0,50	0,22	0,11	0,67	
farm8568	0,13	0,33	0,00	0,00	0,00			
farm8569						0,07	0,12	
farm8570	0,64	0,73	0,75	0,36	0,62	1,57	1,08	

Valid observation period

- **Definition of valid observation period**
  - per farm
    - continuity of recording (vet, employee performance recording organisation, farm) taken into consideration
  - per cow: begin / end of period per cow
    - time on farm with reliable registration of direct health data
    - respective type of use (dairy cow)

Incidence rate of farms

- Minimum requirement 0.1 first diagnoses per cow and year
- On average 0.5 first diagnoses per cow and year in validated data set for genetic evaluation
- 0.7 diagnoses per cow and year in dataset with elect. transmission of diagnoses directly by veterinarian
- Average incidence rate of farm by year calculated and relevant periods considered

Criteria for validation

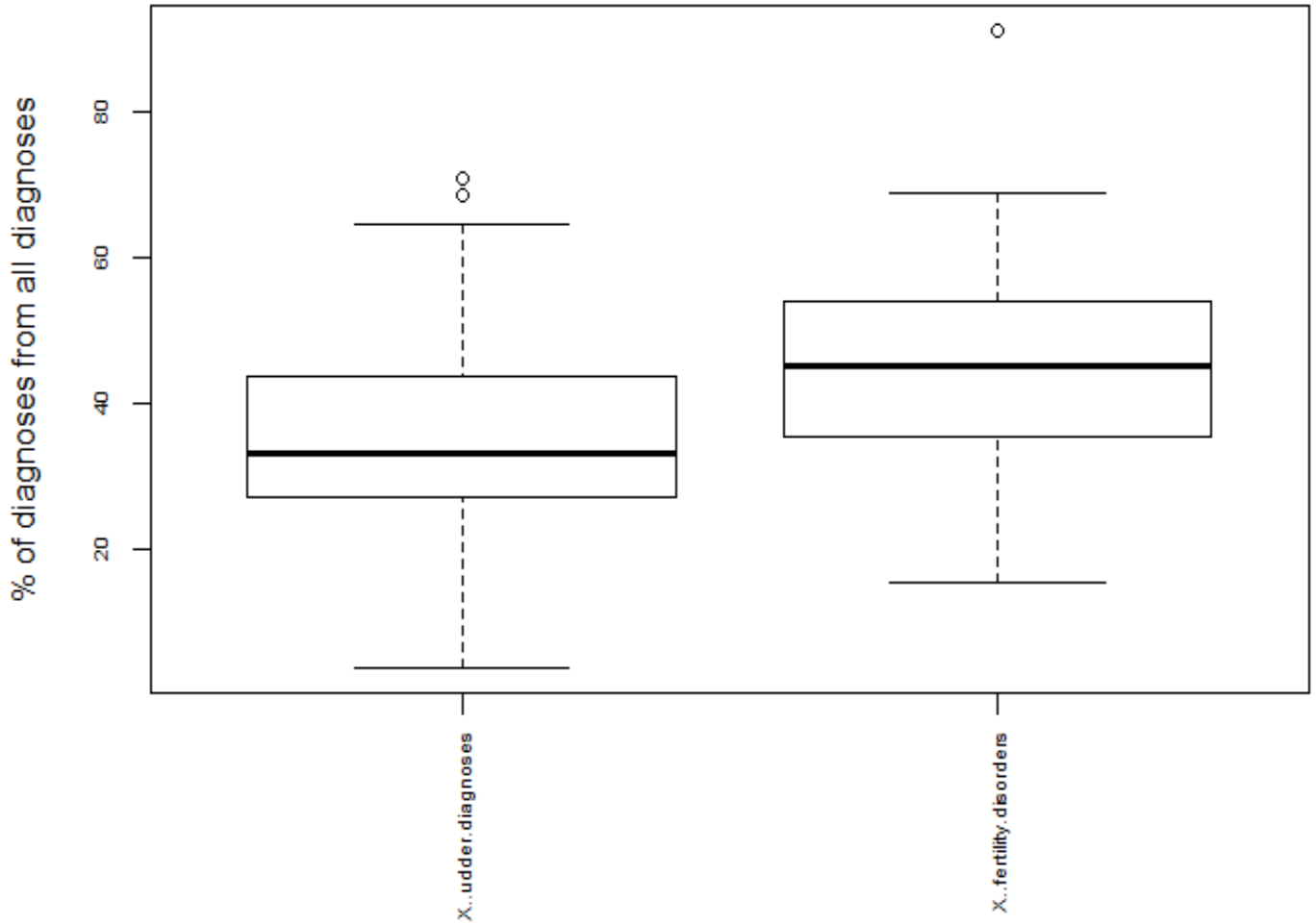
- Definition of valid observation period
  - per farm
  - per cow - begin / end of period per cow (time on farm with reliable registration of direct health data)
  - continuity of recording

- Incidence rates of farms
- **Coding of diagnoses**
- Use of data – impact on strictness

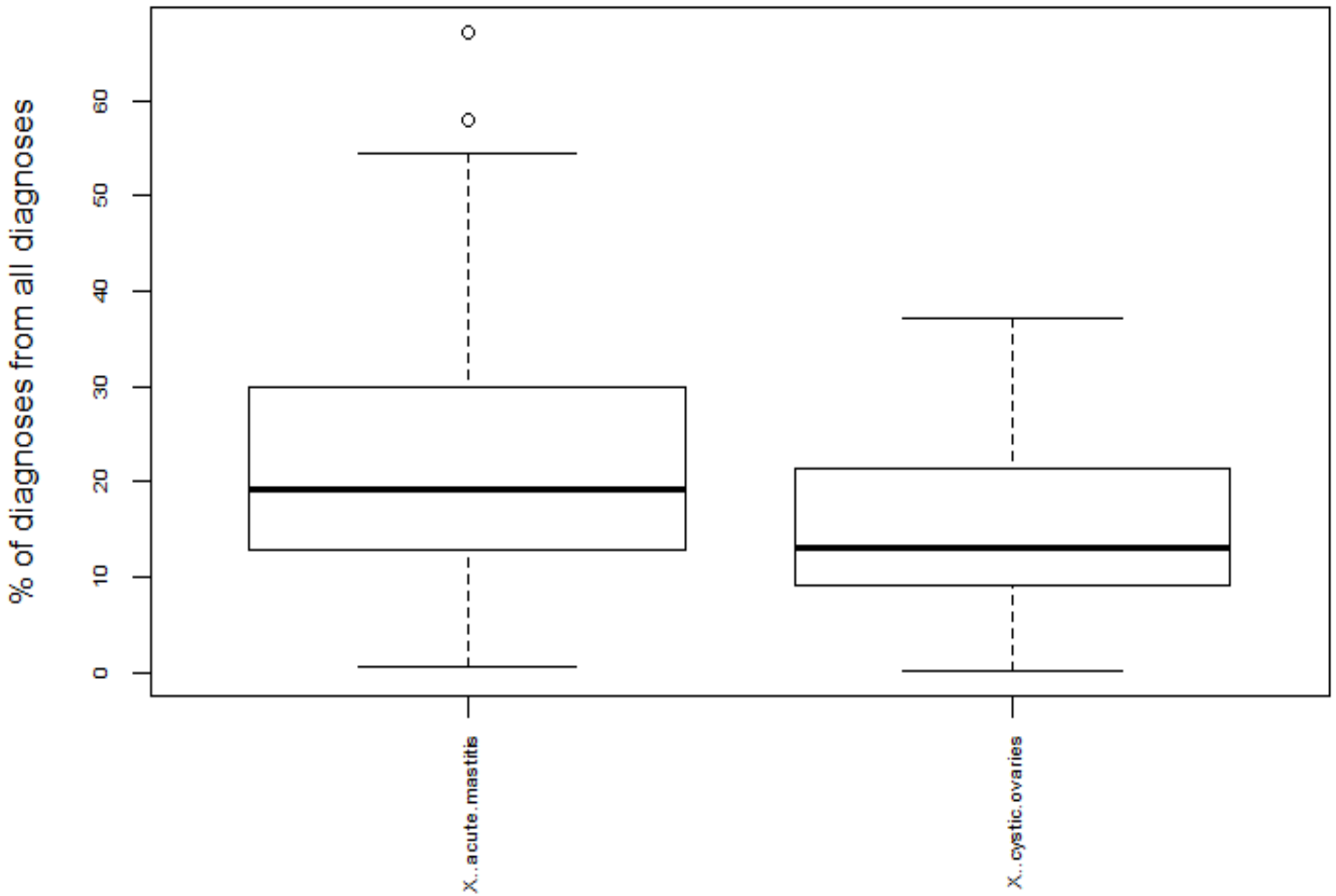
Distribution of most commonly recorded diagnoses of dairy cows, based on validated data for different lactations (lact), in percent (all breeds) Egger-Danner et al. 2012

	n=51,814	n=42,851	n=38,180	n=31,789	n=56,428
Diagnoses	% 1st lact	% 2nd lact	% 3rd lact	% 4th lact	% 5th lact+
Reproductive disorders	**42.56**	**45.81**	**42.55**	**40.28**	**37.32**
Metritis (MET)	6.84	6.19	5.65	5.64	5.46
Anoestrus (ESTRUS)	13.10	13.02	10.86	9.72	7.96
Cystic ovaries (CYST)	12.61	15.32	14.51	13.62	12.21
Retained placenta (RP)	5.49	6.55	6.61	6.57	6.82
Puerperal disorder (PUERP)	3.80	3.29	3.49	3.34	3.5
Udder disorders	**31.91**	**34.66**	**35.73**	**35.77**	**37.02**
Acute mastitis (AcM)	18.77	20.4	21.59	21.66	22.35
Chronic mastitis (CrM)	9.51	11.24	11.29	11.41	12.1
Digestive disorders	**2.96**	**3.90**	**7.07**	**10.24**	**12.48**
Milk fever (MF)	0.83	2.04	4.56	7.71	10.06
Ketosis (KET)	1.61	1.37	1.78	1.86	1.64
Hoof and claw disorders	**7.90**	**5.88**	**6.14**	**6.24**	**6.10**
Panaritium, DD (PAN/DD)	3.84	2.87	2.86	2.93	2.68
Hoof ulcer (HU)	1.82	1.41	1.55	1.57	1.74
Others	**14.67**	**9.75**	**8.51**	**7.47**	**7.08**

Variation in the percentage of fertility disorders and udder diseases per farm based on mainly electronically transmitted diagnostic data (VET)



Variation in the percentage of acute mastitis and cystic ovaries per farm based on mainly electronically transmitted diagnostic data (VET)



Coding of diagnoses - Summary

- Diagnoses are standardized by codes for 65 diagnoses
- Veterinarians working with practice management software often use a more detailed list of diagnoses for their own documentation. To link this a list of synonyms is provided.

- <u>Possible reasons for bigger differences in distribution of codes:</u>
  - Some veterinarian more specialized in certain diseases
  - veterinarian working more in prevention (e.g. ultrasound standard are standard)
  - incorrect diagnoses code or mistake in linkage of codes
  - different judgement of diagnoses
  - higher incidence of specific disease in certain farms / regions by time,...

Criteria for validation

- Definition of valid observation period
  - per farm
  - per cow - begin / end of period per cow (time on farm with reliable registration of direct health data)
  - continuity of recording

- Incidence rates of farms
- Coding of diagnoses
- **Use of data – impact on strictness**

Use of data

Feedback for farmers and veterinarians – reports/action lists e.g. for herd management

- only plausibility checks

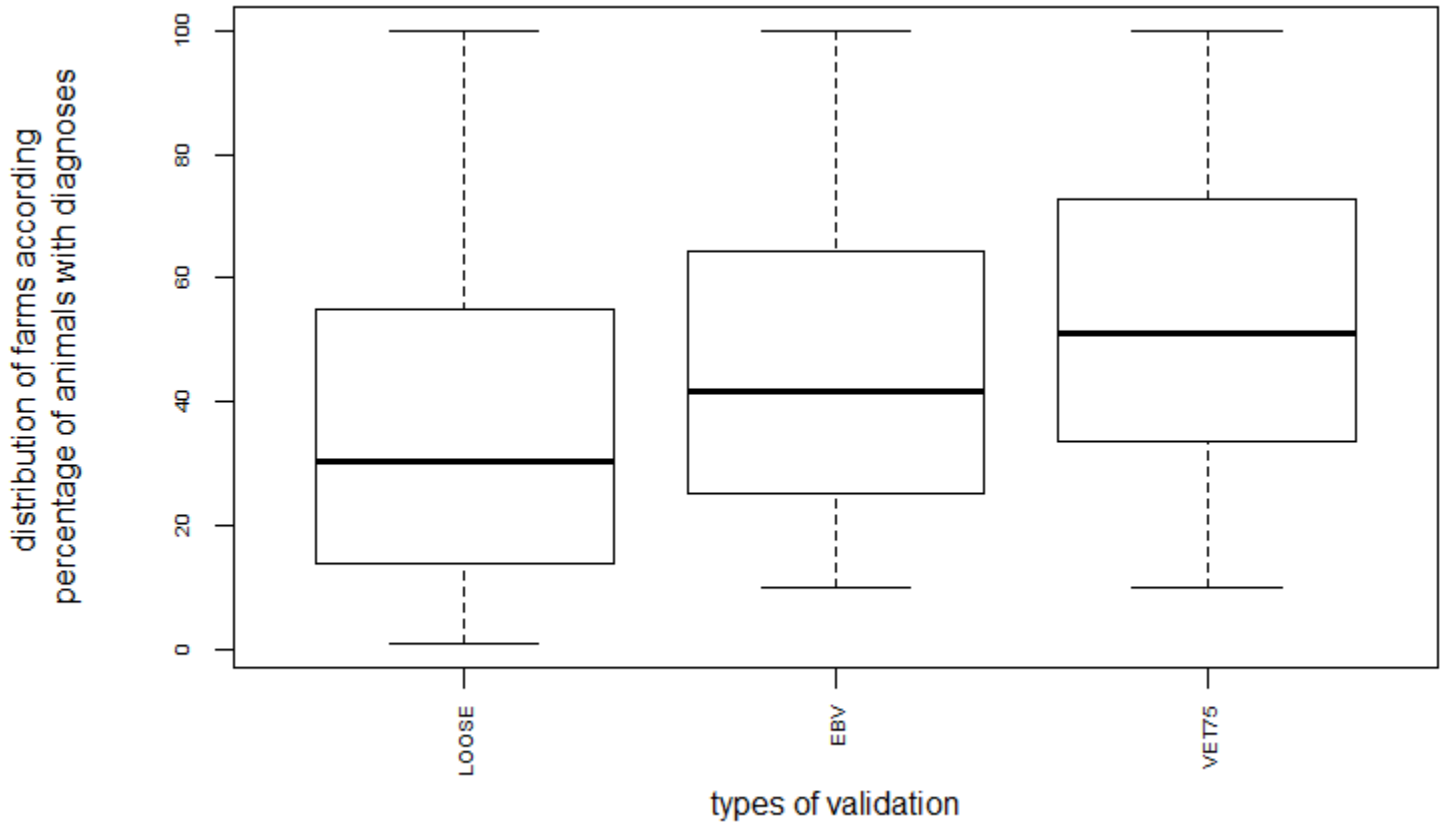
Genetic evaluation

- big amount of data needed (heritability versus quantity of data)
- correction for environmental effects
- combination of traits – higher frequencies and more stable breeding values

Benchmarks / monitoring of diseases

- observed incidences should reflect real incidence
- stringent data validation
- limited number of farms needed

Distribution of farms according to percentage of animals with diagnoses in 2012



6,111 farms **4,579 farms** **1,522 farms**

Measures to improve data quality

Recording of health data

- First priority - veterinarian diagnoses: base for joint use and use of synergies!
- farmer can record based on the same health key – distinction between diagnoses from veterinarians and observations of farmers in database
- since 2012: recording of observations around calving by performance recording organisations

Monitoring of recording

- further improvement of validation
- evaluation and feedback

Continuous information and motivation

Further development of benefits of recording
(together with partners from Bavaria and Baden-Württemberg)

- genetic evaluation
- further services for herd management for farmers
- online-platform for veterinarians,... (www.progesund.de)
- services in combination with regulations on documentation requirements

Conclusions

- Benefit/use of data is important for good data quality!
- Constant monitoring, feedback and training of people involved!
- Commitment of performance recording organisations is very important – link to central cattle data base very valuable!
- Emphasis on validation especially when system is newly established – takes time!
- Different requirements for validation depending on the use of data!
- Field data: chance of big amounts of data with limited costs of recording per diagnoses, but effort on validation!

Data from broad health monitoring systems are very valuable for herd health management, genetic evaluations and for surveillance purposes.

Acknowledgement

The Federation of Austrian Cattle Breeders thanks

•the Ministry of Health, the Ministry of Agriculture, Forestry, Environment and Water management and as well as the Federal States for their financial support.

•the partner organisations for their good cooperation and their valuable advise.


Thank you for your attention!