



**Universität für Bodenkultur Wien**  
University of Natural Resources  
and Life Sciences, Vienna

# Masterarbeit

## **Zuverlässigkeit der Erfassung von Lahmheit und weiteren tierwohlbezogenen Indikatoren in Milchviehbetrieben**

verfasst von

**Riegler Bianca, BSc**

im Rahmen des Masterstudiums  
**Nutztierwissenschaften**

zur Erlangung des akademischen Grades  
**Diplom-Ingenieurin**

Wien, August 2022

Betreut von:  
Univ.Prof. Dr.med.vet. Winckler Christoph  
Institut für Nutztierwissenschaften  
Department für Nachhaltige Agrarsysteme



## **Eidesstattliche Erklärung**

Ich versichere an Eides statt, dass ich diese Masterarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Gedanken, die im Wortlaut oder in grundlegenden Inhalten aus unveröffentlichten Texten oder aus veröffentlichter Literatur übernommen wurden, sind ordnungsgemäß gekennzeichnet, zitiert und mit genauer Quellenangabe versehen.

Die vorliegende Arbeit wurde bisher weder ganz noch teilweise in gleicher oder ähnlicher Form an einer Bildungseinrichtung als Voraussetzung für den Erwerb eines akademischen Grades eingereicht. Sie entspricht vollumfänglich den Leitlinien der Wissenschaftlichen Integrität und den Richtlinien der Guten Wissenschaftlichen Praxis.

Wien, 08.08.2022

Bianca RIEGLER (eigenhändig)

## **Danksagung**

An dieser Stelle möchte ich mich bei all jenen bedanken, die mich während der Anfertigung dieser Masterarbeit, aber auch während des Studiums unterstützt und motiviert haben.

Allen voran ist dies Herr Winckler Christoph und Frau Suntinger Marlene, welche meine Arbeit betreut und mir mit hilfreichen Anregungen und auch konstruktiver Kritik stets weitergeholfen haben.

Besonders bedanken möchte ich mich bei den Landwirt\*innen und Kontrollassistent\*innen, die sich bereit erklärt haben, am Projekt Klauen-Q-Wohl als Beobachter\*in teilzunehmen. Großer Dank gilt auch den Landwirt\*innen und deren Familien, die mich bei den Erhebungen auf ihren Betrieben herzlich empfangen haben.

Nicht zuletzt möchte ich mich auch bei meinen Freundinnen und Freunden bedanken, die mich während der Zeit des Studiums begleitet und dieses zu einer abwechslungsreichen und spannenden Erfahrung gemacht haben.

## Zusammenfassung

Lahmheit und damit in Verbindung stehende Merkmale wie Hautveränderungen sind häufig Bestandteil von Protokollen zur Beurteilung des Tierwohls auf landwirtschaftlichen Betrieben. Zur Unterstützung des Herdenmanagements auf Betriebsebene oder für die Berücksichtigung in der Zuchtwertschätzung mit dem Ziel der genetischen Verbesserung der Tiergesundheit auf Populationsebene ist jedoch eine zuverlässige Datenerfassung erforderlich. Um eine standardisierte Aufzeichnung zu erreichen, könnte die Bewertung im Rahmen der Milchleistungskontrolle durch den\*die Kontrollassistent\*in durchgeführt werden.

Ziel dieser Masterarbeit war es daher, (1) die Zuverlässigkeit der Beurteilung von Lahmheit und die damit zusammenhängenden Tierschutzindikatoren durch Landwirt\*innen und Kontrollassistent\*innen zu ermitteln und (2) die Übereinstimmung der Lahmheitsbeurteilung im Gang mit der Beurteilung von Lahmheit im Stehen im Hinblick auf die Identifikation hochgradig lahmer Tiere zu erheben. Zu diesem Zweck beurteilten fünf geschulte Kontrollassistent\*innen in 22 Milchviehbetrieben mit Melkständen an zwei aufeinanderfolgenden Milchleistungskontrollen die Lahmheit im Stehen (z. B. Belastung von nur einem Teil der Klaue, wiederholtes Anheben), abweichende Fußstellung, Verschmutzung der unteren Hinterbeine sowie haarlose Stellen, Verletzungen und Schwellungen am Sprunggelenk. Die Bewertung erfolgte im Melkstand parallel zur Milchleistungskontrolle. Zeitgleich erhoben eine externe Beurteilerin (Silberstandard) und die Landwirt\*innen selbst dieselben Merkmale, während sich die Kühe im Laufgang befanden. Sie erfassten zusätzlich Lahmheit im Gehen, Verschmutzungen am oberen Hinterbein und die Körperkondition (BCS). Die Übereinstimmung mit dem Silberstandard erreichte bei den Kontrollassistent\*innen einen PABAK von 0,24 bis 0,87 und bei den Landwirt\*innen von 0,32 bis 0,93. Es lag eine gute bis sehr gute Übereinstimmung ( $>0,6$ ) für Lahmheit, Verletzungen und Verschmutzung der oberen Hinterbeine, eine noch akzeptable Übereinstimmung (0,4-0,6) für abweichende Fußstellung und haarlose Stellen und eine unbefriedigende Übereinstimmung ( $<0,4$ ) für Verschmutzung am unteren Hinterbein und BCS vor. Es lag eine noch akzeptable Übereinstimmung ( $\geq 0,4$ ) bei 30 bis 100 % der Kontrollassistent\*innen und 58 bis 100 % der Landwirt\*innen vor. Im Hinblick auf die Erkennung hochgradig lahmer Tiere wies das Merkmal Lahmheit im Stand zwar eine hohe Spezifität auf, aber die Sensitivität war niedrig.

Die Studie zeigte, dass geschulte Landwirt\*innen in der Lage waren, Lahmheits- und andere Tierwohlmerkmale meist mit akzeptabler bis hoher Übereinstimmung zu beurteilen. Der Melkstandtyp beeinflusste die Durchführbarkeit der Erhebung erheblich. Für die Zukunft stellt sich die Frage, ob es sinnvoll ist, Lahmheit im Gehen im Laufgang durch geschulte Beobachter\*innen erheben zu lassen, da durch die Lahmheitsbeurteilung im Stand viele hochgradig lahme Tiere nicht erkannt wurden.

## Abstract

Lameness and related measures such as skin alterations have been repeatedly part of on-farm welfare assessment protocols. However, to support herd management on farm level or for the consideration in breeding value estimation with the aim to genetically improve animal health on population level, reliable data recording is needed. To achieve standardised recording the assessment might be carried out by persons of milk recording organisations in the course of milk performance recording visits. The objectives of this study were thus to determine (1) the inter-observer reliability of assessing lameness and related welfare indicators through farmers and milk recording personnel and (2) the agreement between gait scoring and a standing score for lameness assessment with regard to identifying severely lame animals. For this purpose, on 22 dairy farms using milking parlours, five trained persons of milk recording organisations recorded the standing lameness score (e.g. exposure of only one part of the claw, resting a foot), claw position as well as cleanliness of the lower hind legs and hairless spots, injuries and swellings of the hock on two consecutive milk performance recording visits. The assessment was done in the milking parlour parallel to test-day sampling. In a timely manner, an external person (silver standard) and the farmers themselves assessed the same traits while cows were in the loose housing. They additionally scored locomotion, cleanliness of the upper hind legs and body condition score (BCS). Agreement with the silver standard ranged from PABAK 0.24 to 0.87 across control assistants and from 0.32 to 0.93 across farmers. It was high to very high ( $>0.6$ ) for gait score, injuries and cleanliness of the upper hind legs, still moderate (0.4-0.6) for claw position, and hairless spots and low ( $<0.4$ ) for cleanliness of the lower hind legs and BCS. There was still moderate agreement ( $\geq 0.4$ ) for 30 to 100 % of the control assistants and 58 to 100 % of the farmers. Using a standing score specificity of detecting severely lame animals was high but sensitivity low.

The study showed that trained farmers were able to score lameness and other animal welfare traits mostly with moderate to high agreement. The milking parlour type considerably affected feasibility of the assessment. The question for future is, whether it makes sense if other trained observers assess measures like lameness in the loose housing because the standing score did not detect many severely lame animals.

# Inhaltsverzeichnis

Eidesstattliche Erklärung .....	1
Danksagung .....	2
Zusammenfassung .....	3
Abstract .....	4
Inhaltsverzeichnis .....	5
Abbildungsverzeichnis .....	7
Tabellenverzeichnis .....	9
Abkürzungen .....	11
1 Einleitung und Problemstellung .....	12
2 Ziele und Forschungsfragen .....	14
3 Literatur .....	15
3.1 Berücksichtigung von Lahmheit und anderen Merkmalen des Tierwohls in der Zuchtwertschätzung.....	15
3.2 Tierbezogene Indikatoren .....	17
3.2.1 Auswahl der Merkmale .....	17
3.2.2 Validität .....	17
3.2.3 Zuverlässigkeit der Erfassung.....	18
3.2.4 Durchführbarkeit/Praktikabilität .....	23
3.3 Beobachter*innentraining .....	25
4 Material, Methode und Tiere.....	28
4.1 Hintergrund Projekt-Klauen-Q-Wohl .....	28
4.2 Auswahl der Beobachter*innen und Betriebe .....	28
4.2.1 Beobachter*innen .....	28
4.2.2 Pilotbetriebe .....	28
4.3 Merkmale.....	29
4.3.1 Lahmheit im Gehen (AssureWel) .....	29
4.3.2 Lahmheit im Stand (Brinkmann et al., 2016) .....	30
4.3.3 Abweichende Fußstellung (Hulek, 2015) .....	31
4.3.4 Schäden am Sprunggelenk - Integumentschäden (Brinkmann et al., 2016) .....	32
4.3.5 Verschmutzung (Kofler und Winckler, 2018).....	34
4.3.6 Körperkondition - Body Condition Score .....	35
4.4 Erhebungsprotokolle für die Merkmalerhebung .....	37
4.5 Training der externen Beurteilerin .....	39
4.6 Schulungen .....	39
4.7 Durchführung der Erhebungen .....	42
4.7.1 Tiere.....	43
4.7.2 Vorgehensweise Kontrollassistent*in .....	43
4.7.3 Vorgehensweise Landwirt*in und externe Beurteilerin.....	44
4.8 Auswertung der Daten und Interpretation der Ergebnisse.....	45
4.8.1 Berechnung der Prävalenzen (DG 1 auf DG 2).....	46
4.8.2 Berechnung der Beobachter*innenübereinstimmung.....	46
4.8.3 Abweichung in %-Punkten auf Herdenebene von der externen Beurteilerin (Abw. in %-Punkte).....	48
4.8.4 Zuverlässigkeit der Identifikation hochgradig lahm eingestufte Tiere durch Standbeurteilung durch KA*in/LW*in .....	48
4.8.5 Übermittlung der Ergebnisse an die Landwirt*innen .....	50
4.8.6 Befragung der Teilnehmer*innen .....	53

5	Ergebnisse .....	54
5.1	Übereinstimmung externe Beurteilerin mit erfahrenem Beurteiler .....	54
5.2	Auftretenshäufigkeiten (Prävalenzen) im ersten und im zweiten Durchgang (Erhebung durch externe Beurteilerin) .....	56
5.3	Übereinstimmung der Stakeholdergruppen mit der externen Beurteilerin ...	58
5.3.1	Landwirt*innen .....	58
5.3.2	Kontrollassistent*innen .....	60
5.3.3	Abweichung in %-Punkten von den Ergebnissen der externen Beurteilerin (Abw. in %-Punkte).....	61
5.4	Erfassung ‚hochgradig lahmer‘ Tiere: Beurteilung im Stand (KA*in bzw. LW*in) vs. Gangbeurteilung (externe Beurteilerin).....	68
5.5	Aussagen der Teilnehmer*innen .....	70
5.5.1	Aussagen Landwirt*innen und Kontrollassistent*innen beim leitfadengestützten Interview .....	70
5.5.2	Aussagen bei den Abschlussveranstaltungen zur den Piloterhebungen im Projekt Klauen-Q-Wohl.....	71
6	Diskussion .....	72
6.1	Schulung und Beobachter*innentraining .....	72
6.2	Auftretenshäufigkeiten (Prävalenzen) im ersten und im zweiten Durchgang (Erhebung durch externe Beurteilerin) .....	73
6.3	Übereinstimmung externe Beurteilerin mit erfahrenem Beurteiler .....	74
6.4	Übereinstimmung der Stakeholdergruppen mit der externen Beurteilerin ...	74
6.4.1	Landwirt*innen .....	75
6.4.2	Kontrollassistent*innen .....	75
6.5	Erfassung ‚hochgradig lahmer‘ Tiere: Beurteilung im Stand (KA*in bzw. LW*in) vs. Gangbeurteilung (externe Beurteilerin).....	76
6.6	Aussagen der Teilnehmer*innen zum Projekt Klauen-Q-Wohl.....	77
7	Literaturverzeichnis .....	78
	Anhang .....	88

## Abbildungsverzeichnis

Abb. 1: Vorgeschlagenes Protokoll nach Mullan et al. (2011) für das Testen der Beobachter*innen-übereinstimmung (IOR) von Beurteiler*innen der FA (farm assurance).....	27
Abb. 2: Merkmal Lahmheit im Stand – lahm (1). Quelle: © Solveig-March. <a href="https://www.bio-austria.at/lahmheit/">https://www.bio-austria.at/lahmheit/</a> .....	30
Abb. 3: Merkmal „Gute/akzeptable Fußstellung“ (links - 0) und „abweichende Fußstellung“ (rechts - 1) an den hinteren Extremitäten beim Rind. Die rote Linie am Fersenbein dient als Hilfsmittel zu Erhebung der abweichenden Fußstellung. Quelle: © Riegler.....	31
Abb. 4: Merkmal Schäden am Sprunggelenk – Schwellung rechts (1) .....	32
Abb. 5: Merkmal Schäden am Sprunggelenk – Haarlose Stellen (1).....	32
Abb. 6: Merkmal Schäden am Sprunggelenk – Wunde verkrustet (1). .....	33
Abb. 7: Merkmal Schäden am Sprunggelenk – Wunde frisch (1). Quelle: © BOKU Winckler.....	33
Abb. 8: Merkmal Schäden am Sprunggelenk – Haarlose Stelle auf der linken Körperseite (1). Die Markierung zeigt den untersuchten Bereich (Außenseite des linken Sprunggelenks sowie Innenseite des gegenüberliegenden, rechten Sprunggelenks). Quelle: © BOKU Winckler.....	33
Abb. 9: Merkmal Verschmutzung – Sauber (0) - Die Markierung zeigt das obere und untere Hinterbein auf der linken Körperseite. ....	34
Abb. 10: Merkmal Verschmutzung – Verschmutzt am linken oberen Hinterbein (1). Verschmutzt am linken unteren Hinterbein (1), sowohl auf der linken Außenseite als auch auf der gegenüberliegenden Innenseite Die Markierung zeigt die Abgrenzung zwischen dem oberen und unteren Hinterbein auf der linken Körperseite. ....	34
Abb. 11: 4 Körperbereiche der optischen Bewertung von Körperkondition.....	36
Abb. 12: BCS – Optische Beurteilung. Optische Bewertung der subkutanen Fettauflage an vier Körperregionen. Punktevergabe 1,0 – 5,0 in 0,5 Punkt-Schritte. ....	36
Abb. 13: Schulungsraum der zweiten Schulung in der Hblfa Francisco Josephinum. Quelle: © Riegler .....	41
Abb. 14: Praktischer Teil der ersten Schulung am Lehrbetrieb der Lfs Hafendorf. Quelle: © Suntinger .....	41
Abb. 15: Externe Beurteilerin - Erklärung der Merkmale direkt am Tier am Lehrbetrieb der Hblfa Francisco Josephinum. Quelle: © Suntinger .....	41
Abb. 16: Praktischer Teil der zweiten Schulung am Lehrbetrieb der Hblfa Francisco Josephinum. Quelle: © Suntinger .....	41
Abb. 17: Zeitlicher Ablauf der Durchführung des Beobachterabgleichs, der Schulungen und der Erhebungen auf den Pilotbetrieben .....	42
Abb. 18: Auftretenshäufigkeit in % für die Merkmale Lahmheit im Gehen (LH_Gang), Lahmheit hochgradig (LH_Hochgradig), Lahmheit im Stand (LH_Stand) und abweichende Fußstellung (Abw. Fußst.) auf Basis der Erhebung durch die externe Beurteilerin; x=Mittelwert .....	56
Abb. 19: Auftretenshäufigkeit in % für die Merkmale haarlose Stellen (Haarlos), Wunde, Schwellung, Verschmutzung im oberen hinteren Bereich (Verschm_OH) und Verschmutzung im unteren hinteren Bereich (Verschm_UH) auf Basis der Erhebung durch die externe Beurteilerin; x=Mittelwert .....	57
Abb. 20: Abweichung in %-Punkte der Landwirt*innen/Kontrollassistent*innen zur externen Beurteilerin für die Merkmale Lahmheit gesamt/Gangbeurteilung	

(LH_Gang), Lahmheit hochgradig/Gangbeurteilung (LH_Hochgradig), Lahmheit/Beurteilung im Stand (LH_Stand) sowie abweichende Fußstellung (Abw. Fußst.) in Durchgang 1; $x$ =Mittelwert.....	61
Abb. 21: Abweichung in %-Punkte der Landwirt*innen/Kontrollassistent*innen zur externen Beurteilerin für die Merkmale Lahmheit gesamt/Gangbeurteilung (LH_Gang), Lahmheit hochgradig/Gangbeurteilung (LH_Hochgradig), Lahmheit/Beurteilung im Stand (LH_Stand) sowie abweichende Fußstellung (Abw. Fußst.) in Durchgang 2; $x$ =Mittelwert.....	62
Abb. 22: Abweichung in %-Punkte der Landwirt*innen/Kontrollassistent*innen zur externen Beurteilerin für die Merkmale haarlose Stellen (Haarlos), Wunde, Schwellung, Verschmutzung im oberen hinteren Bereich (Verschm_OH) und Verschmutzung im unteren hinteren Bereich (Verschm_UH) in Durchgang 1; $x$ =Mittelwert .....	63
Abb. 23: Abweichung in %-Punkte der Landwirt*innen/Kontrollassistent*innen zur externen Beurteilerin für die Merkmale haarlose Stellen (Haarlos), Wunde, Schwellung, Verschmutzung im oberen hinteren Bereich (Verschm_OH) und Verschmutzung im unteren hinteren Bereich (Verschm_UH) in Durchgang 2; $x$ =Mittelwert .....	63
Abb. 24: Mittlere Abweichungen in %-Punkte von KA*in 1, KA*in 2, KA*in 3 und KA*in 4 zur externen Beurteilerin der Merkmale LH_Stand, abw. Fußst., Verschm_UH, Haarlos, Wunde und Schwellung in Durchgang 1 .....	67
Abb. 25: Mittlere Abweichungen in %-Punkte von KA*in 1, KA*in 2, KA*in 3 und KA*in 4 zur externen Beurteilerin der Merkmale LH_Stand, abw. Fußst., Verschm_UH, Haarlos, Wunde und Schwellung in Durchgang 2 .....	67
Abb. 26: Gegenüberstellung Sensitivität und Spezifität der Übereinstimmung der Beurteilung hochgradiger Lahmheit zwischen Gangbeurteilung (LH_Hochgradig) externer Beurteilerin und Standbeurteilung (LH_Stand) durch LW*in (N=18)/KA*in (N=6), die Übereinstimmung wird als Spezifität und Sensitivität abgebildet.....	69
Abb. 27: Schulungsunterlagen; laminierte Erhebungskarte von den Merkmalen Haarlos, Wunden, Schwellung und Verschmutzung für die Erhebung direkt im Stall; Seite 1 .....	90
Abb. 28: Schulungsunterlagen; laminierte Erhebungskarte von den Merkmalen Lahmheit Gangbeurteilung, Lahmheit Beurteilung im Stand und abweichende Fußstellung für die Erhebung direkt im Stall; Seite 2.....	90

## Tabellenverzeichnis

Tab. 1: Beobachter*innenübereinstimmung (Inter-Beobachter-Reliabilität) für das 5-Punkte-Lahmheits Score und nach dem Zusammenlegen der einzelnen Scores in ein 4-, 3-, und 2-Punkte-Lahmheits-Score; ausgedrückt in Weighted KAPPA (kw), KAPPA (k) sowie Prozent-Übereinstimmung (PA) (Schlageter-Tello et al., 2014) .....	22
Tab. 2: Erhebungsprotokoll für Landwirte*innen und externe Beurteilerin .....	37
Tab. 3: Erhebungsprotokoll für Kontrollassistent*innen .....	38
Tab. 4: Beispielhafte Darstellung des Schulungsergebnisses eine*r Teilnehmer*in der Schulung .....	40
Tab. 5: Erfassung der Ergebnisse für die Berechnung des Kappas (a-d: beobachtete Häufigkeiten) .....	47
Tab. 6: Beispiel für Spezifität vs. Sensitivität sowie positive (PPV) und negative predicitive value (NPV) am Betrieb 1 bei der LW*in Erhebung.....	49
Tab. 7: Beispiel für die Darstellung und Interpretation der Ergebnisse für den*die Landwirt*in.....	51
Tab. 8: Ergebnisse des Beobachter*innenabgleichs zwischen der externen Beurteilerin (Masterstudentin) und einem erfahrenen Beurteiler (Silberstandard) im März 2019 (N=37) sowie Oktober 2019 (N=36); die Beobachter*innenübereinstimmung wird als prozentuale Übereinstimmung (%ÜE), Prevalence Adjusted Bias Adjusted Kappa (PABAK) sowie Abweichung in der Herdenprävalenz (%-P. Abw.) ausgedrückt; grün (%ÜE 100-80 %/PABAK >0,6) = sehr gute bis gute Übereinstimmung, orange (%ÜE 79-70 %/PABAK (0,60-0,41) = Übereinstimmung noch akzeptabel, rot (%ÜE <70 %/PABAK 0,40-0,21) = Übereinstimmung unbefriedigend .....	54
Tab. 9: Abweichung in %-Punkten der externen Beurteilerin vom Silberstandard für die (gerundete) Gesamtnote BCS (Durchschnitt der 4 Körperregionen) .....	55
Tab. 10: Ergebnisse des Beobachter*innenabgleichs zwischen Landwirt*innen und der externen Beurteilerin (Masterstudentin) in DG 1 und DG 2; die Beobachter*innenübereinstimmung wird als prozentuale Übereinstimmung (%ÜE), Prevalence Adjusted Bias Adjusted Kappa (PABAK) sowie Abweichung in der Herdenprävalenz (%-P. Abw.) ausgedrückt; grün (%ÜE 100-80 %/PABAK >0,6) = sehr gute bis gute Übereinstimmung, orange (%ÜE 79-70 %/PABAK (0,60-0,41) = Übereinstimmung noch akzeptabel, rot (%ÜE <70 %/PABAK 0,40-0,21) = Übereinstimmung unbefriedigend .....	58
Tab. 11: Ergebnisse des Beobachter*innenabgleichs zwischen Kontrollassistent*innen (N = 22) und der externen Beurteilerin (Masterstudentin) in DG 1 und 2; die Beobachter*innenübereinstimmung wird als prozentuale Übereinstimmung (%ÜE), Prevalence Adjusted Bias Adjusted Kappa (PABAK) sowie Abweichung in der Herdenprävalenz (%-P. Abw.) ausgedrückt; grün (%ÜE 100-80 %/PABAK>0,6) = sehr gute bis gute Übereinstimmung, orange (%ÜE 79-70 %/PABAK (0,60-0,41) = Übereinstimmung noch akzeptabel, rot (%ÜE <70 %/PABAK 0,40-0,21) = Übereinstimmung unbefriedigend .....	60
Tab. 12: Zusammenfassung der mittleren Abweichung in %-Punkten der LW*innen und KA*innen zur externen Beurteilerin von den Merkmalen LH_Gang, LH_Hochgradig, LH_Stand, abw. Fußst., Haarlos, Wunde, Schwellung, Verschm_OH und Verschm_UH; (+): positive Abweichung = höhere Prävalenz, (-): negative Abweichung = geringere Prävalenz, (*/**): signifikante Abweichung (p<0,05/0,01) .....	64

Tab. 13: Mittlere Abweichungen (%-Punkte) zur externen Beurteilerin für die Kontrollassistent*innen 1, 2, 3 und 4 (KA*in 1,2, 3, 4) in DG 1 und 2 für die Merkmale Lahmheitsbeurteilung im Stand (LH_Stand), abweichende Fußstellung (abw. Fußst.), haarlose Stellen (Haarlos), Wunde, Schwellung und Verschmutzungen im unteren hinteren Bereich (Verschm_UH) .....	66
Tab. 14: Spezifität, Sensitivität sowie positive predictive value (PPV) und negative predictive value (NPV) für die Erfassung hochgradiger Lahmheit mittels Standbeurteilung (LH_Stand) durch KA*in/LW*in gegenüber der Gangbeurteilung (LH_Hochgradig) durch die externe Beurteilerin .....	68
Tab. 15: Unterschiede der Auftretenshäufigkeit in Prozent der Merkmale vom 1. auf den 2. Durchgang vom Ergebnis der externen Beurteilerin .....	88
Tab. 16: Mittlere Abweichung in %-Punkte aller LW*innen zur externen Beurteilerin mit p-Werte und Prävalenzen (Auftretenshäufigkeiten) .....	88
Tab. 17: Mittlere Abweichung in %-Punkte aller KA*innen zur externen Beurteilerin mit p-Werte und Prävalenzen (Auftretenshäufigkeiten) .....	89

## Abkürzungen

%ÜE _____	Prozent Übereinstimmung zur externen Beurteilerin
Abw. Fußst. _____	abweichende Fußstellung, Kühe mit ‚zu weiter‘ Fußstellung (Note 0/1)
%-P. Abw. _____	Abweichung in der Herdenprävalenz, Abweichungen in Prozent Punkte von der externen Beurteilerin
BCS _____	Body Condition Score (Note 1 bis 5 in 0,5-Punkte Schritte)
EB _____	Externe Beurteilerin
Haarlos _____	Haarlose Stellen im Bereich des Sprunggelenkes am rechten und linken Hinterbein (Note 0/1)
IOR _____	Inter Beobachter Reliability
KA*in _____	Kontrollassistent*in
KA*innen _____	Kontrollassistent*innen
LH_Gang _____	Lahmheitsbeurteilung im Gang (Note 0/1/2), Lahmheit gesamt/Gangbeurteilung
LH_Hochgradig _____	Kühe mit Lahmheitsscore (Gang) von 2, Lahmheit hochgradig/Gangbeurteilung
LH_Stand _____	Lahmheitsbeurteilung im Stand (Note 0/1)
LW*in _____	Landwirt*in
LW*innen _____	Landwirt*innen
NPV _____	Negative predictive value; negativer Vorhersagewert
PABAK _____	Prevalence Adjusted Bias Adjusted Kappa
PPV _____	Positive predictive value; positiver Vorhersagewert
Schwellung _____	Schwellungen im Bereich des Sprunggelenkes am rechten und linken Hinterbein (Note 0/1)
Verschm_OH _____	Verschmutzung im oberen hinteren Bereich auf der rechten und linken Körperseite (Note 0/1)
Verschm_UH _____	Verschmutzung im unteren hinteren Bereich auf der rechten und linken Körperseite (Note 0/1)
Wunde _____	Wunden und Krusten im Bereich des Sprunggelenkes am rechten und linken Hinterbein (Note 0/1)

# 1 Einleitung und Problemstellung

Klauen- und Gliedmaßenkrankungen (7,4 %) sind nach Fruchtbarkeitsstörungen (24,1 %) und Eutererkrankungen (13 %) die dritthäufigste Abgangsursache bei Rindern (ZuchtData, 2020). Rouha-Mülleder et al. (2009) ermittelten eine durchschnittliche Lahmheitsprävalenz von 36 % auf österreichischen Betrieben; – diese variierte stark zwischen den Betrieben und reichte von 0 % bis 77 %. Cha et al. (2010) kalkulierten die durchschnittlichen Kosten einer lahmdenden Kuh, welche mit den Behandlungskosten, Milchverlust und der sinkenden Fruchtbarkeit verbunden sind. Eine lahmdende Kuh kostet im Durchschnitt 205 € bei Sohlengeschwür bzw. 125 € bei Digitaler Dermatitis. Dolecheck und Bewley (2018) verglichen mehrere Studien im Zusammenhang mit Lahmheit im Hinblick auf die damit verbundenen Kosten. Die Kosten für einen unspezifische Lahmheit reichten dabei von 62 € bis 450 €. Um die von Lahmheit verursachten Schmerzen der Tiere sowie wirtschaftliche Verluste zu vermeiden, ist die frühzeitige Erkennung von Klauenerkrankungen ein wichtiger Faktor (Lorenzini et al., 2018).

Einzeltierbeobachtungen sind für die einzelbetriebliche Erkennung von gesundheitlichen Störungen und Tierwohlproblemen, wie zum Beispiel Lahmheiten oder auch Hautveränderungen und Tierverschmutzung, und die Ableitung entsprechender Verbesserungsmaßnahmen wichtig (Lorenzini et al., 2018). Kofler (2014) weist darauf hin, dass eine regelmäßige Kontrolle (mind. 1x/Woche im Laufgang bzw. am Weg zum/vom Melkstand) vorgenommen werden sollte, um Lahmheiten frühzeitig zu erkennen und rasch behandeln zu können. Auch die Tierhaltungsverordnung schreibt vor, dass der Zustand der Klauen regelmäßig zu überprüfen und bei Bedarf ist eine Klauenpflege durchzuführen ist (1. ThVO, Anlage 2, 2.7, 2017). Unter Umständen kann die Beurteilung auch am stehenden Rind vorgenommen werden.

Ein österreichischer Bauernhof hielt vor rund 20 Jahren im Durchschnitt 11 Milchkühe. Diese Herdengröße nahm seitdem kontinuierlich zu (Statistik Austria, 2020). Laut Grüner Bericht (2021) wurden 2021 22 Milchkühe pro Betrieb gehalten, es erfolgte eine Aufstockung der Herde um 2 % im Vergleich zum Vorjahr. Aufgrund der steigenden Anzahl der Kühe pro Herde nimmt die Zeit ab, welche dem\*der Landwirt\*in für die Einzeltierbeobachtung und damit auch die Durchführung der Gangbeurteilung zur Verfügung steht (Schlageter-Tello et al., 2014; Lorenzini et al., 2018).

Eine mittel- bis langfristige Verbesserung der Tiergesundheit kann auch über züchterische Maßnahmen erfolgen. Van der Waaij et al. (2005) beschreibt zwei Möglichkeiten, um die Klauengesundheit auf züchterischer Ebene zu verbessern. Zum einen gibt es die Selektion auf Klauengesundheit, welche bislang in den meisten Ländern noch nicht standardmäßig erfasst wird. Andererseits gibt es die indirekte Selektion auf mit Klauengesundheit korrelierenden Zuchtmerkmalen; letztere werden zum Teil bereits routinemäßig erfasst, z. B. Fundamentmerkmale. In allen Fällen ist eine entsprechende Datengrundlage erforderlich, wie z. B. direkte Merkmale aus der Klauenpflege oder tierärztliche Diagnosen. Weiters können Hilfsmerkmale wie Lahmheiten, Merkmale der linearen Nachzuchtbeschreibung sowie Daten aus elektronisch gesteuerten Herdenmanagementwerkzeugen verwendet werden.

Aufgrund der niedrigen Erbllichkeit von Fitness- und Gesundheitsmerkmalen ist eine fundierte Datenbasis wichtig (Egger-Danner, 2015).

Die Daten müssen, insbesondere für die Zuchtwertschätzung, in großen Mengen erhoben werden und zuverlässig erfassbar sein. Der Fokus dieser Masterarbeit liegt daher auf der Reliabilität – also der Zuverlässigkeit von Messungen. Ziel ist es, die Zuverlässigkeit der Erhebung von Lahmheit und weiteren tierbezogenen Merkmalen des Tierwohls von Milchkühen in Laufstallhaltung durch Landwirt\*innen und Kontrollassistent\*innen zu untersuchen.

## 2 Ziele und Forschungsfragen

Es ist Ziel dieser Arbeit, die Zuverlässigkeit der Erhebung von Lahmheit und weiteren tierbezogenen Merkmalen des Tierwohls von Milchkühen in Laufstallhaltung durch Landwirt\*innen und Kontrollassistent\*innen zu untersuchen. Des Weiteren soll ermittelt werden, ob Zeichen für Lahmheit am Stand als Hilfsmerkmal für bei der Gangbeurteilung als hochgradig lahm beurteilte Tiere herangezogen werden können. Die Ergebnisse können sowohl bei der Entwicklung einer Zuchtwertschätzung für solche Merkmale als auch beim einzelbetrieblichen Herdenmanagement berücksichtigt werden.

Daraus ergeben sich folgende Fragestellungen:

- **Wie zuverlässig können Landwirt\*innen bzw. Kontrollassistent\*innen Lahmheit und tierbezogene Indikatoren des Wohlergehens wie Verschmutzung, Schäden am Sprunggelenk und BCS in der Milchviehhaltung erfassen?**
- **Wie gut können Landwirt\*innen bzw. Kontrollassistent\*innen mittels Beurteilung im Stand Tiere identifizieren, die auf Basis Gangbeurteilung als hochgradig lahm eingestuft wurden?**

## 3 Literatur

### 3.1 Berücksichtigung von Lahmheit und anderen Merkmalen des Tierwohls in der Zuchtwertschätzung

In der Rinderzucht beschränkte sich über Jahrzehnte die Selektion auf Produktionsmerkmale, wie zum Beispiel Milch und Fleisch. So wurde zum Beispiel im Jahr 1994 bei Fleckvieh eine Doppelnutzung mit einer Gewichtung von 55 % Milch, 35 % Fleisch und 10 % sekundären Merkmalen (heute als funktionale Merkmale bezeichnet) angestrebt (Miesenberger et al. (1996), zit. n. AGÖF, 1994). Aufgrund der züchterischen Erfolge und konsequenten Verbesserung im Bereich Fütterung und Haltung stieg in den letzten Jahren die Einzeltierleistung, bezogen auf die produzierten Fett- und Eiweißmengen. Weniger im Vordergrund standen kostensenkende Merkmale, wie z. B. die Resistenz gegen Eutererkrankungen. Die einseitige Zucht auf Leistung hat erhebliche Verschlechterungen der Tiergesundheit mit sich gezogen. Ein Beispiel dafür ist eine negative Energiebilanz, welche bei steigender Milchleistung in der Früh lactation die Ursache ist (Tetens, 2019) und einen entscheidenden Einfluss auf die Stoffwechselbelastung der Tiere hat. Bei zu extremer Erschöpfung der Körperreserven, z. B. bei Holstein-Kühen, können Stoffwechsel-, Fruchtbarkeitsstörungen und Lahmheiten entstehen (Brade, 2013). Die Sättigung der Märkte, mit den einhergehenden Preisrückgängen für Milch und Fleisch, aber auch die gesellschaftliche Kritik an den Produktionssystemen der landwirtschaftlichen Tierhaltung (Bennewitz et al. 2019), ließ die Bedeutung von funktionalen Merkmalen ansteigen. Dies ging mit einer Veränderung des Gesamtzuchtwertes unter stärkerer Berücksichtigung von Merkmalen des Tierwohls und der Tiergesundheit einher. Derzeit machen beim Zuchtziel für Fleckvieh die funktionalen Merkmale, also jene Merkmale, die vorliegen müssen, damit das Tier die eigentliche Leistung erbringen kann, 44 % (Fitness) aus, während Milch- und Fleischleistung nur noch mit 38 % bzw. 18 % in den Gesamtzuchtwert eingehen (ZuchtData, 2021).

Funktionale Merkmale sind in Merkmalskomplexen wie Gesundheit, Fruchtbarkeit, Verhalten und Ressourceneffizienz zu finden. Die züchterische Bearbeitung ist im Vergleich zu anderen Merkmalen erschwert, da sie eine niedrigere Erblichkeit als klassische Leistungsmerkmale haben und ihre phänotypische Ausprägung schwerer zu erfassen ist. In Bezug auf Tierwohl, Wirtschaftlichkeit und gesellschaftliche Akzeptanz der Nutztierproduktion haben funktionale Merkmale eine hohe Bedeutung. Mitinbegriffen sind hier auch Merkmale, welche aus der veränderten Erwartungshaltung der Verbraucher entstehen. Zur Identifizierung solcher Merkmalskomplexe sind einfach erfassbare Hilfsmerkmale von Bedeutung, welche eine einfache züchterische Bearbeitung ermöglichen (Tetens, 2019).

In Österreich erfolgt die Zuchtwertschätzung für folgende Fitness-Merkmale (ZuchtData 2019):

- Exterieur
- Nutzungsdauer
- Fruchtbarkeit
- Kälberverlauf und Totgeburtenrate
- Aufzuchtverluste

- Zellzahl und Melkbarkeit

Seit 2010 fließen zusätzlich Gesundheitsmerkmale (Eutergesundheitswert 10 %, Fruchtbarkeitswert 14 %) ein (ZuchtData, 2021). Die Datengrundlage für die Zuchtwertschätzung basiert auf tierärztlichen Diagnosen und geburtsnahen Beobachtungen und wird entweder vom\*der Tierarzt\*ärztin übermittelt oder im Rahmen der Leistungskontrolle erfasst. Dabei wird für die Zuchtwertschätzung überprüft, ob das Rind zum jeweiligen Zeitpunkt gesund war oder vom\*der Tierarzt\*ärztin behandelt wurde. Es werden von validierten Betrieben Daten für die Zuchtwertschätzung verwendet. Bei den geburtsnahen Beobachtungen handelt es sich um Nachgeburtverhalten, Festliegen, Mastitis und Lahmheit.

Die folgenden Merkmale gehen in die ZWS mit ein:

- Mastitis
- Frühe Fruchtbarkeitsstörungen
- Zysten
- Milchfieber (ZuchtData 2019)

Merkmalskomplexe aus dem Klauen- und Fundamentbereich werden bisher in der gemeinsamen Zuchtwertschätzung in Österreich und Deutschland ausschließlich durch die Verwendung der Ergebnisse aus der Exterieurbeurteilung im Rahmen der Nachzuchtbeschreibung berücksichtigt. Die Leistungsbeschreibung von Exterieurmerkmalen erfolgt durch die Beschreibung/Bewertung von etwa 30 – 60 zufällig ausgewählten Töchtern eines Stiers, welcher im Prüf- bzw. Ersteinsatz eingesetzt wird. Es erfolgt eine Vergabe von 4 – 6 Hauptnoten, weiters werden 20 Einzelmerkmale linear auf einer Skala von 1 bis 9 beschrieben. Beim Fundament werden Gesamtnoten vergeben. Die Einzelmerkmale Sprunggelenk (Ausprägung, Winkelung), Fessel und Trachten werden linear beschrieben (Fürst-Waltl et al., 2015). Durch die Auswahl von klauengesunden und langlebigen Vererbern kann die Widerstandsfähigkeit einer Kuhherde gegenüber klauenschädlichen Umwelteinflüssen verbessert werden (Kümper, 2008).

Eine Umfrage im Rahmen des Projektes OptiGene (Steininger et al., 2012) zum Thema „Welche Merkmale, für die es derzeit keine Zuchtwerte gibt, wären für Sie besonders interessant, um sie züchterisch verbessern zu können?“ ergab, dass Züchter\*innen an der Verbesserung des Klauen- und Fundamentkomplexes interessiert sind und dies bei der Auswahl des Besamungstieres berücksichtigen würden. Die Exterieurzuchtwertschätzung allein wird nicht als ausreichend erachtet, um Klauengesundheit züchterisch zu verbessern (König und Swalve, 2006). Fürst-Waltl et al. (2015) schlugen daher vor, Daten aus Gesundheitsmonitoring, Nachzuchtbeschreibung und Klauenpflege zu kombinieren, u.a. indem die Klauenpflegeprotokolle elektronisch erfasst und des Weiteren Lahmheit bei jeder Milchleistungskontrolle erfasst werden. Damit Rinder-Klauendaten langfristig und länderübergreifend vergleichbar sind, benötigt es vereinheitlichte Definitionen und Vorgaben für eine standardisierte Dokumentation, welche bereits vom International Committee for Animal Recording (ICAR) bereitgestellt wird (Christen et al., 2015). Derzeit gibt es in Österreich außer den tierärztlichen Diagnosen und Hilfsmerkmalen aus der linearen Beschreibung keine österreichweite systematische Erfassung von Klauenpflagedaten (Egger-Danner, 2015). Beim Projekt Klauen-Q-Wohl, das 2017 begonnen hat, wurde in Zusammenarbeit mit Landwirt\*innen, Klauenpfleger\*innen und Tierarzt\*innen versucht, eine standardisierte Dokumentation und elektronische zentrale Erfassung sowie Auswertung von Klauenpflagedaten weiterzuentwickeln,

um die Klauengesundheit züchterisch zu verbessern. Mittels elektronischer Hilfsmittel (wie z. B. *App Klauenprofi*) können mögliche Probleme in der Klauengesundheit frühzeitig erkannt und behandelt werden. Weiters kann die mobile Datenerfassung wesentliche Vereinfachung für den\*die Landwirt\*in und Klauenpfleger\*in bringen (Rinderzucht Austria, 2017).

## **3.2 Tierbezogene Indikatoren**

### **3.2.1 Auswahl der Merkmale**

Keeling (s.a.) beschreibt mehrere Typen von Parametern, welche bei der Tierwohlbeurteilung verwendet werden. Sie werden eingeteilt in

- ressourcenbezogene Parameter (z. B. Boxengröße, Art des Bodens (Keeling, s.a.), Wasser-, Futtermittellieferung, Stallausstattung (Welfare Quality®, 2009), usw.
- managementbasierte Parameter (z. B. Klauenpflegemaßnahmen (Keeling, s.a.), Besatzdichte, Enthornung, Einstreumaterial, Zugang zu Weide und Auslauf (Welfare Quality®, 2009), usw. und
- tierbasierte Parameter (z. B. Lahmheit, Verletzungen, Verhalten).

Die ersten beiden Indikatorengruppen werden generell als „Input-Merkmale“ bezeichnet und die letztgenannte als „Output-Merkmale“. Für die Bewertung von Tierwohl sollten so viel tierbasierte Merkmale wie möglich verwendet werden. Dabei werden ressourcenbezogene Merkmale ebenfalls als wichtig angesehen, damit Risiken für das Wohlbefinden ermittelt werden können. Durch die Verwendung von tierbasierten Merkmalen kann das Wohlbefinden von Tieren in verschiedenen Haltungssystemen verglichen werden (Forkman, 2009; Keeling, s.a.).

Grundsätzlich sind Merkmale erforderlich, die 1) zuverlässig und valide sind, 2) von geschultem Personal leicht erhoben werden können, 3) einen begrenzten Zeitaufwand fordern. Bei der Entscheidung für Merkmale werden als zentrale Kriterien daher die Validität, Reliabilität und Durchführbarkeit genannt (Knierim und Winckler, 2009).

### **3.2.2 Validität**

Ein zentraler Aspekt bei der Beurteilung von Tierwohl ist, inwieweit die einzelnen Indikatoren valide sind, d.h. das messen, was gemessen werden soll. Scott et al. (2001) nennen verschiedene Formen der Validität, um die Gültigkeit des Merkmals zu messen. *Face validity* (Plausibilität/Anschauungvalidität) ist eine subjektive Beurteilung, wobei das Merkmal/Skala von einem\*r oder mehreren Expert\*innen als gültig angesehen wird.

Bei *construct validity* (Konstruktvalidität) werden zunächst Hypothesen über die Beziehung von Tierwohl und anderen Variablen formuliert – diese Hypothesen können dann untersucht werden (Cohen et al., 1996; Streiner und Norman, 1995). Wenn die erwartete Beziehung aufrecht erhalten bleibt, dann hat die Messskala Konstruktvalidität (Scott et al., 2001). Eine Skala/Maßstab hat *criterion validity* (Kriteriumsvalidität), wenn sie eine nachweisbare Beziehung entweder zu einer bestehenden Messgröße oder zu einem ‚Goldstandard‘-Merkmal hat (Cohen et al., 1996; Streiner und Norman, 1995). Im Bereich des Tierwohls existiert jedoch kein ‚Goldstandard‘, dadurch kann die Kriteriumsvalidität nur durch den Vergleich mit

anderen gültigen (validen) und zuverlässigen (reliablen) Methoden nachgewiesen werden (Cohen et al., 1996).

Waiblinger et al. (2001) nennen Gesundheit und Verhalten als direkte Tierwohlintikatoren, welche eine hohe Wahrscheinlichkeit für Validität aufweisen. Der Nachteil dieser Tierwohlintikatoren ist, dass bei der Erhebung Expert\*innen vorausgesetzt werden und es keine Angaben zur Ursache für die Beeinträchtigung des Tierwohl gibt. Beim Tierwohlintikator Gesundheit werden folgende Parameter vorgeschlagen, welche bei Erhebungen eingebunden werden sollen: Hautverletzungen, Störung des Bewegungsapparates, Mastitis, Körperkondition, Merzungsrate aufgrund von Krankheiten, Sauberkeit und Gehalt an somatischen Zellen in der Milch. Das Zeitbudget betreffend Liege-, Steh- und Fressverhalten, Verhaltensrhythmen, Aufstehverhalten, soziale Interaktionen und Ausweichdistanz gegenüber dem Menschen sind verhaltensbezogene Tierwohlintikatoren.

### **3.2.3 Zuverlässigkeit der Erfassung**

Geeignete Merkmale sind nicht nur valide (Gültigkeit), sondern zeichnen sich auch durch ein hohes Maß an Wiederholbarkeit aus (Knierim und Winckler, 2009). Im Vordergrund steht hierbei die Wiederholbarkeit innerhalb und zwischen Beobachter\*innen sowie die kurzfristige Wiederholbarkeit (z. B. innerhalb desselben Tages). Die Intra-Beobachter-Reliabilität beschreibt die Übereinstimmung bei aufeinanderfolgenden Messungen des\*der gleichen Beobachters\*in (Scott et al. 2001). Es sollen aber auch verschiedene Beobachter\*innen mit einem bestimmten Grad an Ausbildung bzw. Training möglichst das gleiche Ergebnis erzielen (Inter-Beobachter-Reliabilität = die Übereinstimmung der Ergebnisse verschiedener Beobachter\*innen). Weiters sollen Beobachter\*innen bei einer Wiederholung des Tests/der Erhebung zu gleichen Ergebnissen kommen (Test-Retest-Reliability) (Knierim und Winckler, 2009). Die Frage der langfristigen Wiederholbarkeit ist umstrittener. Bei der Beurteilung von Tierwohl im Rahmen von Zertifizierungen auf landwirtschaftlichen Betrieben hat die Wiederholbarkeit (Repeatability) über die Zeit eine hohe Priorität. Um Kosten einzusparen, finden solche Beurteilungen in längeren Abständen (z. B. Monate bis 1x jährlich) statt. Das bedeutet, dass die Beurteilungsergebnisse für eine längerfristige Betriebssituation repräsentativ sein müssen. Wenn demnach keine größeren Veränderungen, welche das Tierwohl beeinflussen, am Betrieb durchgeführt werden, soll ein ähnliches Ergebnis erzielt werden können (Knierim und Winckler, 2009). Einige Messwerte verändern sich allerdings im Laufe der Zeit und mit zunehmendem Alter der Tiere; zum Beispiel Lahmheitswerte bei Masthühnern. In diesem Fall kann durch die Beurteilung mehrerer aufeinanderfolgender Durchgänge im gleichen Alter erfasst werden, ob ein Problem konsistent auftritt (Keeling, s.a.).

#### **3.2.3.1 Analyse der Inter- und Intra-Beobachter-Reliabilität**

In Studien, an denen mehrere Beobachter\*innen beteiligt sind, sollte die Zuverlässigkeit der Beobachter\*innen überprüft werden, damit kein ‚Observer bias‘ vorliegt und es zu keiner Verzerrung der Daten kommt. Diese Zuverlässigkeit kann von Übung, Erfahrung, Training, Schnelligkeit des Verhaltens, Aufmerksamkeitslevel

des\*der Beobachters\*in sowie auch die Klarheit der Definition von Merkmalen beeinflusst werden. Des Weiteren läuft ein\*e einzelne\*r Beobachter\*in oft Gefahr ein Merkmal konsequent falsch zu erheben (Martin und Bateson, 1993). Deshalb ist es notwendig, einen Vergleich zwischen den Beobachter\*innen durchzuführen. Auch im Falle der Durchführung eines Experimentes mit nur einem\*r Beobachter\*in ist es wichtig, die Zuverlässigkeit zu bewerten, um mögliche Verzerrungen aufzuweisen (Kaufman und Rosenthal, 2009).

Es sollten eine vielfache Auswahl an Verhaltensweisen verwendet werden, um eine genaue Bewertung der Zuverlässigkeit über alle Situationen/Gegebenheiten zu gewährleisten (Martin und Bateson, 1993). Rosenthal (1976) beschreibt die Wichtigkeit der Neutralität und der Unvoreingenommenheit. Es wird aber zugegeben, dass es durchaus möglich ist, dass ein\*e Beobachter\*in (also ein Mensch) unbewusst sieht, was er gerne sehen möchte. Des Weiteren ist es möglich, dass in Feldsituationen in denen es häufig zu Sichtbehinderungen oder sich schnell ändernden Verhaltensweisen kommt, sich die Beobachter\*innen irren.

Für die Beschreibung der Beobachter\*innenübereinstimmung existieren, auch in Abhängigkeit vom Typ der untersuchten Messgröße (kontinuierlich vs. kategoriell), verschiedene Kenngrößen, u.a.:

- Übereinstimmung in Prozent (kategoriell)
- Kappa-Koeffizienten (kategoriell)
- Korrelation (kontinuierlich)
- Kendall's W (kontinuierlich, mehrere Beobachter\*innen)

Für die Übereinstimmung von verschiedenen Beobachter\*innen in Prozent wird die Summe der Übereinstimmungen durch die Gesamtzahl der Beobachtungen dividiert (Pritchard et al., 2007; Channon et al., 2009).

Als weiteres Maß für die Beobachter\*innenübereinstimmung wird häufig der Kappa-Koeffizient verwendet. Dieser Wert berücksichtigt die beobachtete Übereinstimmung ( $p_o$ ) sowie die zufällig erwartete Übereinstimmung ( $p_e$ ) (Cohen, 1960). Eine weitere Möglichkeit die Beobachter\*innenübereinstimmung auszudrücken, ist der PABAK (prevalence adjusted bias adjusted kappa). Der PABAK berücksichtigt die Häufigkeit/Vorkommen eines Merkmals sowie die Voreingenommenheit (Bias) des\*der Beobachters\*in. Kappa sowie PABAK können Werte zwischen -1 und 1 annehmen (Byrt et al., 1993). Bei einem Kappa/PABAK von 0,00 entspricht die Übereinstimmung der Zufallswahrscheinlichkeit. Bei einem Kappa/PABAK von 1 bedeutet dies eine 100 %-ige Übereinstimmung (Fleiss et al., 2003). Der Kappa (einschließlich PABAK) kann in folgende Kategorien unterteilt werden: 0,61 – 1 Übereinstimmung gut bis sehr gut, 0,41 - 0,60 Übereinstimmung noch akzeptabel, <0,40 Übereinstimmung unbefriedigend (Landis und Koch, 1977). Ein Kappa Koeffizient von 0,4 bedeutet im extremen Fall, dass zwei Beobachter\*innen bei einer Übereinstimmung von 70 % eine Auftretenshäufigkeit von 50 % aber auch 80% Lahmheit ermitteln können (Knierim und Winckler, 2009). Bei Tierwohlbeurteilungen mit dem Ziel, das Tierwohllevel auf landwirtschaftlichen Betrieben zu klassifizieren, was in weiterer Folge ökonomische Konsequenzen für den\*die Landwirt\*in haben könnte, soll eine höhere Reliabilität angestrebt werden (de Passillé and Rushen, 2005).

Die Zuverlässigkeit (Reliabilität) innerhalb und zwischen den Beobachter\*innen kann für kontinuierliche Messgrößen mittels Korrelationen zwischen den Ergebnissen für

Beobachter\*innenpaare gemessen werden. Dabei ist es wichtig, mehrere zufällig ausgewählte Beobachtungen in der Stichprobe zu verwenden. Dadurch wird eine genaue Bewertung der Zuverlässigkeit unter allen Bedingungen gewährleistet (Martin und Bateson, 2007). Bei der Interpretation des Korrelationskoeffizienten ( $r$ ) gibt es unterschiedliche Grenzwerte. Bei einem Koeffizient  $r < 0,1$  liegt ein vernachlässigbarer Zusammenhang vor und bei einem Koeffizient  $r > 0,9$  liegt ein starker Zusammenhang vor (Schober et al., 2018). Martin und Bateson (2007) schlagen für wichtige Merkmale als Grenzwert für eine akzeptable Übereinstimmung eine Korrelation von mind. 0,7 vor.

Neben dem Kappa Koeffizienten ( $k$ ), welcher für diskrete Daten verwendet wird, gibt es den Kendall's coefficient of concordance (Kendall's  $W$ ). Kendall's  $W$  wird bei Rangfolgen mit kontinuierlichen Daten angewandt. Sowohl der Kappa, als auch Kendall's  $W$  sind zufallsbereinigte Maße für die Übereinstimmung zwischen mehr als zwei Beobachter\*innen, welche dasselbe beobachten (Cohen, 1960; Siegel and Castellan, 1988).

Winckler und Willen (2001) untersuchten die Übereinstimmung zwischen drei Beobachter\*innen bei gleichzeitiger Erhebung des Gangs bei Milchkühen ( $n=147$ ) mittels 5-stufigen Lahmheit-Scores. Dabei kam es im Mittel zu einer Beobachter\*innenübereinstimmung von 68 % (63 – 74 %). Bei 30 % (25 – 34 %) der Kühe unterschied sich die Lahmheitsbewertung um einen Score. Fast zwei Drittel (62 %) der Abweichungen fanden innerhalb der Lahmheit-Scores 1 (normaler Gang) und 2 (unebener Gang) statt. Die Schwierigkeit der Unterscheidung zwischen normalem Gang und möglicher subklinischer Lahmheit (unebener Gang) bei Kühen zeigt die Grenzen bei der subjektiven Lahmheitsbewertung auf (Whay et al., 1998).

Hollenbeck (1978) merkt an, dass Beobachter\*innen auf verschiedene Arten die Wiederholbarkeit (Repeatability) beeinflussen können. Ergebnisse können einerseits aufgrund von Beobachterverzerrung (Bias) beeinflusst werden, z. B. Fehler durch Nichtberücksichtigung von Kriterien (das aufgetretene Verhalten wurde nicht erhoben) oder durch eine bestimmte Erwartungshaltung (das Tier wurde nicht als lahm eingestuft, da die restliche Herde gesund ist). Andererseits kann die Erfahrung des\*der Beobachters\*in das Ergebnis beeinflussen. In der Studie von Main et al. (2000) kam es bei der Beurteilung von Lahmheit bei Schweinen zu einer Übereinstimmung zwischen den Beobachtungen bei unerfahrenen Beobachter\*innen von 26 % - 53 %, während diese bei erfahrenen Beobachter\*innen bei 94 % lag.

### 3.2.3.2 Verbesserung der Inter- und Intra-Beobachter-Reliabilität

In der Literatur werden verschiedene Ansätze für die Verbesserung der Übereinstimmung zwischen und innerhalb Beobachter\*innen beschrieben.

Eine Verfeinerung von Definitionen bzw. verbesserte Beschreibung von Kategorien kann zu einer Verbesserung der Reliabilität führen, wie es z. B. in der Studie Thomsen et al. (2008) versucht wurde. Dazu wurden zwei Experimente mit Expert\*innen und Direktbeobachtungen bei Milchkühen auf Spaltenböden durchgeführt, welche bereits ausreichend Erfahrung in der Lahmheitsbeurteilung hatten. Es wurde ein Lahmheitsscoring-System mit fünf Stufen verwendet. Das Scoring-System wurde im Hinblick auf Inter- und Intra-Beobachter\*innenübereinstimmung unter der Verwendung der Kappa-Statistik evaluiert. Die Beobachter\*innenübereinstimmung wurde vor und nach einem Training erhoben. Dabei kam heraus, dass das Training nur einen geringen positiven Effekt auf die (Inter- und Intra-) Beobachter\*innenübereinstimmung hatte. Die Weighted Kappa-Werte (Inter-Beobachter-Reliabilität) von allen Beobachter\*innen reichten von 0,24 bis 0,68 und lagen im Mittel bei 0,48 und 0,52 vor bzw. nach dem Training. Die Autor\*innen weisen darauf hin, dass es wichtig ist, klinische Zeichen zu verwenden, welche in wenigen Sekunden von Beobachter\*innen erhoben werden können (z. B. können sich die Gelenke frei bewegen oder nicht). Bestimmte klinische Anzeichen sind in den meisten Fällen bei den Rindern vorhanden und können somit dem Lahmheitscore zugeordnet werden, aber manchmal ist das nicht der Fall. Die meisten lahmen oder stark lahmen Kühe (Lahmheits-Scores 4 und 5) haben beim Stehen und Gehen einen gewölbten Rücken. Es wurden jedoch auch oft Kühe beobachtet, die offensichtlich lahmten (keine oder nur eingeschränkte Gewichtsbelastung auf einem Bein), die aber keine Anzeichen eines gewölbten Rückens beim Stehen (und manchmal beim Gehen) zeigten. Mit dem Zusatz „in den meisten Fällen“ soll eine falsche Klassifizierung vermieden werden (hat z. B. eine offensichtlich lahme Kuh keinen gewölbten Rücken im Stehen würde sie trotzdem als lahm eingestuft werden können) und macht die Beschreibung weniger starr und verständlicher.

Während bei Thomsen et al. (2008) keine Verbesserung der Beobachter\*innenübereinstimmung gefunden wurde, konnte bei anderen Arbeiten eine Erhöhung der Übereinstimmung mittels Zusammenlegens von Lahmheits-Scores festgestellt werden. In den Arbeiten von Brenninkmeyer et al. (2007) und March et al. (2007) wird eine effiziente Methode beschrieben, um den Anteil der Beobachter\*innen, welche die gleichen Scores erheben, zu erhöhen. Sie legten eine 5-Punkte-Lahmheitsskala zu einer 2-Punkte-Lahmheitsskala (lahm/nicht lahm) zusammen. Dadurch erhöhte/verbesserte sich der Prevalence Adjusted Bias Adjusted Kappa (PABAK) um bis zu 0,3.

Schlageter-Tello et al. (2014) untersuchten die Übereinstimmung von 10 Expert\*innen (Intra- und Inter-Beobachter-Reliabilität) mittels Videobeurteilung von Lahmheit bei Rindern anhand eines 5-Punkte-Lahmheits-Scores. Die Grenzwerte für eine zuverlässige Beobachter\*innenübereinstimmung (Intra- und Inter-rater) wurde mit einem *Weighted Kappa-Wert* von  $\geq 0,6$  ( $k_w$  und  $k$ ) und einer Prozent-Übereinstimmung  $\geq 75$  % festgelegt. Die Übereinstimmung innerhalb der

Beobachter\*innen (Inter-rater) reichte von 0,28 – 0,84 *kw* bzw. von 22,6 – 81,8 % Übereinstimmung beim 5-Punkte-Lahmheits-Score. Nach dem Zusammenlegen der einzelnen Scores in einen 4-, 3-, und 2-Stufen-Lahmheits-Score wurde der Weighted Kappa bzw. Kappa sowie die Prozent-Übereinstimmung neu berechnet. Verglichen mit anderen Studien (March et al., 2008 und Channon et al., 2009) hatte das Zusammenfassen der Scores keinen signifikanten Effekt auf die Zuverlässigkeit der Beobachter\*innenübereinstimmung. Jedoch stieg die Gesamtübereinstimmung (Inter-Beobachter-Reliabilität) der Beobachter\*innen um etwa 10 %-Punkte nach der Zusammenlegung jeweils eines Lahmheits-Scores (siehe Tabelle 1) (Schlageter-Tello et al., 2014).

Tab. 1: Beobachter\*innenübereinstimmung (Inter-Beobachter-Reliabilität) für das 5-Punkte-Lahmheits-Score und nach dem Zusammenlegen der einzelnen Scores in ein 4-, 3-, und 2-Punkte-Lahmheits-Score; ausgedrückt in Weighted KAPPA (*kw*), KAPPA (*k*) sowie Prozent-Übereinstimmung (PA) (Schlageter-Tello et al., 2014)

Lahmheits-Score	Kombinationen (vom original 5-Punkte-Score)	Beobachter*innenübereinstimmung (Inter-Beobachter-Reliabilität)	
		<i>kw</i> (Weighted KAPPA) oder <i>k</i> (KAPPA für 2-Punkte-Score)	PA (Prozent-Übereinstimmung)
5-Punkte	--	0.65 (0.64–0.66)	57.1 (55.7–58.4)
4-Punkte	(12)345	0.67 (0.66–0.69)	69.7 (68.3–70.9)
	1(23)45	0.61 (0.60–0.63)	70.6 (69.3–71.8)
	12(34)5	0.63 (0.61–0.64)	67.4 (66.1–68.7)
	123(45)	0.67 (0.65–0.68)	62.1 (60.7–63.4)
3-Punkte	(12)3(45)	0.70 (0.69–0.72)	74.6 (73.4–75.8)
	1(23)(45)	0.64 (0.62–0.66)	75.5 (74.4–76.7)
	(12)(34)5	0.66 (0.64–0.68)	80.0 (79.0–81.1)
	12(345)	0.65 (0.64–0.67)	72.6 (71.4–73.8)
	1(234)5	0.53 (0.51–0.56)	81.7 (80.7–82.8)
	(123)45	0.64 (0.62–0.66)	83.6 (82.6–84.6)
2-Punkte	1(2345)	0.57 (0.54–0.60)	86.9 (86.0–87.8)
	(12)(345)	0.70 (0.68–0.72)	85.2 (84.2–86.2)
	(123)(45)	0.70 (0.67–0.72)	88.6 (87.7–89.5)
	(1234)5	0.42 (0.37–0.48)	94.7 (94.2–95.5)

Neben der Zusammenfassung von Kategorien wie bei Schlageter-Tello et al. (2014) kann sich auch die Verwendung einer kontinuierlichen Skala positiv auf die Übereinstimmung zwischen Beobachter\*innen auswirken. Engel et al. (2003) untersuchten die Ergebnisse von Beobachter\*innen, welche den Gang von Kühen anhand von Videoaufnahmen klassifizierten. Die visuelle Beurteilung erfolgte auf zwei Arten: anhand von diskreten/geordneten Scores (1 = normal bis 9 = stark abnormal) und als kontinuierliche Scores durch die Markierung auf einem Papierstreifen (links = 0 = normal und rechts = 1 = stark abnormal). In der Studie von Engel et al. (2003) wurde herausgefunden, dass die Verbesserung der Übereinstimmung durch das Zusammenfassen von Scores mit einer geringeren Unterscheidung von Lahmheits-Scores (z. B. Unterscheidung zwischen Score 1 und 2) verbunden ist. Je nach Ziel der Erhebung kann dies ein Problem darstellen oder nicht. Einen ähnlichen Effekt für eine verbesserte Übereinstimmung bzw. eine geringere Unterscheidungsfähigkeit wurde erzielt, wenn größere Unterschiede zwischen den Beobachter\*innen zugelassen werden. So stieg die Übereinstimmung zwischen den Beobachter\*innen und dem\*der Expert\*in (Goldstandard) unter Anwendung der diskreten Scores von 47 % auf 80 %, wenn eine Abweichung von einem Score toleriert wurde. Sowohl bei den diskreten als auch bei den kontinuierlichen Scores wurde eine Verbesserung der Übereinstimmung zwischen Beobachter\*innen und Expert\*in nach einem weiteren Training festgestellt. Die Ergebnisse der kontinuierlichen Scores wurden als mittlere Abweichung zum\*r Expert\*in dargestellt. Verglichen zu den diskreten Scores gab es eine höhere Variabilität innerhalb der Ergebnisse bei den kontinuierlichen Scores.

Eine ähnliche Studie wurde auch für die Gangbeurteilung bei Sauen durchgeführt. Nalon et al. (2014) verwendeten für die Beurteilung von Videoaufnahmen drei verschiedene Skalen (visuelle Analogskala mit 5 Intervallen, 5- und 2-Punkte-Skala). Die höchste Wiederholbarkeit zeigte die visuelle Analogskala (Inter-Beobachter-Reliabilität 0,73 und Intra-Beobachter-Reliabilität 0,8). Eine geringere Wiederholbarkeit zeigte die 2-Punkte Skala (Inter-Beobachter-Reliabilität 0,6 und Intra-Beobachter-Reliabilität 0,67). Aus der Perspektive des Tierwohls ist zu erwähnen, dass eine Verringerung der Anzahl von Lahmheit-Scores, insbesondere wenn diese auf lahm/nicht lahm reduziert wird, mit dem Verlust wichtiger Informationen über den Lahmheitsstatus der einzelnen Sau und der ganzen Sauenherde verbunden sein kann. Ähnliche Aussagen trafen auch Tuytens et al. (2009). Die Anwendung von einer kontinuierlichen Skala wies eine höhere Sensitivität auf, um Änderungen beim Lahmheitsgrad bei Milchkühen festzustellen.

### **3.2.4 Durchführbarkeit/Praktikabilität**

Knierim und Winckler (2009) nennen auch die Machbarkeit und die Praktikabilität als wichtige Eigenschaften/Aspekte bei Tierwohlerhebungen auf landwirtschaftlichen Betrieben.

Einige Studien haben sich explizit mit der Durchführbarkeit und Praktikabilität von Tierwohlerhebungen beschäftigt. Knierim und Winckler (2009) erwähnen, dass bei Erhebungen die Herdengröße, eine repräsentative Stichprobe sowie die Identifikation der Tiere problematisch sein können. Bei Anwendung des Welfare Quality® Protokolls wird die Gesamtzeit, zur Erhebung von tierwohlbasierenden Merkmalen,

welche im Zusammenhang mit der Herdengröße steht, als Haupthindernis angesehen. Bei 60 Rindern wurden etwa 5,5 Stunden für die Erhebungen und dem Interview mit dem\*der Landwirt\*in durch eine\*n Beurteiler\*in benötigt. Zu einem ähnlichen Ergebnis kamen auch die Autor\*innen Gieseke et al. (2014), welche auch die Praktikabilität und Anwendbarkeit des Welfare Quality® Protokolls untersuchten. Dabei wurde die Gesamtdauer der Betriebsbesuche erfasst und eine Einschätzung des erforderlichen Zeitaufwands konnte erfolgen. Demnach hatte die betriebsindividuelle Bestandsgröße den höchsten Einfluss auf die Dauer der Datenerhebung. Die Erhebungsdauer des Einzeltiers konnte reduziert werden, wenn eine Fixierung, z. B. ein Fressgitter, vorhanden war. Den höchsten Zeitaufwand erforderten Verhaltensbeobachtungen beim Einzeltier, welche nach Möglichkeit vereinfacht werden sollten. Um die Akzeptanz der Landwirt\*innen zu erhöhen sowie Personalkosten im Rahmen von Zertifizierungsprogrammen einzusparen, sollte eine Reduktion des erforderlichen Zeitaufwands erfolgen.

Im Hinblick auf die Anpassung der Erhebung an die Betriebsabläufe muss beachtet werden, dass z. B. Beobachtungen des Sozialverhaltens zu bestimmten Tageszeiten, z. B. orientiert an dem Zeitpunkt der Futtervorlage, durchgeführt werden müssen. Die Erfassung von Merkmalen wie Integumentschäden ist dagegen nicht abhängig von der Tageszeit (Welfare Quality®, 2009). Bei der Lahmheitserhebung gibt es andere zeitabhängige Faktoren, welche das Ergebnis beeinflussen können. Bei der Erhebung von Lahmheit wurde von Flower et al. (2006) festgestellt, dass die Euterfüllung einen Einfluss auf den Gang der Rinder hatte. Sie erfassten den Gang von Milchkühen vor und nach dem Melken. Nach dem Melken gingen Rinder schneller und machten längere Schritte.

Bahr et al. (2008) nennt automatische Aufzeichnungen wie z. B. Beschleunigungssensoren oder Bildanalysetechniken, welche für die Lahmheitserkennung entwickelt wurden, als vielversprechende Ansätze. Flower und Weary (2008) schlagen eine technische Ausrüstung (zur kinetischen/kinematischen Messung) vor, um die Subjektivität der Lahmheitsbeurteilung zu verringern und um die Beobachter\*innenzuverlässigkeit zu erhöhen. Dieses Equipment ist aber teuer und unpraktisch für on-farm Erhebungen. Weiters können nicht alle Sensortechnologien, welche in einer Laborumgebung entwickelt wurden, erfolgreich auf einem Milchviehbetrieb implementiert werden. Sie funktionieren besser in einer experimentellen Umgebung, wo kontrollierte Bedingungen mit einer kleinen Stichprobe vorherrschen (Stygar et al., 2021). Schlageter-Tello et al. (2018) untersuchten die Übereinstimmung zwischen on-farm-Lahmheitsscoring durch 2 Expert\*innen (Goldstandard) und einer computerbasierenden automatischen Lahmheitserkennung (3D-ALS). Es kam bei der Identifikation von Rindern als lahm; zwischen den Expert\*innen (Sensitivität von 74,2 % bei Vergleich Expert\*in 1 mit 2) und 3D-ALS (Sensitivität 73,9 bzw. 71,8 % bei Vergleich von Expert\*innen 1 bzw. 2 mit 3D-ALS) zu ähnlichen Ergebnissen, wobei eine geringere Prozentübereinstimmung (lahm und nicht lahm) zwischen Expert\*innen (84,2 %) und 3D-ALS (67,7 – 69,2 %) vorlag. Eine Erklärung für die Unterschiede in der Übereinstimmung ist die technische Einschränkung der 3D-ALS (enger Blickwinkel und Probleme bei der Rinderbewegung).

Main et al. (2010) untersuchten in einer Studie zur Lahmheitsbeurteilung bei Rindern, wie groß die Stichprobe sein muss, um ein valides Ergebnis zu erhalten. Abhängig von der Herden-/Betriebsgröße bedarf es zu viel Erhebungszeit, um Lahmheit von der gesamten Herde sowie unterschiedlichen Aspekte des landwirtschaftlichen Betriebes zu erfassen. Wird nur bei einem Teil der Rinder z. B. beim Verlassen des

Melkstands, Lahmheit erhoben, ist zu beachten, dass meistens eher die hochgradig lahmen Tiere gegen Ende des Melkprozesses den Melkstand verlassen. Es wird vorgeschlagen, dass in der Mitte des Melkprozesses Lahmheit erfasst werden sollte. Die Gangbeurteilung nach dem Verlassen des Melkstandes konnte nicht durchgeführt werden, wenn 1) der Melkstandtyp oder sein Ausgang es zu schwierig machten, die Rinder in der freien Bewegung zu beurteilen, 2) die Tieridentifikation nicht möglich war, 3) die Herde zu klein war. Auf diesen Betrieben wurde die Gangbeurteilung in einem Paddock durchgeführt.

Sorensen et al. (2007) beschäftigten sich mit den tatsächlichen Kosten (€) für on-farm Tierwohlerhebungen bei einem Schweinebetrieb, einem AMS-Betrieb, einem Legehennenbetrieb und einer Nerzfarm. Neben den bei der Erhebung anfallenden Kosten entstehen Kosten für die Datenverarbeitung sowie für die Präsentation der Ergebnisse der Tierwohlerhebung für den\*die Landwirt\*in. Um Kosten zu reduzieren, wird empfohlen, die zu erhebenden Parameter in einem Protokoll zu reduzieren. Es wird z. B. vorgeschlagen, anstatt der klinischen Untersuchung von Gewicht und Fußgesundheit bei Hühnern das Gefieder der Herde zu beurteilen. Demnach können 30 % der Zeit und 20 % Kosten eingespart werden. Weiters wird erwähnt, dass die Häufigkeit und die Anzahl der Tiere bei Erhebungen reduziert werden kann, ohne dabei die Validität zu verlieren (Waiblinger und Menke, 2003; Sorensen et al., 2007). Eine weitere Möglichkeit Kosten zu sparen ist die Durchführung der Erhebungen durch den/die Landwirt\*in und nicht mehr durch externe Personen (Sorensen et al., 2007).

### **3.3 Beobachter\*innentraining**

Damit eine akzeptable Zuverlässigkeit und Übereinstimmung bei Beurteilungen von Lahmheit über die Zeit hinweg bestehen, benötigen Beobachter\*innen Schulungen und Training.

In dem Artikel von Van Reenen und Engel (2009) wird die Entwicklung einer standardisierten Methode beschrieben, um Beobachter\*innen zu trainieren. Demnach soll ein ‚Goldstandard‘ für Tierwohl-Monitoring-Systeme definiert werden, welcher fachlich begründet ist und allgemeine Anerkennung findet. Es werden Tierwohlmerkmale verwendet, die valide, wiederholbar und durchführbar sind. Die Beobachter\*innen sollen in zwei aufeinanderfolgenden Phasen trainiert werden. Die erste Phase ist die sogenannte „Laborphase“, in der visuelle Materialien (z. B. Fotos, Videos) verwendet werden. In der zweiten Phase findet die Direktbeobachtung auf Betrieben statt – die „on-farm“-Phase. Während der Laborphase wird vorausgesetzt, dass für jedes Foto oder Video ein ‚wahrer‘ Messwert vorhanden ist, der bevorzugt von einem Forum von Expert\*innen festgelegt werden sollte. Dieser ‚wahre‘ Messwert ist der sogenannte ‚Goldstandard‘, gegen welchen die Beobachter\*innen trainiert werden sollen. Nach einer Einführungsphase sollen die Beobachter\*innen eine bestimmte Anzahl an Fotos und/oder Videos bewerten und ein Mindestmaß an Übereinstimmung mit dem ‚Goldstandard‘ erreichen. Diese Mindest-Anzahl von korrekten Antworten wird durch statistische Überlegungen ermittelt. Fotos und Videos sollten in einer repräsentativen und ausreichenden Anzahl vorhanden sein, weil sonst keine ausreichende Grundlage für die statistische Evaluierung vorliegt oder die Ergebnisse aufgrund des ‚Memory Effekts‘ verzerrt sein können. Die Autoren schlagen vor, dass der\*die Beobachter\*in die Direktbeobachtung im Stall („on-farm“-Phase) nach einem erfolgreichen Abschluss der Laborphase – z. B. nach einer

bestandenene Prüfung - durchführen soll. Während der ‚on-farm‘-Trainingsphase werden individuelle Tiere und/oder Ställe analysiert. Ähnlich wie in der Laborphase ist es während der ‚on-farm‘-Phase notwendig, eine ausreichende Anzahl an Tieren und/oder Ställen (abhängig von dem Merkmal, welches beobachtet wird) zur Verfügung zu haben, damit ein ausreichend genaues Ergebnis erhalten wird. Zwischen den Trainer\*innen (jene, die das Training beim ‚on farm‘-Training anleiten und betreuen) sollte mindestens ein\*e Experte\*in sein, welche\*r in der statistischen Beurteilung bei den ‚on-farm‘ Ergebnissen beteiligt ist. Diese\*r Experte\*in kann als Silberstandard betrachtet werden. Dieser Silberstandard zeigt eine ähnliche Beurteilung wie der Goldstandard. Nachdem die Ergebnisse vom\*von der Experten\*in beim ‚on-farm‘-Training als weniger wiederholbar betrachtet werden können (schnelle Entscheidungen müssen getroffen werden, unter nicht einfachen Bedingungen etc.), verglichen mit der Laborphase mit Fotos und Videos (ohne Zeitdruck, ideale Bedingungen, es gibt die Möglichkeit für eine Diskussion mit anderen Expert\*innen), wird im Zusammenhang mit dem ‚on-farm‘-Training der Begriff Silberstandard verwendet. Nach Beendigung des ‚on-farm‘-Trainings sollte der\*die erfolgreiche Teilnehmer\*in eine ausreichende Anzahl an Tieren oder Ställen bewertet haben, und eine ausreichende Übereinstimmung mit jener Übereinstimmung des Expert\*innenergebnisses erzielt haben. Die Autor\*innen betonen, dass das Training in engem zeitlichen Zusammenhang mit der Datenerhebung stattfinden soll. Auch im Laufe der Datenerhebung sollte überprüft werden, ob Abweichungen vom Standard vorliegen.

Während Van Reenen und Engel (2009) sich mit der Beobachter\*innenübereinstimmung auf Einzeltierebene beschäftigen, beschreibt Mullan et al. (2011) die Übereinstimmung auf Herdenebene (Prävalenzen; siehe Abbildung 1). Es sollen dabei Grenzen festgelegt werden, wieviel von der erfassten Beobachtung von der Prävalenz (erfasst durch einen Trainer) abgewichen werden darf, um eine ausreichende Übereinstimmung zu erreichen. Bei weniger als 2 % ( $\leq 20$  % Auftretenshäufigkeit innerhalb der Herde) bzw. 5 % ( $\geq 20$  % Auftretenshäufigkeit) Abweichung der erfassten Prävalenzen gegenüber dem vom Trainer ermittelten Wert, wurde eine hohe prozentuelle Übereinstimmung angenommen. Erreichte der\*die Beobachter\*in keine hohe Übereinstimmung (das Ergebnis lag oberhalb der vorgeschlagenen Grenze), sollte ein weiteres Training durchgeführt werden. Die Autor\*innen erwähnten, dass aufgrund der wiederholten Trainings der Beobachter\*innen die Gefahr der Beobachterverzerrung (Observer bias) auftreten kann.

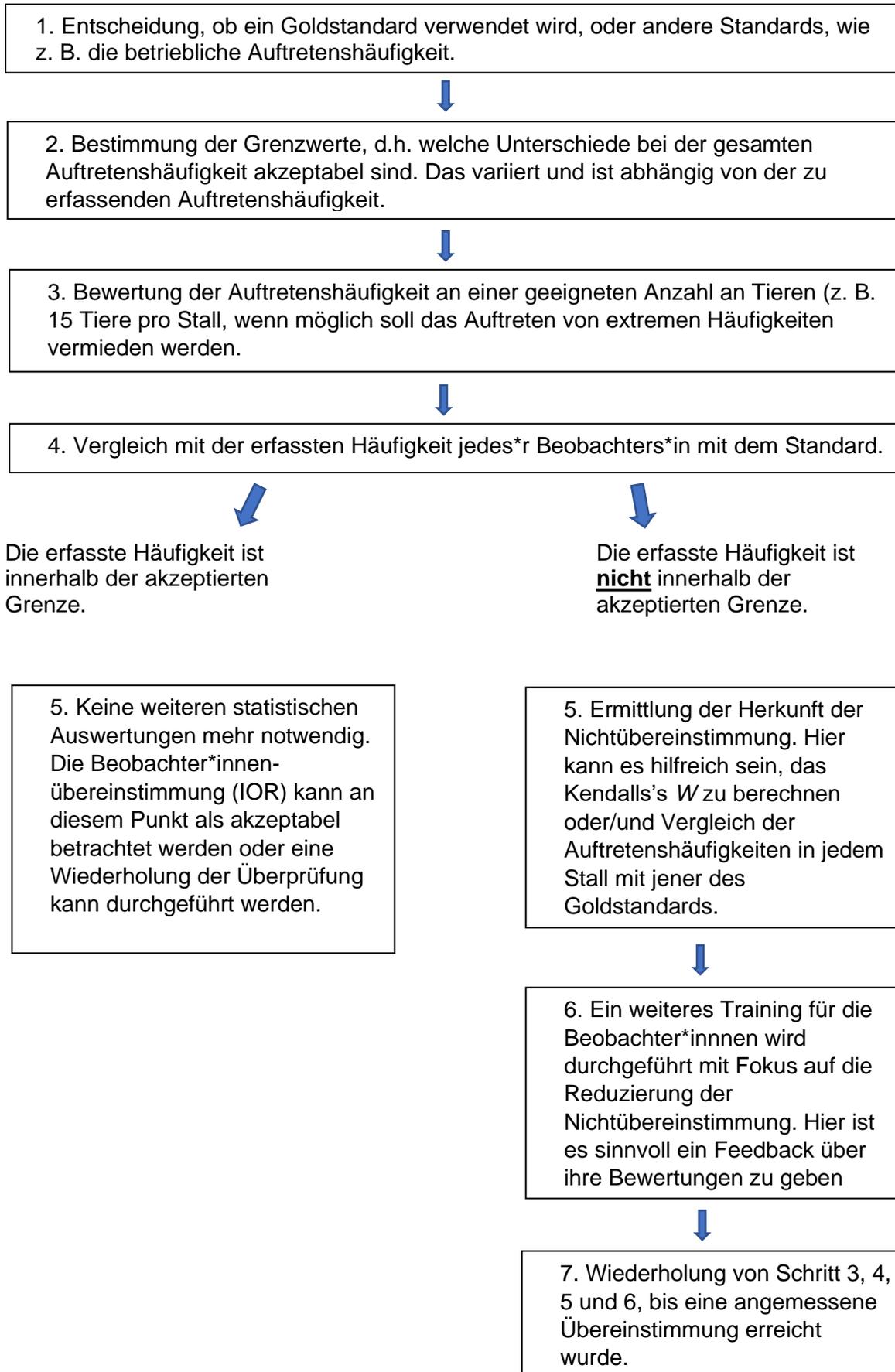


Abb. 1: Vorgeschlagenes Protokoll nach Mullan et al. (2011) für das Testen der Beobachter\*innen-übereinstimmung (IOR) von Beurteiler\*innen der FA (farm assurance)

## **4 Material, Methode und Tiere**

### **4.1 Hintergrund Projekt-Klauen-Q-Wohl**

Das Projekt „Klauen-Q-Wohl“ fand im Rahmen der Europäischen Innovationspartnerschaft „Landwirtschaftliche Produktivität und Nachhaltigkeit“ (EIP-AGRI) statt. Es wurde von der ZuchtData, einer Tochterfirma der zentralen Arbeitsgemeinschaft Österreichischer Rinderzüchter (ZAR), koordiniert und die Laufzeit war Oktober 2017 bis Oktober 2020 (Rinderzucht Austria, 2017). „Klauen-Q-Wohl“ hatte zum Ziel, eine österreichweite Infrastruktur zur zentralen standardisierten Erfassung und Auswertung von Daten zur Klauengesundheit, Lahmheit und Tierwohl und der Ableitung von betriebsspezifischen Empfehlungen zur Verbesserung aufzubauen. Dazu sollten u.a. verschiedene Merkmale, die in Zusammenhang mit Klauengesundheit und Tierwohl stehen, erfasst werden, um ihre Eignung für die Zuchtwertschätzung und die Entwicklung von neuen Tools für das Herdenmanagement zu untersuchen. Dabei stellte sich die grundsätzliche Frage, wer die Erfassung der potenziellen Merkmale für die Zuchtwertschätzung künftig in der Routine durchführen könnte, um einen validen Datensatz zu erhalten. Die Zuverlässigkeit der Datenerhebung steht hier im Vordergrund.

### **4.2 Auswahl der Beobachter\*innen und Betriebe**

#### **4.2.1 Beobachter\*innen**

Die Beteiligten, die die Beobachtung durchführten, setzten sich zusammen aus

- Landwirt\*innen (LW\*innen)
- Kontrollassistent\*innen (KA\*innen) und
- der Autorin als externer Beurteilerin (EB).

Voraussetzung war die Teilnahme an einer Schulung (siehe Kapitel 4.6).

Die Auswahl der Kontrollassistent\*innen wurde vom Landeskontrollverband (LKV) vorgenommen. Einige der Kontrollassistent\*innen führten schon vor der Erhebung die Milchmessungen auf den Pilotbetrieben durch. Die Landwirt\*innen waren die Betriebsführer\*innen der Pilotbetriebe.

#### **4.2.2 Pilotbetriebe**

Die Auswahl der Pilotbetriebe wurde vom Landeskontrollverband (LKV) durchgeführt. Die Kriterien für die Teilnahme als Pilotbetrieb sowie als Beobachter\*in waren

- Milchviehstall mit Laufstallhaltung ohne Melkroboter
- mind. 20 Milchkühe
- mind. 1 bis 2-malige Durchführung einer Klauenpflege/Jahr (Klauenpfleger\*in)
- die Teilnahme an einer Schulung.

Des Weiteren sollten die Pilotbetriebe die Bereitschaft zu

- einer 2-maligen, selbstständigen Erhebung von Lahmheit und Tierwohlparametern durch Landwirt\*in

- einer 2-maligen Erhebung auf dem Pilotbetrieb durch eine externe Beurteilerin (Masterstudentin) und
- einer 2-maligen Erhebung auf dem Pilotbetrieb durch einen\*r eingeschulte\*n Kontrollassistent\*in im Rahmen der Milchleistungsprüfung

haben.

Die Pilotbetriebe und Kontrollassistent\*innen erhielten einen finanziellen Ausgleich für den zeitlichen Aufwand der Untersuchung. Sie eigneten sich praktisches Wissen in der Erhebung von Parametern zur Tiergesundheit an und erhielten einen Bericht zur Nutzung im Herdenmanagement.

An der Studie nahmen insgesamt 22 Landwirt\*innen sowie 6 Kontrollassistent\*innen teil. Die Betriebe der Landwirt\*innen befanden sich in Niederösterreich (Ybbsitz, Aschbach, Sankt Peter in der Au, Amstetten, Zeillern, St Oswald, Raxendorf, Artstetten, Euratsfeld), Oberösterreich (Waldhausen im Strudengau, Grein) und in der Steiermark (Krieglach, Trofaiach, Passail, Birkfeld, Fladnitz an der Teichalm, St. Kathrein am Offenegg, Kapfenberg). Die Betriebe werden im Folgenden als Pilotbetrieb 1-22 bezeichnet.

### **4.3 Merkmale**

Die Auswahl der zu erhebenden Merkmale lehnte sich an eine Piloterhebung von Winckler und Kofler (2018) zu tierbezogenen Parametern des Wohlergehens im Rahmen des Projektes Klauen-Q-Wohl, an Indikatoren des Tierwohls bei Milchkühen von AssureWel (Advancing Animal Welfare Assurance) und an den Erhebungsbogen Tierwohl Rind von Bio Austria (Bio Austria, 2018) an.

Die Merkmale, die hinsichtlich Zuverlässigkeit der Erhebung berücksichtigt wurden, sind:

- Lahmheit im Gehen (AssureWel)
- Lahmheit im Stand (Brinkmann et al., 2016)
- Abw. Fußstellung (Hulek, 2015)
- Schäden am Sprunggelenk (Brinkmann et al., 2016)
- Verschmutzung (Kofler und Winckler, 2018)
- Körperkondition (Edmondson et al., 1989).

#### **4.3.1 Lahmheit im Gehen (AssureWel)**

- Nicht lahm (0)

Das Rind hat einen elastischen Gang und raumgreifende Schritte. Es sind keine Unregelmäßigkeiten in der Schrittfolge zu beobachten. Alle vier Gliedmaßen werden gleichmäßig belastet.

Das Rind wird ebenfalls als nicht lahm eingestuft, wenn ein so genannter unebener Gang beobachtet wird, z. B. in Form von steifer Hinterhand oder vorsichtigem Fußen. Es sind auch ein langsamer Gang oder insgesamt kurze Schritte möglich; es liegt aber noch keine unregelmäßige Schrittfolge vor.

- Lahm (1)  
Das Rind hat eine unregelmäßige Schrittfolge durch Entlastung eines Beines. Es kann eine ungleichmäßige Belastung eines unmittelbar erkennbaren und/oder offensichtlich verkürzten Beines beobachtet werden. Diese beeinträchtigte Mobilität kann mit einem aufgekrümmten Rücken verbunden sein.
- Hochgradig lahm (2)  
Das Rind zeigt deutliches Widerstreben ein Bein zu belasten oder Entlastung von mehr als einem Bein. Das Rind hat Anzeichen einer stark eingeschränkten Mobilität.

Lahme Tiere zeigen möglicherweise einen aufgekrümmten Rücken; dies kann jedoch bei Rindern der Rasse Fleckvieh häufig nicht beobachtet werden (Bio Austria).

#### Erhebung des Merkmals Lahmheit im Gehen:

Der Bewegungsablauf des Tieres wurde im Gehen auf festem Untergrund im Laufgang beurteilt. Das Rind wurde, wenn ein Fressgitter vorhanden, aus der Fixierung gelöst und zum Gehen animiert. Wenn das Tier nicht fixiert werden konnte, z. B. bei Fressplatzbegrenzung nach vorne nur mit Nackenriegel oder wenn sich das Rind in einem anderen Bereich des Stalls befand, wurden die Tiere am jeweiligen Aufenthaltsort animiert zu gehen. Der\*die Beobachter\*in beurteilte den Bewegungsablauf des Rindes von hinten bzw. von der Seite und ging ggfs. hinter dem Tier her. Lag eine Lahmheit vor, wurde dies im Erhebungsprotokoll mit 1 (lahm) bzw. 2 (hochgradig lahm) festgehalten. Wurde keine Lahmheit beobachtet, wurde im Erhebungsprotokoll in der Zeile für das beobachtete Rind eine 0 eingetragen.

#### **4.3.2 Lahmheit im Stand (Brinkmann et al., 2016)**

- Nicht lahm (0)
- Lahm (1)  
Das Rind wird als lahm beurteilt, wenn eines der drei folgenden Kriterien zutrifft:
  - ‚Herumtrippeln‘ - Wiederholte Gewichtsverlagerung von einem Bein auf das andere,
  - wiederholtes Anheben einer Gliedmaße oder
  - Belastung nur eines Teils der Klaue (z. B. Aufsetzen der Klauenspitze).



Abb. 2: Merkmal Lahmheit im Stand – lahm (1). Quelle: © Solveig-March.  
<https://www.bio-austria.at/lahmheit/>

#### Erhebung des Merkmals Lahmheit im Stand:

Das Merkmal wurde in der Regel am fixierten Tier (im Fressgitter bzw. im Melkstand) erhoben. Falls eine Fixierung im Laufstall nicht möglich war, wurden die Tiere in jenem Stallbereich beobachtet, wo sie sich aufhielten. Es wurden die hinteren Extremitäten des Rindes beobachtet. Lag eines der oben genannten Kriterien für Lahmheit im Stand vor, so wurde das Rind als lahm eingestuft und im Erhebungsprotokoll mit 1 (lahm) notiert. Zeigte das Rind keine Lahmheit im Stand, notierte der\*die Beobachter\*in im Erhebungsprotokoll in der entsprechenden Zeile für das Tier die Ziffer 0 (nicht lahm).

#### **4.3.3 Abweichende Fußstellung (Hulek, 2015)**

- Gut/akzeptabel (0)
  - Gut: Die Hinterbeine stehen (fast) parallel.
  - Akzeptabel: Die Hinterbeine sind leicht nach außen gestellt.
- Abweichende Fußstellung (1): Die Hinterbeine sind nach außen gestellt und Drehung der Klauen nach außen.

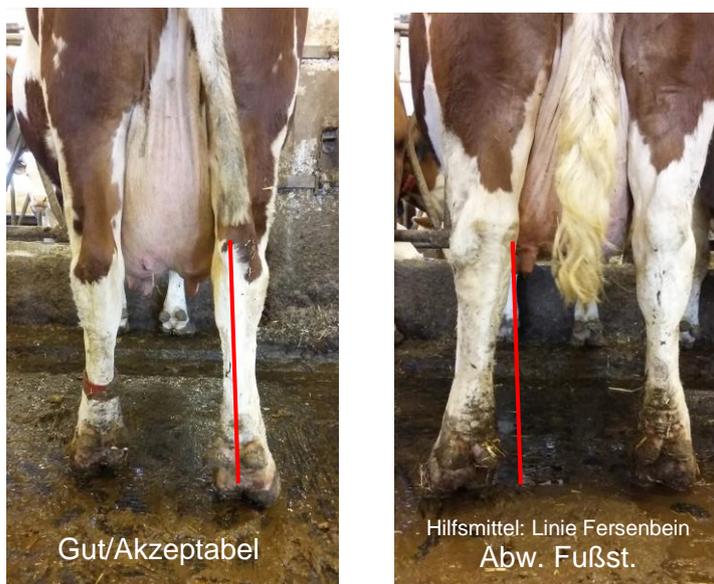


Abb. 3: Merkmal „Gute/akzeptable Fußstellung“ (links - 0) und „abweichende Fußstellung“ (rechts - 1) an den hinteren Extremitäten beim Rind. Die rote Linie am Fersenbein dient als Hilfsmittel zu Erhebung der abweichenden Fußstellung. Quelle: © Riegler

### Durchführung der Erhebung des Merkmals Beurteilung von abweichender Fußstellung:

Das Merkmal abweichende Fußstellung wurde entweder beim fixierten Tier (im Fressgitter oder im Melkstand) oder beim stehenden Tier im Laufgang (falls eine Fixierung nicht möglich war) erhoben. Es wurden die hinteren Extremitäten des Rindes beobachtet. Die Messung der von Hulek (2015) vorgeschlagenen Winkel zur Einstufung einer abweichenden Fußstellung als ‚gut (0 -17°)/akzeptabel (17 – 23°)‘ bzw. ‚zu weit (> 23°)‘ war bei der Direktbeobachtung im Stall nicht möglich. Als Hilfsmittel (siehe Abbildung 3) für die Erhebung der abweichenden Fußstellung der Hinterbeine diente eine gedachte, von der Mitte des Fersenbeinhöckers zum Boden führende Senkrechte. Berührte diese gedachte Linie die Klauen, so wurde eine ‚gute/akzeptable‘ Fußstellung angenommen. Berührte die Linie die Klauen nicht mehr, kam es zu einer ‚abweichenden Fußstellung‘. Wurde vom\*von der Beobachter\*in das Tier mit einer guten/akzeptablen Fußstellung eingestuft, wurde dies im Erhebungsprotokoll in der Zeile für das Rind mit der Ziffer 0 vermerkt. Wurde eine abweichende Fußstellung festgestellt, notierte der\*die Beobachter\*in eine 1.

#### **4.3.4 Schäden am Sprunggelenk - Integumentschäden (Brinkmann et al., 2016)**

Zu den Integumentschäden zählen

- haarlose Stellen (1): ab einem Durchmesser von 2,5 cm (1 € Münze)
- Wunden (1): frisch oder verkrustet, jeweils >2,5 cm (1 € Münze)
- Schwellungen (1): eindeutig mit bloßem Auge erkennbare Umfangsvermehrung im Vergleich zum Normalzustand
- Keine Schäden am Sprunggelenk (0)



Abb. 5: Merkmal Schäden am Sprunggelenk – Haarlose Stellen (1).

Quelle: © BOKU Winckler



Abb. 4: Merkmal Schäden am Sprunggelenk – Schwellung rechts (1)



Abb. 6: Merkmal Schäden am Sprunggelenk – Wunde verkrustet (1).  
Quelle: © BOKU Winckler



Abb. 7: Merkmal Schäden am Sprunggelenk – Wunde frisch (1).  
Quelle: © BOKU Winckler

Durchführung der Erhebung der Merkmale hinsichtlich Schäden am Sprunggelenk (haarlose Stellen, Wunden und Schwellungen):

Schäden im Bereich des Sprunggelenkes wurden beim fixierten Tier (im Fressgitter bzw. im Melkstand) oder beim stehenden Tier im Laufgang (falls eine Fixierung nicht möglich war) erhoben. Es wurden beide Körperseiten im Bereich des Sprunggelenks des Rindes beobachtet und Schäden am Sprunggelenk sowohl auf der rechten, sowie auf der linken Seite erhoben. Zur rechten Seite zählte die rechte Außenseite des Sprunggelenks sowie die gegenüberliegende Innenseite. Zur linken Seite zählte die linke Außenseite des Sprunggelenks sowie die gegenüberliegende Innenseite (siehe Abbildung 8). Als Hilfsmittel für die Erhebung zur Abschätzung von haarlosen Stellen und Wunden/Krusten diente eine 1-Euro-Münze (2,5 cm Durchmesser). Wurde eine haarlose Stelle und/oder eine Wunde und/oder eine Schwellung festgestellt, so trug der\*die Beobachter\*in die Ziffer 1 auf der rechten und/oder auf der linken Körperseite in das Erhebungsprotokoll ein. Wurde keine haarlose Stelle und/oder Wunde und/oder Schwellung beobachtet, so trug der\*die Beobachter\*in die Ziffer 0 auf der rechten und/oder auf der linken Körperseite in das Erhebungsprotokoll ein.

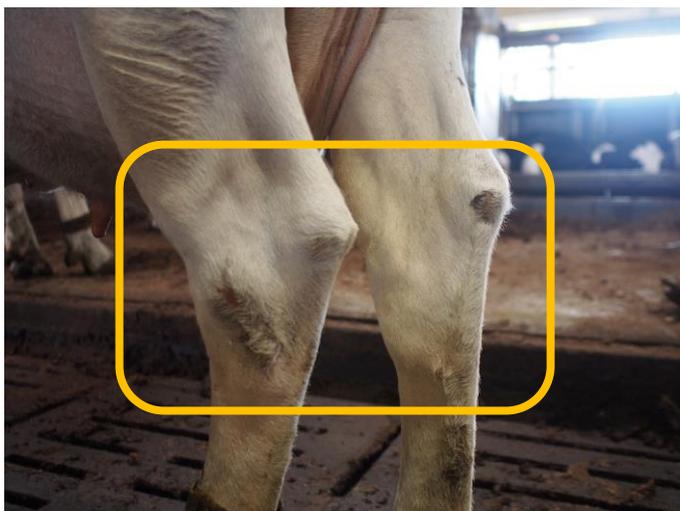


Abb. 8: Merkmal Schäden am Sprunggelenk – Haarlose Stelle auf der linken Körperseite (1). Die Markierung zeigt den untersuchten Bereich (Außenseite des linken Sprunggelenks sowie Innenseite des gegenüberliegenden, rechten Sprunggelenks). Quelle: © BOKU Winckler

#### 4.3.5 Verschmutzung (Kofler und Winckler, 2018)

Bei der Verschmutzung wurde in 2 Körperregionen unterschieden, nämlich in unteres bzw. oberes Hinterbein.

- „Sauber“ (0):  
Die jeweilige Körperregion wird als sauber eingestuft, wenn keine Verschmutzungen bzw. nur nasses Fell oder Verfärbungen des Haarkleids ohne Kotalagerung vorkommen.
- „Verschmutzt“ (1):  
Verschmutzung liegt dann vor, wenn das obere Hinterbein inkl. Schwanz ohne Quaste eine mind. 30 cm lange (ca. Unterarmlänge), durchgehende Kotalage aufweist bzw. das untere Hinterbein inkl. Innenseite des gegenüberliegenden Hinterbeins in Summe mind. handtellergröße Kotalagen aufweist.

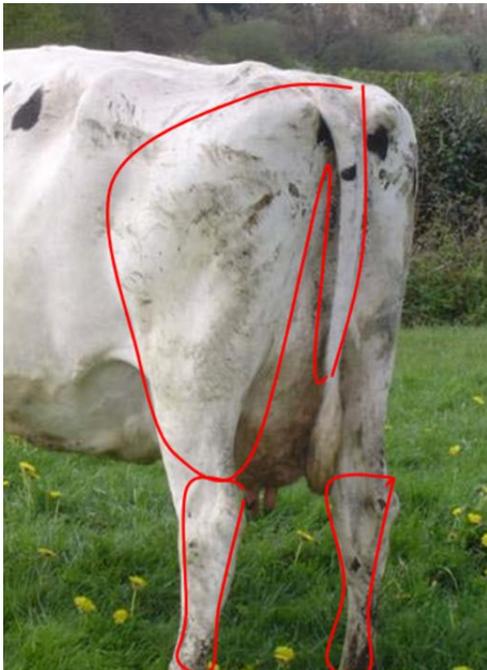


Abb. 9: Merkmal Verschmutzung – Sauber (0) - Die Markierung zeigt das obere und untere Hinterbein auf der linken Körperseite.

Quelle: © BOKU Winckler

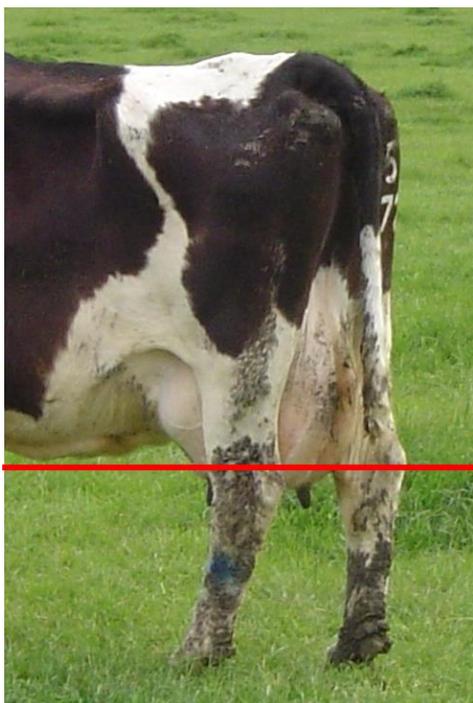


Abb. 10: Merkmal Verschmutzung – Verschmutzt am linken oberen Hinterbein (1). Verschmutzt am linken unteren Hinterbein (1), sowohl auf der linken Außenseite als auch auf der gegenüberliegenden Innenseite Die Markierung zeigt die Abgrenzung zwischen dem oberen und unteren Hinterbein auf der linken Körperseite.

Quelle: © BOKU Winckler

#### Durchführung der Erhebung des Merkmals Verschmutzung (am oberen und unteren Hinterbein):

Das Merkmal Verschmutzung wurde beim fixierten Tier (im Fressgitter bzw. im Melkstand) oder beim stehenden Tier im Laufgang (falls eine Fixierung nicht möglich war) erhoben. Im Laufgang wurden beide Körperseiten am oberen und unteren Hinterbein beim Rind beobachtet und Verschmutzung sowohl auf der rechten als auch auf der linken Körperseite erhoben. Im Melkstand wurden beide Körperseiten am unteren Hinterbein beim Rind beobachtet und Verschmutzung sowohl auf der rechten als auch auf der linken Körperseite erhoben.

Zum oberen Hinterbein zählte die (rechte bzw. linke) Außenseite oberhalb des Sprunggelenks bis zum Schwanzansatz inkl. Schwanz ohne Quaste. Zum unteren Hinterbein zählte die (rechte bzw. linke) Außenseite vom Kronsaum bis einschließlich Sprunggelenk sowie die Innenseite des jeweils gegenüberliegenden Hinterbeins (siehe Abbildung 9).

Wurde eine Verschmutzung am oberen und/oder unteren Hinterbein festgestellt, so trug der\*die Beobachter\*in die Ziffer 1 auf der rechten und/oder auf der linken Körperseite in das Erhebungsprotokoll ein. Wurde keine Verschmutzung beobachtet, so trug der\*die Beobachter\*in die Ziffer 0 auf der rechten und/oder auf der linken Körperseite in das Erhebungsprotokoll ein.

#### **4.3.6 Körperkondition - Body Condition Score**

Die Erhebung der Körperkondition wurde von den Landwirt\*innen und der externen Beurteilerin im Laufgang durchgeführt. Dieses Merkmal wurde nicht von den Kontrollassistent\*innen im Melkstand erhoben. Die Bewertung der subkutanen Fettauflage erfolgte an vier Körperregionen mittels visueller Beurteilung (Edmondson et al., 1989):

1. Quer- und Dornfortsätze der Wirbelsäule - Verbindungslinie Quer- und Dornfortsätze
2. Bereich zwischen den Hüftbeinhöckern
3. Bereich zwischen Hüft- und Sitzbeinhöcker
4. Vertiefung zwischen Schwanzansatz und Sitzbeinhöcker

Es wurden für jede Körperregion Punkte von 1,0 (hochgradig mager) bis 5,0 (hochgradig verfettet) in 0,5-Punkt-Schritten vergeben.

#### Durchführung der Erhebung des Merkmals Körperkondition:

Das Merkmal Körperkondition wurde am fixierten Tier (im Fressgitter) oder beim stehenden Tier im Laufgang (falls eine Fixierung nicht möglich war) von den Landwirt\*innen und der externen Beurteilerin erhoben. Die einzelnen Körperregionen wurden separat beurteilt. Der\*die Beobachter\*in schrieb den entsprechenden Wert in das Erhebungsprotokoll (z. B. Spalte für Körperregion Quer- und Dornfortsätze) und erhob die nächste Körperregion. Zum Abschluss konnte in der Spalte für Durchschnitt im Erhebungsprotokoll vom\*von der Beobachter\*in eine durchschnittliche Körperkondition eingetragen werden.



Abb. 11: 4 Körperbereiche der optischen Bewertung von Körperkondition.  
Quelle: © Raumberg Gumpenstein Häusler. Kuhsignale – Körperkondition

BODY CONDITION SCORE		Verbindungs- linie Dorn- zu Querfort- sätzen		Hinteransicht Hüftbeinhöcker		Seitenansicht der Verbindungs- linie zw. Hüft- u. Sitzbeinhöcker		Höhle zwischen Schwanzansatz u. Sitzbeinhöcker	
								Hinter- ansicht    Seiten- ansicht	
<b>1</b>	hochgradig abgemagert								
<b>2</b>	Knochenvor- sprünge sichtbar								
<b>3</b>	Knochenvor- sprünge gut abgedeckt								
<b>4</b>	Knochenvor- sprünge angedeutet								
<b>5</b>	hochgradig verfettet								

Abb. 12: BCS – Optische Beurteilung. Optische Bewertung der subkutanen Fettschicht an vier Körperregionen. Punktevergabe 1,0 – 5,0 in 0,5-Punkt-Schritten.  
Quelle: Edmondson et al., 1989





## **4.5 Training der externen Beurteilerin**

Die externe Beurteilerin wurde vom Betreuer der Masterarbeit, der gleichzeitig Experte für Tierwohlerhebungen sowie Ansprechpartner für tiergerechte Tierhaltung und Tierschutz ist, eingeschult. Diese Einschulung fand auf der Vetfarm Kremesberg der Veterinärmedizinischen Universität statt. Dabei wurden alle Merkmale an den Milchkühen erhoben und die Ergebnisse miteinander verglichen und diskutiert. Die externe Beurteilerin nahm ebenfalls an den Schulungen für die Beobachter\*innen teil (s. 4.6) – insgesamt erfolgte die Teilnahme an vier Schulungen. Diese Schulungen wurden vom Experten und teilw. von der externen Beurteilerin abgehalten.

Bevor die Erhebungen auf den Pilotbetrieben starteten, wurde die Beobachter\*innenübereinstimmung der externen Beurteilerin mit dem erfahrenen Beurteiler (Silberstandard) überprüft. Dazu wurde durch beide Personen eine voneinander unabhängige Erhebung aller Merkmale im März 2019 bei 37 Rindern auf Pilotbetrieb 6 durchgeführt. Im Oktober 2019 erfolgte zum Abschluss der Praxisphase bei 36 Rindern auf Pilotbetrieb 22 erneut ein Beobachter\*innenabgleich.

## **4.6 Schulungen**

Voraussetzung für die Teilnahme am Projekt war die Teilnahme an einer halbtägigen Schulung der Beobachter\*innen. Die Schulung der Landwirt\*innen, Kontrollassistent\*innen sowie der externen Beurteilerin erfolgte an insgesamt vier Terminen, wovon zwei in der Steiermark (LFS Hafendorf) und zwei in Niederösterreich (Hblfa Francisco Josephinum) stattfanden. Die Teilnehmer\*innenanzahl lag jeweils zwischen 6 und 14 Personen.

Die Schulung bestand aus drei Teilen. Begonnen wurde mit einer theoretischen Vorbesprechung (ca. 1 Stunde). Es wurden das Projekt Klauen-Q-Wohl vorgestellt, die zu erhebenden Merkmale anhand von Fotos und Videos erklärt, das Erhebungsprotokoll gezeigt und erläutert, wer welche Merkmale, in welchen zeitlichen Abständen an welchem Ort zu erfassen hat. Die Schulungsteilnehmer\*innen erhielten Schulungsunterlagen (Erklärung und Anleitung der zu erhebenden Merkmale inklusive Fotos und Bilder) und eine laminierte Scoring Card (siehe Anhang). Im zweiten Teil der Schulung wurden die Merkmale im Stall direkt am Tier erhoben (ca. 1 Stunde). Im dritten Teil der Schulung wurde ein Abgleich der Beurteilung anhand von Videos/Fotos durchgeführt (ca. 1 Stunde). Die Zusammenstellung der Videos/Fotos erfolgte durch den Experten. Dabei wurden von den Teilnehmer\*innen das Vorkommen (1) bzw. nicht Vorkommen (0) von Lahmheit im Gehen und im Stand, Verschmutzungen, Wunden/Krusten, haarlose Stellen und Schwellungen erfasst. Jede\*r Teilnehmer\*in beurteilte für sich das am Video oder Foto zu sehende Merkmal und dokumentierte das Ergebnis. Der Experte diente für den Beobachter\*innenabgleich im Rahmen der Schulungen als Silberstandard.

Das Ergebnis hinsichtlich Beobachter\*innenübereinstimmung (Prozent Übereinstimmung mit Silberstandard, Abweichungen in Prozentpunkten vom Silberstandard, Kappa; Auswertung s. Kap. 5) wurde den Teilnehmer\*innen im Anschluss übermittelt (siehe Tabelle 4).

Tab. 4: Beispielhafte Darstellung des Schulungsergebnisses eine\*r Teilnehmer\*in der Schulung

Beobachter XY	Lahmheit Beurteilung im Gehen	Lahmheit Beurteilung im Stand	Sprungelenk Haarlose Stellen	Sprungelenk Wunden/Krusten	Sprungelenk Schwellungen	Verschmutzung Oberes Hinterbein	Verschmutzung Unteres Hinterbein
Prozentuelle Übereinstimmung (Zielwert mind. 80%)	85%	90%	60%	80%	73%	95%	89%
Kappa (Maß für Übereinstimmung, Zielwert mind. 0,40)	0,82	0,80	0,25	0,53	0,44	0,88	0,79
Abweichung vom Silberstandard (Prozentpunkte)	5	10	27	7	27	-2	-11

Zusätzlich zu den Schulungsergebnissen wurde den Teilnehmer\*innen eine Erklärung zur Interpretation der Ergebnisse und ein Empfehlungsschreiben beigelegt, um sie in Abhängigkeit von den individuellen Ergebnissen auf Aspekte hinzuweisen, auf die bei der eigentlichen Datenerhebung besonders geachtet werden sollte:

*Bei sechs von sieben Merkmalen besteht gute bis sehr gute Übereinstimmung; auch das zentrale Merkmal Lahmheit wurde sehr gut bis gut erkannt. Bitte noch einmal die Kriterien für haarlose Stellen und Schwellungen am Sprunggelenk beachten. Hier liegt jeweils eine Überschätzung vor (positive Werte in der Zeile ‚Abweichung vom Silberstandard‘), d.h. es wurden mehr Tiere als betroffen eingestuft als durch einen erfahrenen Beurteiler (=Silberstandard). Dabei ist allerdings auch zu beachten, dass haarlose Stellen vom Foto schwer zu beurteilen sind.*



Abb. 13: Schulungsraum der zweiten Schulung in der Hblfa Francisco Josephinum. Quelle: © Riegler



Abb. 14: Praktischer Teil der ersten Schulung am Lehrbetrieb der Lfs Hafendorf. Quelle: © Suntinger



Abb. 16: Praktischer Teil der zweiten Schulung am Lehrbetrieb der Hblfa Francisco Josephinum. Quelle: © Suntinger



Abb. 15: Externe Beurteilerin - Erklärung der Merkmale direkt am Tier am Lehrbetrieb der Hblfa Francisco Josephinum. Quelle: © Suntinger

## 4.7 Durchführung der Erhebungen

Die Merkmalerhebungen durch KA\*in, LW\*in und externer Beurteilerin fanden auf den Pilotbetrieben in insgesamt zwei Durchgängen statt. Nachdem die Beobachter\*innen an der Schulung teilgenommen hatten, begann ausgehend von der Milchmessung die erste Erhebung durch den\*die KA\*in auf den Pilotbetrieben. Innerhalb von maximal einer Woche erhoben dann der\*die LW\*in und die externe Beurteilerin an unterschiedlichen Tagen die Merkmale. Die zweite Erhebungsrunde fand bei der nächsten Milchmessung (im Abstand von 4 bis 5 Wochen) statt.

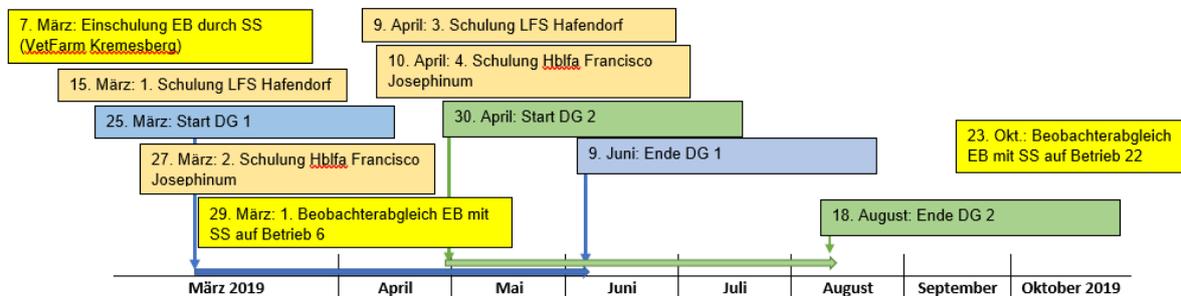


Abb. 17: Zeitlicher Ablauf der Durchführung des Beobachterabgleichs, der Schulungen und der Erhebungen auf den Pilotbetrieben

#### **4.7.1 Tiere**

Die Erhebungen wurden auf den Pilotbetrieben von den Beobachter\*innen an allen laktierenden Kühen durchgeführt. Es handelte sich dabei hauptsächlich um Tiere der Rassen Fleckvieh, Holstein Frisian und Braunvieh. Trockenstehende Kühe und Kalbinnen wurden bei der Erhebung nicht berücksichtigt. Die Identifikation erfolgte durch das Ablesen der 4-stelligen Ohrmarkennummer. Beim ersten Durchgang wurde die Erhebung an 814 Milchkühen durchgeführt (im Mittel 37 Tiere pro Betrieb, 19-60 Tiere). Im zweiten Durchgang wurde die Erhebung an 723 Milchkühen durchgeführt (im Mittel 33 Tiere pro Betrieb, 16-57 Tiere).

#### **4.7.2 Vorgehensweise Kontrollassistent\*in**

Die Erhebung der Merkmale erfolgte durch den\*die Kontrollassistent\*in im Melkstand. Nach Anlegen des Melkzeugs wurden folgende Merkmale erhoben:

1. Lahmheit - Beurteilung im Stand (Melkstand)
2. Abweichende Fußstellung
3. Schäden am Sprunggelenk
  - Haarlose Stellen
  - Wunden/Krusten
  - Schwellungen
4. Verschmutzung (unteres Hinterbein und gegenüberliegende Innenseite)

Aufgrund des eingeschränkten Sichtfeldes im Melkstand konnte der\*die Kontrollassistent\*in, im Vergleich zur externen Beurteilerin und Landwirt\*in, nicht alle Merkmale erheben. Die Beurteilung des Ganges war nicht möglich, da die Tiere vor dem Melkstand standen und keine Lauffläche zur Verfügung hatten. Des Weiteren konnte es Stufen und enge Wendungen geben, was die Lahmheitserkennung beim Eintrieb in den Melkstand erschwerte. Die Erfassung der Verschmutzung war aufgrund des eingeschränkten Sichtfeldes im Melkstand teilweise nicht möglich. Es konnte nur das untere Hinterbein erfasst werden, aber nicht das obere Hinterbein. Die Beurteilung der Körperkondition war ebenfalls durch das eingeschränkte Sichtfeld nicht möglich. Deshalb wurden von dem\*der Kontrollassistent\*in nur die oben genannten Merkmale erhoben. Die Reihenfolge der Merkmalerhebung konnte vom\*von der Beobachter\*in frei ausgewählt werden. Die Ergebnisse wurden an die ZuchtData geschickt, und danach zur externen Beurteilerin für die Auswertung weitergeleitet.

### 4.7.3 Vorgehensweise Landwirt\*in und externe Beurteilerin

Die Erhebung der Merkmale wurde durch den\*die LW\*in und die externe Beurteilerin wie folgt durchgeführt:

1. Lahmheit - Beurteilung im Stand
  2. Abweichende Fußstellung
  3. Körperkondition
  4. Schäden am Sprunggelenk
    - Haarlose Stellen
    - Wunden/Krusten
    - Schwellungen
  5. Verschmutzung
    - oberes Hinterbein inkl. Schwanz ohne Quaste
    - unteres Hinterbein und gegenüberliegende Innenseite
  6. Lahmheit - Beurteilung im Gehen
- 
1. Die Kühe wurden fixiert und die Merkmale bei jeder Kuh nacheinander erhoben
2. Jede Kuh wurde einzeln aus der Fixierung gelöst und im Laufgang beobachtet (ggfs. verfolgt bzw. zum Gehen animiert)

Bevor die Merkmalerhebung begann, wurden die Kühe im Fressgitter fixiert. Bei der Tieridentifikation konnte ausgewählt werden, ob die Stallnummer oder die Ohrmarkennummer (4 Stellen) angegeben wurde. Variante 1: begonnen wurde bei der ersten fixierten Kuh und die Merkmale 1 – 5 wurden erhoben. Danach wurde die Kuh aus der Fixierung gelöst und das Merkmal 6 (Lahmheit - Beurteilung im Gehen) erhoben. Dann wurden die Merkmale 1 – 6 bei der zweiten Kuh erhoben und dieser Vorgang nacheinander bei allen Kühen im Stall durchgeführt. Variante 2: bei allen Kühen wurden zuerst die Merkmale 1 - 5 erhoben. Danach wurde Merkmal 6 (Lahmheit im Gehen) bei allen vorher beurteilten Tieren erhoben.

Das Merkmal 4 „Schäden am Sprunggelenk“ sowie das Merkmal 5 „Verschmutzung“ sollte jeweils auf der rechten und auf der linken Körperseite erhoben werden. Die Ergebnisse sollten für die entsprechende Körperseite im Erhebungsprotokoll eingetragen werden. Nachdem die Erhebung abgeschlossen war, wurde das Erhebungsprotokoll an die externe Beurteilerin übergeben. Die Ergebnisse der Landwirte\*innen wurden entweder am Tag der Erhebung der externen Beurteilerin übergeben (falls der\*die Landwirt\*in die Erhebung bereits durchgeführt hatte) oder per Mail zur Auswertung zugeschickt.

## **4.8 Auswertung der Daten und Interpretation der Ergebnisse**

Die Ergebnisse der Schulungs-Teilnehmer\*innen wurden in Excel-Tabellen eingetragen und mittels Excel ausgewertet. Berechnet wurden die Auftretenshäufigkeit in den Pilotbetrieben (s. 4.8.1), die Beobachter\*innenübereinstimmung (s. 4.8.2.), die Abweichung in %-Punkten Prävalenz vom Silberstandard (s. 4.8.3) sowie die Genauigkeit der Erfassung hochgradig lahmer Tiere (auf Basis Gangbeurteilung durch externe Beurteilerin) mittels Beurteilung im Stand durch KA\*in/LW\*in (s. 4.8.4).

Die Ergebnisse der Erhebungen durch Landwirt\*in, Kontrollassistent\*in und externe Beurteilerin wurden in Excel-Tabellen eingetragen und mittels Excel und SAS ausgewertet. Berechnet wurden die Auftretenshäufigkeiten (Prävalenzen), Beobachter\*innenübereinstimmung in Prozent, Abweichungen in der Prävalenz (%-Punkte) und PABAK. Die Übereinstimmung Lahmheitsbeurteilung im Stand mit der Erfassung von hochgradiger Lahmheit anhand Gangbeurteilung erfolgte mittels der Berechnung von Sensitivität, Spezifität sowie positive and negative predictive value. Da bei der Erhebung der Merkmale haarlose Stellen, Wunden und Krusten, Schwellungen sowie Verschmutzung zwischen rechter und linker Körperseite unterschieden wurde, lagen für die genannten Merkmale pro Tier und Durchgang zwei Datensätze vor; diese wurden in der Auswertung als unabhängige Beobachtungen betrachtet.

Von allen 22 Betrieben wurden im ersten und zweiten Durchgang die Auftretenshäufigkeiten (Prävalenzen - erhoben durch die externe Beurteilerin) berechnet.

Von den 22 Betrieben mussten 3-4 Betriebe aufgrund fehlender Daten ausgeschlossen werden. Entweder fehlten Daten (Betriebe 3, 11 und 12 in beiden Durchgängen, Betrieb 19 in DG 1) oder die Dokumentation der Ergebnisse ließ keine Auswertung zu. Betrieb 19 konnte im ersten DG nicht ausgewertet werden. Bei Betrieb 11 und 20 wurde vom\*von der KA\*in die Tierseite für die Integumentveränderungen am Sprunggelenk und Verschmutzung nicht angegeben. Bei der Auswertung des Erhebungsprotokolls des\*der KA\*in konnten daher bei diesen Betrieben nur die Daten für die Merkmale Lahmheit im Stand und abw. Fußstellung verwendet werden.

Im ersten Durchgang konnten daher nur für 18 Pilotbetriebe und im zweiten Durchgang nur für 19 Pilotbetriebe die Auswertungen für die Beobachter\*innenübereinstimmung, die Abweichung in %-Punkten sowie die Genauigkeit der Erfassung hochgradig lahmer Tiere (auf Basis Gangbeurteilung durch externe Beurteilerin) mittels Beurteilung im Stand durch KA\*in/LW\*in durchgeführt werden.

KA\*in 5 und 6 führten jeweils nur eine Erhebung pro Durchgang durch. Daher wurde für diese KA\*innen keine Abweichung in %-Punkten berechnet (s. Kapitel 5.3.3.).

#### **4.8.1 Berechnung der Prävalenzen (DG 1 auf DG 2)**

Zur Beschreibung der Auftretenshäufigkeiten (Prävalenzen) in den Pilotbetrieben wurden die Ergebnisse der externen Beurteilerin herangezogen. Die Auftretenshäufigkeit wurde berechnet, indem die Gesamtanzahl der betroffenen Tiere (z. B. LH\_Stand) durch die Gesamtanzahl der beobachteten Tiere dividiert wurde. Die Abbildung dieser Ergebnisse erfolgt in Boxplots. Angezeigt wird der Median, Mittelwert, die größte Beobachtung (ohne Ausreißer), die kleinste Beobachtung, die Quantile (25 %, 50 % und 75 %) und ggfs. Ausreißer.

Um zu überprüfen, ob sich die Prävalenzen zwischen erstem und zweitem Durchgang signifikant veränderten, wurde jeweils die Differenz zwischen erstem und zweitem Durchgang berechnet (LH\_Stand 1 – LH\_Stand 2) und dann getestet, ob sich die Differenzen von Null unterscheiden. Die Auswertung wurde mittels Prozedur TTEST für verbundene Stichproben in SAS 9.4 durchgeführt.

#### **4.8.2 Berechnung der Beobachter\*innenübereinstimmung**

##### **4.8.2.1 Beobachter\*innenübereinstimmung in Prozent (%ÜE):**

In verschiedenen Studien wird die Übereinstimmung der Ergebnisse der jeweiligen Beobachter\*innen mit einem Silberstandard in Prozent ausgedrückt (Pritchard et al., 2007; Channon et al., 2009). Burn et al. (2009) geben für die Tierwohlbeurteilung bei Pferden mittels nominaler Variablen einen Grenzwert von 75 % für ausreichende Übereinstimmung an.

Das Projekt AssureWel (2016) bietet auf einem geschützten Bereich der Webseite <http://assurewel.org/> ein Trainings-Tool für Auditor\*innen aus Qualitätssicherungsprogrammen an. Es stehen im Bereich Legehennen, Milchkühe und Schweine detaillierte Informationen, Erhebungsprotokolle sowie interaktive Tests zu ausgewählten tierbezogenen Parametern zur Verfügung. Die Auditor\*innen müssen verschiedene Online-Tests durchführen, wobei die Übereinstimmung für die Zulassung zu den Betriebserhebungen nach dem Protokoll von AssureWel mindestens 80% betragen muss (AssureWel, 2016, zit. n. Schenkenfelder, 2016).

Zur Interpretation der % Übereinstimmung wurden daher folgende Kategorien verwendet:

- 100 – 80 %= Übereinstimmung sehr gut bis gut
- 79 – 70%= Übereinstimmung noch akzeptabel
- <70 % = Übereinstimmung unbefriedigend

##### **4.8.2.2 Kappa und PABAK**

Cohen's Kappa ist ein Maß für die Übereinstimmung bei der Verwendung von diskreten Daten (z. B.: vorkommen/nicht vorkommen 0/1, ordinale Skalen – BCS), wenn zwei Beobachter\*innen dasselbe beurteilen. Der Kappa-Koeffizient wird häufig für die Ermittlung der Beobachter\*innenübereinstimmung verwendet (Cohen, 1960).

### Formel für Kappa (Cohen, 1960)

Für die Berechnung des Kappa's werden die Ergebnisse der Erhebungen in einer 2x2 Tabelle erfasst (s. Tabelle 5).

Tab. 5: Erfassung der Ergebnisse für die Berechnung des Kappas (a-d: beobachtete Häufigkeiten)

		Beobachter*in 1	
		JA	NEIN
Beobachter*in 2	JA	a	b
	NEIN	c	d

Zelle a enthält die Beobachtungen, für die sich beide Beobachter\*innen über das Vorkommen eines Merkmals einig sind. Zelle b beschreibt die Zahl der Fälle, in denen laut Beobachter\*in 1 das Merkmal vorliegt, laut Beobachter\*in 2 jedoch nicht; umgekehrt in Zelle c. In der Zelle d sind sich beide Beobachter\*innen wieder einig, dass das Merkmal nicht vorkommt. Der Kappa-Koeffizient (k) ergibt sich aus der Formel

$$\text{Kappa (k)} = (p_0 - p_e) / (1 - p_e)$$

wobei sich die beobachtete Übereinstimmung aus  $p_0 = (a+d)/N$  und die erwartete Übereinstimmung aus  $p_e = (a+b)/N * (a+c)/N + ((b+d)/N * (c+d)/N$  berechnet.

Der Koeffizient kann Werte zwischen -1 und 1 annehmen:

- 0 % Übereinstimmung, dann -1
- 100 % Übereinstimmung, dann 1
- wenn Übereinstimmung 50 %, dann 0

Die Interpretation des Kappas kann schwierig sein, da dieser Wert abhängig ist von der Häufigkeit bzw. vom Vorkommen des Merkmals (höhere Wahrscheinlichkeit für Übereinstimmung wenn Vorkommen über 50 %). Des Weiteren ist er abhängig von der Voreingenommenheit (Bias) der Beobachter\*innen – (manche Beobachter\*innen beurteilen niedriger/höher als andere).

Einige Autoren\*innen haben daher den sogenannten PABAK (prevalence adjusted bias adjusted kappa) vorgeschlagen, um diese Aspekte zu berücksichtigen (Byrt et al., 1993; Siegel und Castellan, 1988). Die Formel für den PABAK lautet:  $\text{PABAK} = (2p_0 - 1)$  (Byrt et al., 1993).

In Anlehnung an Landis und Koch (1977) können für Kappa-Koeffizienten (einschließlich PABAK) folgende Kategorien für die Interpretation der Übereinstimmung herangezogen werden:

- 0,61 – 1 Übereinstimmung gut bis sehr gut
- 0,41 - 0,60 Übereinstimmung noch akzeptabel
- <0,40 Übereinstimmung unbefriedigend

Die Interpretation der Ergebnisse der vorliegenden Untersuchung lehnte sich an die Einstufungen nach Landis und Koch (1977) an und berücksichtigte zusätzlich Vorschläge des Kuratoriums für Technik und Bauwesen in der Landwirtschaft - KTBL (2021). Dieses bietet auf der Webseite <https://tierschutzindikatoren-schulung.ktbl.de/> Online-Schulungen sowie Tests der Übereinstimmung für auf landwirtschaftlichen Betrieben zu erhebende Tierschutzindikatoren für Rind, Schwein und Geflügel an.

Als Maß für die Übereinstimmung mit Expert\*innen wird der PABAK herangezogen. Der Test gilt bei einem PABAK  $\geq 0,61$  als bestanden; eine sehr gute Übereinstimmung wird bei einem PABAK  $\geq 0,81$  angenommen. Ein Wert von 0,4 sollte nicht unterschritten werden (Landis und Koch, 1977).

#### **4.8.3 Abweichung in %-Punkten auf Herdenebene von der externen Beurteilerin (Abw. in %-Punkte)**

Wie in der Studie von Mullan et al. (2011) wurde die Abweichung in %-Punkten von der durch die externe Beurteilerin ermittelten Prävalenz auf Herdenebene berechnet. Eine negative Abweichung bedeutet, dass der\*die Landwirt\*in/Kontrollassistent\*in bei weniger Tieren das Merkmal erhoben hat als die externe Beurteilerin. Wurde von dem\*der Landwirt\*in bzw. KA\*in das Merkmal häufiger erhoben, kommt es zu einer positiven Abweichung – insgesamt wurde eine Abweichung von 0 angestrebt.

Um zu überprüfen, ob die mittlere Abweichung sich signifikant von 0 unterschied (als Hinweis auf eine gerichtete Abweichung) wurde in SAS 9.4 die Prozedur UNIVARIATE verwendet.

#### **4.8.4 Zuverlässigkeit der Identifikation hochgradig lahm eingestufte Tiere durch Standbeurteilung durch KA\*in/LW\*in**

Ziel dieser Erhebung war die Untersuchung eines möglichen/geeigneten Hilfsmerkmals für hochgradige Lahmheiten, d.h. die Erkennung von in der Gangbeurteilung als hochgradig lahm beurteilte Tiere anhand einer Beurteilung im Stand durch KA\*in/LW\*in. Dazu wurden die Sensitivität und Spezifität berechnet. Die Sensitivität (Verhältnis von wahren Positiven zur Summe aus wahren Positiven und falsch Negativen) beschreibt die Fähigkeit die wahren Positiven zu erkennen; die Spezifität (Verhältnis von wahren Negativen zur Summe aus wahren Negativen und falsch Positiven) beschreibt die Fähigkeit die wahren Negativen zu erkennen. Dabei wird eine Spezifität und Sensitivität von 1,0 angestrebt (Hoehler, 2019).

Eine dritte Kenngröße/Wert ist der Vorhersagewert (predictive value). Dieser Wert gibt den Anteil der Personen/Tiere mit einem positiven Test an, bei denen bei weiterer Untersuchung die Krankheit tatsächlich festgestellt wird. Der positive Vorhersagewert wird berechnet, indem die Anzahl der positiv getesteten Personen/Tiere, welche später nachweislich die Krankheit haben, durch die Gesamtanzahl positiv getesteter Personen/Tiere dividiert wird. Dieser Wert spiegelt nicht nur die Sensitivität oder Spezifität, sondern auch die Prävalenz/das Vorkommen einer Krankheit in der zu untersuchenden Population/Herde wider. Je größer der Anteil der gesunden Personen/Tieren, desto größer ist die Anzahl der falsch positiven Tests (Forrest, 1990).

Die Berechnung wurde in Excel durchgeführt.

- **True positives:** LW\*in/KA\*in beurteilt Tier im Stand als lahm (Note 1 LH\_Stand) - EB beurteilt Tier im Gang als hochgradig lahm (Note 2 LH\_Gang)
- **True negatives:** LW\*in/KA\*in beurteilt Tier im Stand als nicht lahm (Note 0 LH\_Stand) - EB beurteilt Tier im Gang als nicht hochgradig lahm (Note 0 oder 1 LH\_Gang)
- **False negatives:** LW\*in/KA\*in beurteilt Tier im Stand als nicht lahm (Note 0 LH\_Stand) - EB beurteilt Tier im Gang als hochgradig lahm (Note 2 LH\_Gang)
- **False positives:** LW\*in/KA\*in beurteilt Tier im Stand als lahm (Note 1 LH\_Stand) - EB beurteilt Tier im Gang als nicht hochgradig lahm (Note 0 oder 1 LH\_Gang)

In Tabelle 6 wird ein Beispiel für Sensitivität und Spezifität dargestellt. Die Ergebnisse der Übereinstimmung stammen von Betrieb 1 und berücksichtigen die Erhebungen durch den/die LW\*in. Die LW\*in hat im DG 1 eine Spezifität von 92 % und eine Sensitivität von 50 %. Das heißt 92 % Tiere konnten von dem\*der LW\*in als nicht lahm im Stehen eingestuft werden, die auch laut EB als nicht hochgradig lahm eingestuft wurden (Spezifität). Jedoch wurden vom\*von der LW\*in nur 50 % der Tiere als lahm im Stand erhoben (Sensitivität). D.h. ein Tier, welches von der EB als hochgradig lahm erhoben wurde, wurde als nicht lahm (LH\_Stand) vom\*von der LW\*in erhoben – in diesem Fall kamen laut externer Beurteilerin 2 Tiere von 53 Tiere vor, die hochgradig lahm waren, jedoch erkannte der\*die LW\*in nur ein Tier mit Lahmheit im Stehen. 4 Tiere wurden vom\*von der LW\*in als Lahmheit im Stehen erhoben, aber von der externen Beurteilerin nicht als hochgradig lahm.

Als weitere Werte werden in Tab. 6 der positive (PPV) und negative Vorhersagewert (NPV) aufgelistet. Der PPV für Betrieb 1 lautet 20 % - demnach ist die Trefferwahrscheinlichkeit 20 %, dass jene Tiere, welche vom\*von der LW\*in als lahm im Stehen identifiziert wurden auch tatsächlich hochgradig lahm (von EB als hochgradig lahm eingestuft) sind. Der NPV, hier für Betrieb 1 98 %, gibt an, dass jene Tiere, welche vom\*von der LW\*in als nicht lahm beurteilt wurden auch tatsächlich gesund sind. In diesem Fall wurden vom\*von der LW\*in alle Tiere, bis auf ein Tier richtig eingestuft (gesund – nicht lahm).

Tab. 6: Beispiel für Spezifität vs. Sensitivität sowie positive (PPV) und negative predicitive value (NPV) am Betrieb 1 bei der LW\*in Erhebung

Beurteilung LW*in	LW*in (DG 1) Anzahl Tiere	Spezifität = $TN/(TN+FP)$	Sensitivität = $TP/(TP+FN)$	PPV = $TP/(TP+FP)$	NPV = $TN/(TN+FN)$
True Positives (TP)	1	92%	50%	20%	98%
True Negatives (TN)	47				
False Negatives (FN)	1				
False Positives (FP)	4				

#### **4.8.5 Übermittlung der Ergebnisse an die Landwirt\*innen**

Die Landwirt\*innen erhielten jeweils eine Interpretation der Ergebnisse (siehe Tab. 7) für ihre Betriebe. Diese Ergebnisse auf Einzeltier- und Herdenebene (Prävalenzen - Daten von EB) und Auswertung mit % Übereinstimmung, Abweichung zur externen Beurteilerin wurde den Landwirt\*innen im Jänner 2020 übermittelt.

## Beispiel für die Übermittlung der Ergebnisse für den\*die Landwirt\*in, Betrieb 1

Erhebung externe Beurteilerin im Durchgang 1 am 20.04.2019 und Durchgang 2 am 20.05.2019

Tab. 7: Beispiel für die Darstellung und Interpretation der Ergebnisse für den\*die Landwirt\*in

Betrieb 1	LH_Gang*		LH_Hochgradig*		LH_Stand*		Abw. Fußstellung*		Verschm_OH*		Verschm_UH*		Haarlos*		Wunde*		Schwellung*	
	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
Durchgang (DG)	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
<b>1 Häufigkeiten</b> (vorgekommenen Merkmale in %)	19	40	4	11	4	5	23	25	33	24	30	45	17	13	2	7	0	0
<b>2 Prozentuelle Übereinstimmung</b> (zw. externer Beurteilerin und Landwirt*in – Zielwert mind. 80 %)	74	67	98	86	91	82	79	78	87	79	77	57	86	91	97	93	99	98
<b>3 Abweichungen</b> von externer Beurteilerin (in %-Punkte)	0	-20	-2	0	6	11	-17	-23	10	8	-16	-23	-8	-7	0	-4	1	2

*\*Erläuterung der Abkürzungen:* LH\_Gang = Lahmheit im Gang (Note 0,1,2), LH\_Hochgradig = Kühe mit Lahmheitsscore (Gang) von 2, LH\_Stand = Lahmheitsbeurteilung im Stand (Note 0, 1), Abw. Fußstellung = Kühe mit ‚zu weiter Fußstellung‘ (Note 0,1), Verschm\_OH = Verschmutzung im oberen hinteren Bereich auf der rechten und linken Körperseite (Note 0, 1), Verschm\_UH = Verschmutzung im unteren hinteren Bereich auf der rechten und linken Körperseite (Note 0, 1), Haarlos = Haarlose Stellen im Bereich des Sprunggelenkes rechts und links (Note 0, 1), Wunde = Wunden und Krusten im Bereich des Sprunggelenkes rechts und links (Note 0, 1), Schwellung = Schwellungen im Bereich des Sprunggelenkes rechts und links (Note 0, 1).

### Ergebnisse:

#### 1 Häufigkeiten – Merkmale, welche von der externen Beurteilerin erhoben wurden:

- LH\_Gang – mehr lahme Tiere im 2. DG (40 % der Tiere lahm)
- LH\_Hochgradig – mehr hochgradig lahme Tiere im 2. DG (11 %)
- LH\_Stand – es zeigten 4-5 % Tiere Lahmheit im Stand im DG 1 und 2
- Abw. Fußstellung – ca. ¼ der Tiere mit abweichender Fußstellung im DG 1 und 2

- Verschmutzung – mehr Tiere im unteren hinteren Bereich verschmutzt als im oberen hinteren Bereich in beiden DG
- Haarlose Stellen – weniger Tiere im 2. DG mit haarlosen Stellen (17 % vs. 13 %)
- Wunden – im 2. DG mehr Tiere mit Wunden (7 %)
- Schwellungen – keine Schwellungen vorgekommen

### 2 Prozentuelle Übereinstimmung zur externen Beurteilerin und Abweichungen zur externen Beurteilerin

Bei **fünf** von neun Merkmalen besteht in beiden Durchgängen gute bis sehr gute Übereinstimmung → **grüner Bereich**.

Bei den Merkmalen **Lahmheit im Gehen (DG 1 und 2)**, **abw. Fußstellung (DG 1 und 2)** und **Verschmutzung im oberen (DG 2) und unteren hinteren Bereich (DG 1)** ist die Übereinstimmung unter 80 % → **oranger Bereich**.

Das Merkmal **Verschmutzung im unteren hinteren Bereich** wurde im 2. DG weniger gut erkannt → **roter Bereich**.

### 3 Abweichung von externer Beurteilerin

Bei den Merkmalen **Lahmheit im Gehen (DG 2)**, **abw. Fußstellung (DG 1 und 2)**, **Verschmutzung am unteren Hinterbein (DG 1 und 2)** liegt eine leichte Unterschätzung vor (negative Werte in der Zeile ‚Abweichungen von externer Beurteilerin‘), d.h. es wurden weniger Tiere als betroffen eingestuft als durch die externe Beurteilerin → **gelber Bereich**.

Bei dem Merkmal **Verschmutzung am oberen Hinterbein (DG 2)** liegt eine Überschätzung vor (positive Werte in der Zeile ‚Abweichungen von externer Beurteilerin‘), d.h. es wurden mehr Tiere als betroffen eingestuft als durch die externe Beurteilerin → **gelber Bereich**.

Eine Abweichung von Null von der externen Beurteilerin bedeutet, dass manche Tiere als lahm eingeschätzt wurden, obwohl dies so nicht von der externen Beurteilerin gesehen wurde und umgekehrt Tiere, die als lahm von der externen Beurteilerin eingestuft wurden, als nicht lahm vom\*von der Landwirt\*in angesehen wurden.

**Fazit:** Generell wurden die Merkmale gut erkannt. Wenn in der Tabelle Bereiche **orange** bzw. **rot** markiert wurden, dann bitte die Kriterien für das jeweilige Merkmal noch einmal durchlesen.

## **4.8.6 Befragung der Teilnehmer\*innen**

### **4.8.6.1 Aussagen Landwirt\*innen und Kontrollassistent\*innen**

Nach der Merkmalserhebung durch die externe Beurteilerin im Stall wurden die Erfahrungen mit der Tierwohlerhebung mittels leitfadengestützten Interviews mit dem\*der Landwirt\*in ermittelt. Die externe Beurteilerin erfragte dabei, welche Schwierigkeiten bzw. Unklarheiten bei der Erhebung entstanden sind bzw. welche Vorteile sich durch die Teilnahme am Projekt ‚Klauen-Q-Wohl‘ ergeben.

Bei den Projektsitzungen zum Projekt Klauen-Q-Wohl wurden die Kontrollassistent\*innen gebeten, ein Feedback über die Tierwohlerhebung vom Melkstand aus zu geben.

### **4.8.6.2 Aussagen bei den Abschlussveranstaltungen zu den Piloterhebungen im Projekt Klauen-Q-Wohl**

Anfang Dezember 2019 wurden zwei Abschlussveranstaltungen zu den Piloterhebungen im Projekt Klauen-Q-Wohl durchgeführt. Die erste Abschlussveranstaltung wurde in Karpfenberg für die Betriebe in der Steiermark abgehalten; die zweite Abschlussveranstaltung fand in Wieselburg für die Betriebe in Niederösterreich und Oberösterreich statt. Im Anschluss an die Ergebnispräsentation wurden die Teilnehmer\*innen zu einer Diskussion eingeladen.

## 5 Ergebnisse

### 5.1 Übereinstimmung externe Beurteilerin mit erfahrenem Beurteiler

Tab. 8: Ergebnisse des Beobachter\*innenabgleichs zwischen der externen Beurteilerin (Masterstudentin) und einem erfahrenen Beurteiler (Silberstandard) im März 2019 (N=37) sowie Oktober 2019 (N=36); die Beobachter\*innenübereinstimmung wird als prozentuale Übereinstimmung (%ÜE), Prevalence Adjusted Bias Adjusted Kappa (PABAK) sowie Abweichung in der Herdenprävalenz (%-P. Abw.) ausgedrückt; grün (%ÜE 100-80 %/PABAK >0,6) = sehr gute bis gute Übereinstimmung, orange (%ÜE 79-70 %/PABAK (0,60-0,41) = Übereinstimmung noch akzeptabel, rot (%ÜE <70 %/PABAK 0,40-0,21) = Übereinstimmung unbefriedigend

Merkmale	LH_Gang	LH_Hochgradig	LH_Stand	Abw. Fußst.	Haarlos	Wunde	Schwellung	Verschm_OH	Verschm_UH	BCS gesamt
<b>März 2019</b>										
%ÜE	81	91	94	82	71	89	100	98	77	47
PABAK	0,71	0,82	0,88	0,64	0,42	0,79	1	0,97	0,70	0,33
%-P. Abw. Abw. - SS	-5	-2	1	-4	-20	8	0	-1	6	s. Tab. 9
<b>Okt. 2019</b>										
%ÜE	72	83	75	81	88	83	90	100	82	56
PABAK	0,58	0,67	0,5	0,61	0,75	0,67	0,81	1	0,64	0,45
%-P. Abw. Abw. - SS	0	0	-8	-15	-4	14	-4	0	4	s. Tab. 9

Beim ersten Beobachter\*innenabgleich (DG 1) im März 2019 bestand für die Merkmale LH\_Gang (81 %), LH\_Hochgradig (91 %), LH\_Stand (94 %), abw. Fußst. (82 %), Wunde (89 %), Schwellung (100 %) und Verschmutzung\_OH (98 %) eine gute bis sehr gute Übereinstimmung (Tab. 8). Die Übereinstimmung bei den Merkmalen Haarlos (71 %) und Verschm\_UH (77 %) war noch akzeptabel. Hierbei kam es beim Merkmal Haarlos zu einer Unterschätzung (-20 %-Punkte), das heißt es wurden weniger Tiere als betroffen eingestuft als durch den Silberstandard. Beim Merkmal Verschm\_UH kam es eher zu einer Überschätzung (+6 %-Punkte), das heißt es wurden mehr Tiere als betroffen eingestuft als durch den Silberstandard. Beim zweiten Beobachter\*innenabgleich im Oktober 2019 kam es bei den Merkmalen LH\_Hochgradig (83 %), abw. Fußst. (81 %), Haarlos (88 %), Wunde (83 %), Schwellung (90 %) Verschmutzung\_OH (100 %) und Verschm\_UH (82 %) zu einer guten bis sehr guten Übereinstimmung. Die Übereinstimmung bei den Merkmalen LH\_Stand (75 %) und LH\_Gang (72 %) war noch akzeptabel. Hierbei kam es beim Merkmal LH\_Stand zu einer Unterschätzung (-20 %-Punkte). Beim Merkmal LH\_Gang kam es zu einer Abweichung von 0 %. Eine Abweichung von 0 % vom Silberstandard bei gleichzeitig nicht 100 %iger Übereinstimmung (hier Beobachter\*innenübereinstimmung bei Lahmheitsbeurteilung im Gehen 72 %) bedeutet, dass Tiere als lahm eingestuft wurden, die laut erfahrenem Beurteiler (=Silberstandard) nicht lahm waren, und umgekehrt.

Als weiteres Maß für die Übereinstimmung wurde der PABAK herangezogen. Dieser lag in beiden Durchgängen für die Merkmale LH\_Gang, LH\_Hochgradig, LH\_Stand, abw. Fußst., Haarlos, Wunde, Schwellung, Verschm\_OH und Verschm\_UH zwischen 0,42 und 1 (noch akzeptable Übereinstimmung bis gute/sehr gute Übereinstimmung). Eine lediglich akzeptable Übereinstimmung lag im ersten Durchgang für Haarlos (0,42) sowie im zweiten Durchgang für die Merkmale LH\_Hochgradig (0,5) und LH\_Gang (0,58) vor.

Bei der Körperkonditionsbeurteilung betrug die Beobachter\*innenübereinstimmung in DG 1 47 % (PABAK 0,33) und in DG 2 56 % (PABAK 0,45). Bei der Durchschnittsnote für die vier Körperregionen (BCS gesamt) kam es bei der externen Beurteilerin sowohl zur Überschätzung (positive Werte in der Tabelle 9) als auch zur Unterschätzung (negative Werte in der Tabelle 9).

Tab. 9: Abweichung in %-Punkten der externen Beurteilerin vom Silberstandard für die (gerundete) Gesamtnote BCS (Durchschnitt der 4 Körperregionen)

	Abweichung in %-Punkten						
	BSC 2	BCS 2,5	BCS 3	BCS 3,5	BCS 4	BCS 4,5	BCS 5
DG 1	-2,3	3,1	9,6	10,9	-16,8	-4,5	0
DG 2	2,9	3,0	-1,9	-2,0	3,5	-5,5	0-

Aufgrund der geringen Übereinstimmung zwischen erfahrenem Beurteiler und externer Beurteilerin erfolgte keine weitere Auswertung der in der Untersuchung erhobenen Daten zur Körperkondition.

## 5.2 Auftretenshäufigkeiten (Prävalenzen) im ersten und im zweiten Durchgang (Erhebung durch externe Beurteilerin)

In Abb. 18 sind die Prävalenzen für die Merkmale Lahmheit gesamt/Gangbeurteilung (LH\_Gang), Lahmheit hochgradig/Gangbeurteilung (LH\_Hochgradig), Lahmheit/Beurteilung im Stand (LH\_Stand) sowie abweichende Fußstellung (Abw. Fußst.) dargestellt.

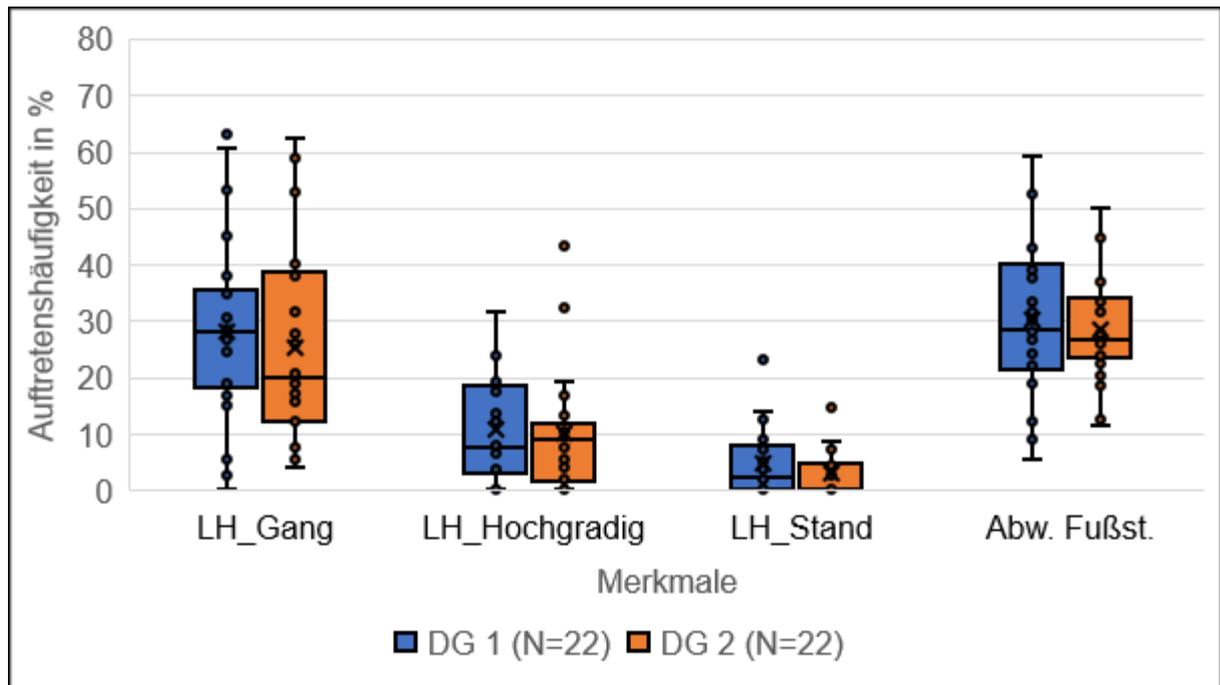


Abb. 18: Auftretenshäufigkeit in % für die Merkmale Lahmheit im Gehen (LH\_Gang), Lahmheit hochgradig (LH\_Hochgradig), Lahmheit im Stand (LH\_Stand) und abweichende Fußstellung (Abw. Fußst.) auf Basis der Erhebung durch die externe Beurteilerin; x=Mittelwert

Im Mittel zeigten 28 % (0-63 %) der Tiere in DG 1 und 25 % (4-62 %) in DG 2 Lahmheit im Gehen. Hochgradige Lahmheit lag im Mittel bei 11 % der Tiere in DG 1 (0-32 %) und 10 % in DG 2 (0-42 %) vor. Anzeichen von Lahmheit im Stand zeigten im DG 1 5 % (0-32 %) und im DG 2 3 % der Tiere (0-16 %). Im Mittel zeigten 30 % der Tiere im DG 1 (6-59 %) und 28 % im DG 2 (12-50 %) eine abweichende Fußstellung. Bei keinem der Merkmale lagen signifikante Unterschiede in der Auftretenshäufigkeit zwischen erstem und zweitem Durchgang vor ( $p=0,161-0,698$ ).

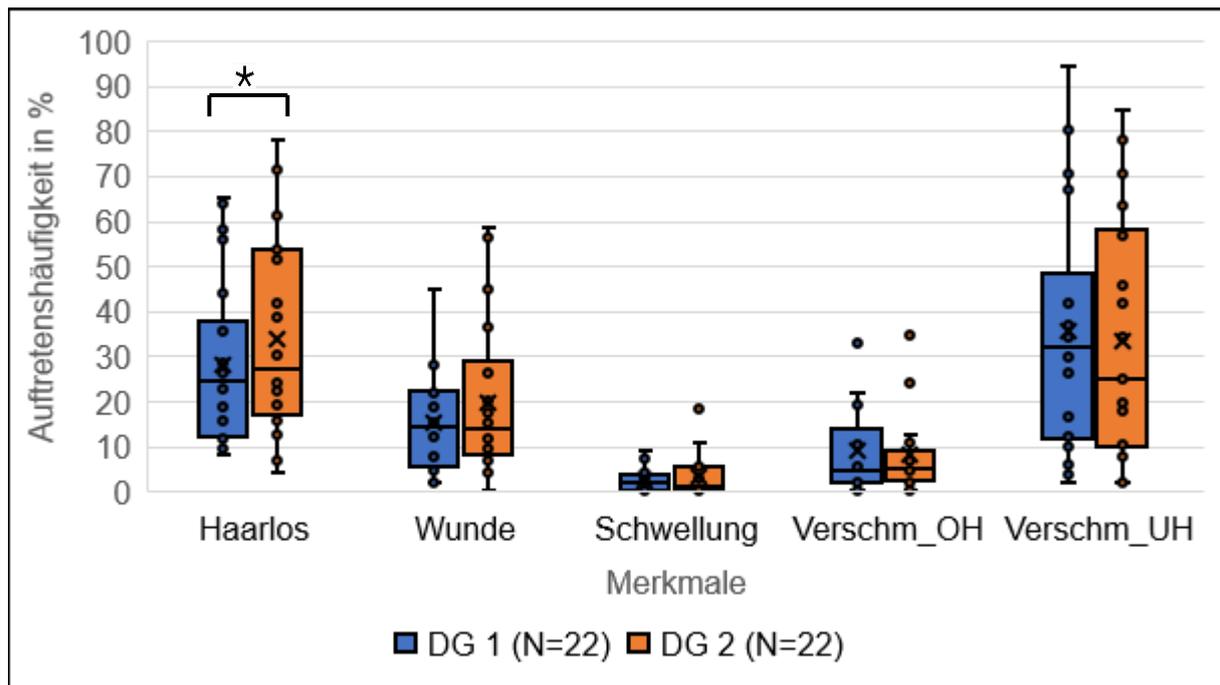


Abb. 19: Auftretenshäufigkeit in % für die Merkmale haarlose Stellen (Haarlos), Wunde, Schwellung, Verschmutzung im oberen hinteren Bereich (Verschm\_OH) und Verschmutzung im unteren hinteren Bereich (Verschm\_UH) auf Basis der Erhebung durch die externe Beurteilerin; x=Mittelwert

In Abb. 19 werden die Prävalenzen für die Merkmale haarlose Stellen (Haarlos), Wunde, Schwellung sowie Verschmutzung am oberen (Verschm\_OH) bzw. unteren Hinterbein (Verschm\_UH) für die beiden Durchgänge dargestellt. Die Prävalenzen des Merkmals BCS (Körperkondition) wurden aufgrund der vielen unterschiedlichen Kategorien/Scores nicht dargestellt.

Im Mittel zeigten 28 % (8-65 %) und 34% (4-78 %) der Tiere in DG 1 bzw. DG 2 haarlose Stellen. Im Mittel hatten 15 % (2-45 %) und 20% (0-58 %) im DG 1 bzw. DG 2 der Tiere Wunden. Das Merkmal Schwellung kam im Mittel in DG 1 (0-9 %) und 2 (0-18 %) 3 % vor. Im Mittel lag bei den Tieren 9 % (0-33 %) und 8 % (0-35 %) im DG 1 bzw. DG 2 Verschmutzungen im oberen hinteren Bereich vor. Das Merkmal Verschmutzung im unteren hinteren Bereich zeigten im Mittel in DG 1 bzw. DG 2 38 % (2-94 %) und in 33 % (2-84 %) der Tiere.

Mit Ausnahme des Merkmals Haarlos, das im zweiten Durchgang signifikant häufiger erfasst wurde als in DG 1 ( $p=0,037$ ), lagen keine weiteren signifikanten Unterschiede in der Auftretenshäufigkeit zwischen erstem und zweitem Durchgang vor ( $p=0,078-0,553$ ).

## 5.3 Übereinstimmung der Stakeholdergruppen mit der externen Beurteilerin

### 5.3.1 Landwirt\*innen

Tab. 10: Ergebnisse des Beobachter\*innenabgleichs zwischen Landwirt\*innen und der externen Beurteilerin (Masterstudentin) in DG 1 und DG 2; die Beobachter\*innenübereinstimmung wird als prozentuale Übereinstimmung (%ÜE), Prevalence Adjusted Bias Adjusted Kappa (PABAK) sowie Abweichung in der Herdenprävalenz (%-P. Abw.) ausgedrückt; grün (%ÜE 100-80 %/PABAK >0,6) = sehr gute bis gute Übereinstimmung, orange (%ÜE 79-70 %/PABAK (0,60-0,41) = Übereinstimmung noch akzeptabel, rot (%ÜE <70 %/PABAK 0,40-0,21) = Übereinstimmung unbefriedigend

Merkmale	LH_Gang	LH_Hochgradig	LH_Stand	Abw. Fußst.	Haarlos	Wunde	Schwellung	Verschm_OH	Verschm_UH
<b>DG 1 (N=18)</b>									
%ÜE mittel	75	90	90	78	78	87	97	87	66
%ÜE min	37	68	71	56	44	66	89	63	25
%ÜE max	100	100	100	94	98	100	100	100	94
PABAK mittel	0,63	0,88	0,79	0,55	0,56	0,73	0,93	0,74	0,32
PABAK min	0,05	0,37	0,43	0,12	-0,11	0,31	0,78	0,26	-0,5
PABAK max	1	1	1	0,89	0,95	1	1	0,96	0,89
%PABAK $\geq 0,4$	83	94	100	72	78	94	100	89	67
<b>DG 2 (N=19)</b>									
%ÜE mittel	77	92	93	72	77	84	96	87	67
%ÜE min	43	66	74	56	59	47	81	49	22
%ÜE max	100	100	100	92	96	96	100	98	97

<b>PABAK mittel</b>	0,66	0,84	0,86	0,45	0,53	0,67	0,92	0,74	0,34
<b>PABAK min</b>	0,15	0,32	0,49	0,12	0,17	-0,06	0,62	-0,03	-0,56
<b>PABAK max</b>	1	1	1	0,83	0,91	0,92	1	1	0,96
<b>%PABAK <math>\geq</math>0,4</b>	89	89	100	58	63	83	100	89	89

In beiden Durchgängen wurde bei den LW\*innen im Mittel eine gute bis sehr gute Übereinstimmung bei den Merkmalen LH\_Hochgradig (DG 1: 90%/68-100%; DG 2: 92%/66-100%), LH\_Stand (DG 1: 90 %/71-100 %; DG 2: 93 %/74-100 %), Wunde (DG 1: 87 %/66-100 %; DG 2: 84 %/47-96 %), Schwellung (DG 1: 97 %/89-100 %; DG 2: 96 %/81-100 %) und Verschm\_OH (DG 1: 87 %/63-100 %; DG 2: 87 %/49-98 %) erreicht (Tab. 10).

Bei den Merkmalen LH\_Gang, abw. Fußst. und Haarlos kam es mit 75 %/77 % (37-100 %/43-100 %), 78 %/72 % (56-94 %/56-92 %) bzw. 78 %/77 % (44-98 %/59-96 %) in beiden Durchgängen zu einer noch akzeptablen Übereinstimmung.

Dagegen war die Übereinstimmung mit der externen Beurteilerin bei den Merkmalen Verschm\_UH mit 66 %/67 % (25-94 %/22-97 %) und BCS mit 35 %/33 % (11-61 %/8-58 %) unbefriedigend.

Die Ergebnisse hinsichtlich Prozentübereinstimmung spiegeln sich im PABAK wider. Eine gute bis sehr gute Übereinstimmung (PABAK) hatten die Landwirt\*innen im Mittel bei den Merkmalen LH\_Hochgradig in DG 1 und 2, LH\_Stand in DG 2 und Schwellung in DG 1 und 2 (0,84-0,93), sowie bei den Merkmalen LH\_Gang in DG 1 und 2, LH\_Stand in DG 1, Wunde in DG 1 und 2 und Verschm\_OH in DG 1 und 2 (0,63-0,79). Eine noch akzeptable Übereinstimmung hatten die Landwirt\*innen im Mittel bei den Merkmalen abw. Fußst. in DG 1 und 2, Haarlos in DG 1 und 2 (0,45-0,56). Eine unbefriedigende Übereinstimmung hatten die Landwirt\*innen im Mittel bei dem Merkmal Verschm\_UH (0,32/0,34) sowie bei dem Merkmal BCS (0,18/0,16).

### 5.3.2 Kontrollassistent\*innen

Tab. 11: Ergebnisse des Beobachter\*innenabgleichs zwischen Kontrollassistent\*innen (N = 22) und der externen Beurteilerin (Masterstudentin) in DG 1 und 2; die Beobachter\*innenübereinstimmung wird als prozentuale Übereinstimmung (%ÜE), Prevalence Adjusted Bias Adjusted Kappa (PABAK) sowie Abweichung in der Herdenprävalenz (%-P. Abw.) ausgedrückt; grün (%ÜE 100-80 %/PABAK>0,6) = sehr gute bis gute Übereinstimmung, orange (%ÜE 79-70 %/PABAK (0,60-0,41) = Übereinstimmung noch akzeptabel, rot (%ÜE <70 %/PABAK 0,40-0,21) = Übereinstimmung unbefriedigend

Merkmale	LH_Stand	Abw. Fußst.	Haarlos	Wunde	Schwellung	Verschm_UH
<b>DG 1</b>						
%ÜE mittel	85	72	74	86	92	64
%ÜE min	62	41	47	61	67	43
%ÜE max	100	92	91	100	100	94
PABAK mittel	0,69	0,44	0,48	0,71	0,84	0,29
PABAK min	0,24	-0,19	-0,06	0,22	0,33	-0,14
PABAK max	1	0,84	0,91	1	1	0,89
%PABAK $\geq 0,4$	91	67	70	90	95	30
<b>DG 2</b>						
%ÜE mittel	86	71	74	81	94	62
%ÜE min	62	52	44	50	69	26
%ÜE max	100	89	100	100	100	96
PABAK mittel	0,71	0,41	0,47	0,62	0,87	0,24
PABAK min	0,23	0,04	-0,13	0	0,38	-0,49
PABAK max	1	0,77	1	1	1	0,92
%PABAK $\geq 0,4$	91	55	65	80	95	45

Es wurde bei den KA\*innen im Mittel eine gute bis sehr gute Übereinstimmung bei den Merkmalen LH\_Stand (DG 1/2: 85 %/86 %, jeweils 62-100 %), Wunde (86 %/81 %; 61-100 %/50-100 %) und Schwellungen (92 %/94 %; 67-100 %/69-100 %) erreicht. Bei den Merkmalen abw. Fußst. (72 %/71 %; 41-92 %/52-89 %) sowie Haarlos (74 %/74 %; jeweils 44-100 %) kam es zu einer noch akzeptablen Übereinstimmung. Bei dem Merkmal Verschm\_UH (64 %/62 %; 43-94 %/26-96 %) war die Übereinstimmung zur externen Beurteilerin unbefriedigend (Tab. 11).

Eine gute bis sehr gute Übereinstimmung (PABAK) hatten die Kontrollassistent\*innen im Mittel bei den Merkmalen Schwellung (0,84/0,87), LH\_Stand und Wunde (0,69/0,71). Eine noch akzeptable Übereinstimmung hatten die Kontrollassistent\*innen im Mittel bei den Merkmalen abw. Fußst. (0,44/0,41). Eine unbefriedigende Übereinstimmung hatten die Kontrollassistent\*innen in beiden Durchgängen im Mittel bei dem Merkmal Verschm\_UH (0,29/0,24).

### 5.3.3 Abweichung in %-Punkten von den Ergebnissen der externen Beurteilerin (Abw. in %-Punkte)

In Abb. 20 und 21 werden die mittleren Abweichungen bei den LW\*innen und KA\*innen von 0 in %-Punkten für die Merkmale Lahmheit gesamt/Gangbeurteilung (LH\_Gang), Lahmheit hochgradig/Gangbeurteilung (LH\_Hochgradig), Lahmheit/Beurteilung im Stand (LH\_Stand) sowie abweichende Fußstellung (Abw. Fußst.) von Durchgang 1 bzw. 2 dargestellt.

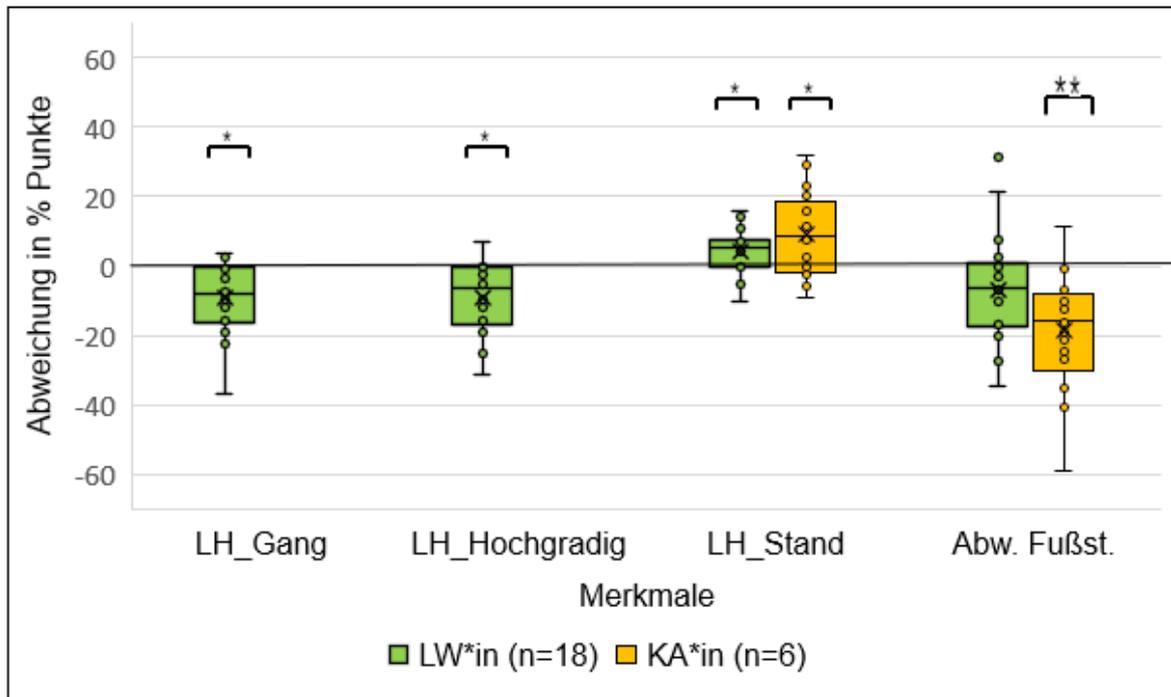


Abb. 20: Abweichung in %-Punkte der Landwirt\*innen/Kontrollassistent\*innen zur externen Beurteilerin für die Merkmale Lahmheit gesamt/Gangbeurteilung (LH\_Gang), Lahmheit hochgradig/Gangbeurteilung (LH\_Hochgradig), Lahmheit/Beurteilung im Stand (LH\_Stand) sowie abweichende Fußstellung (Abw. Fußst.) in Durchgang 1;  $\bar{x}$ =Mittelwert

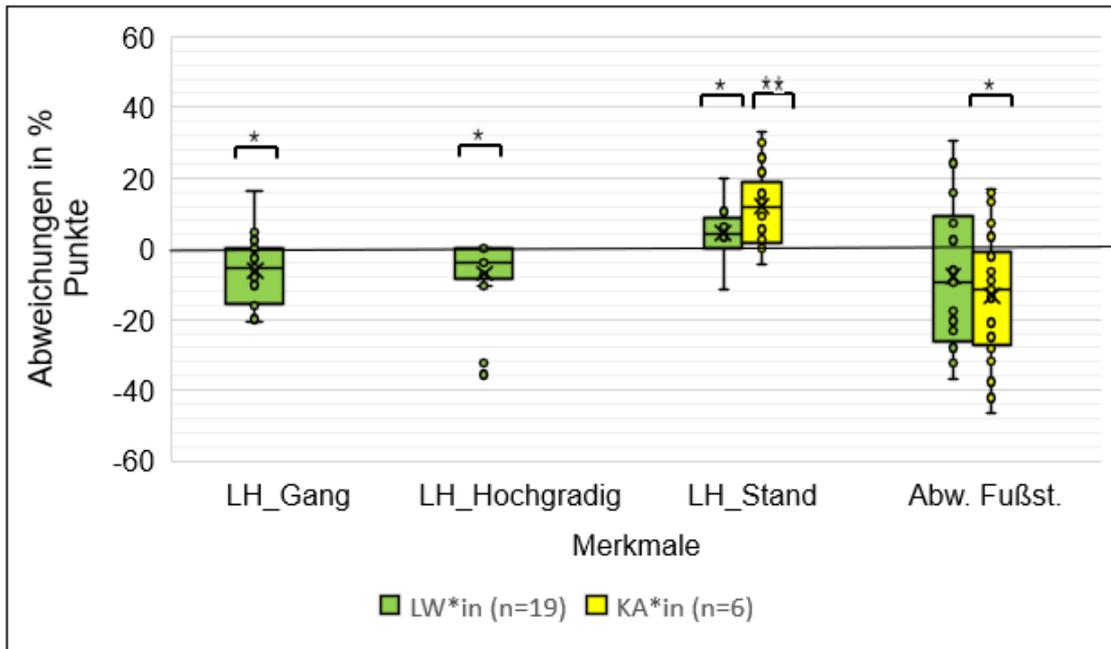


Abb. 21: Abweichung in %-Punkte der Landwirt\*innen/Kontrollassistent\*innen zur externen Beurteilerin für die Merkmale Lahmheit gesamt/Gangbeurteilung (LH\_Gang), Lahmheit hochgradig/Gangbeurteilung (LH\_Hochgradig), Lahmheit/Beurteilung im Stand (LH\_Stand) sowie abweichende Fußstellung (Abw. Fußst.) in Durchgang 2; x=Mittelwert

In beiden Durchgängen wurden bei den Merkmalen LH\_Gang (-9 %/-6 %,  $p=0,002/0,014$ ) sowie LH\_Hochgradig (-9 %/-7 %,  $p=0,002/0,008$ ) von den *LW\*innen* weniger Tiere als lahm bzw. hochgradig lahm als durch die externe Beurteilerin (EB) beurteilt. Demgegenüber wurden in beiden Durchgängen mehr Tiere mit Lahmheit im Stand beurteilt (+4 %/+4 %,  $p=0,018/0,014$ ). Numerisch erhoben die *LW\*innen* in beiden Durchgängen weniger Tiere mit einer abweichenden Fußstellung (-7 %/-8 %) im Vergleich zur externen Beurteilerin; hier lag aber kein signifikanter Unterschied vor (siehe Abbildung 20 und 21).

Im Vergleich zu EB wurde das Merkmal Lahmheit im Stand von den *KA\*innen* in beiden Durchgängen im Mittel häufiger erhoben (+9,3 %/+11,9 %,  $p=0,002/<0,001$ ), während das Merkmal abw. Fußst. seltener erhoben wurde (-18,6 %/-13,3 %,  $p<0,001/0,003$ ) (siehe Abbildung 20 und 21).

In Abb. 22 und 23 werden die mittleren Abweichungen bei den LW\*innen und KA\*innen von den durch die EB erhobenen Prävalenzen in %-Punkten für die Merkmale haarlose Stellen (Haarlos), Wunde, Schwellung, Verschmutzung im oberen hinteren Bereich (Verschm\_OH) und Verschmutzung im unteren hinteren Bereich (Verschm\_UH) vom ersten und zweiten Durchgang dargestellt.

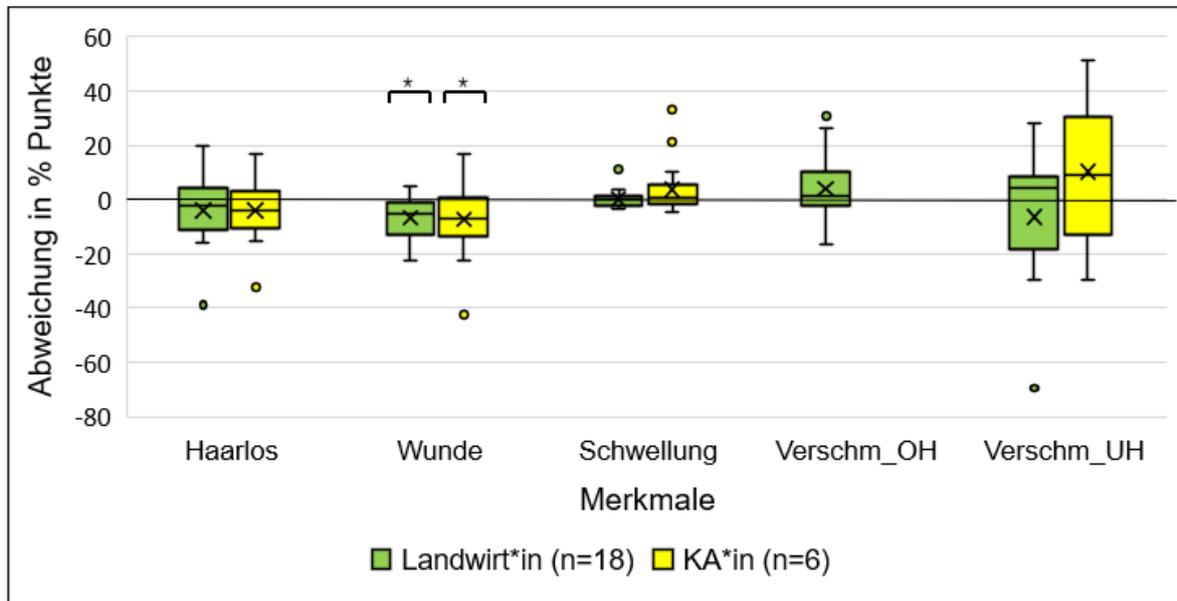


Abb. 22: Abweichung in %-Punkte der Landwirt\*innen/Kontrollassistent\*innen zur externen Beurteilerin für die Merkmale haarlose Stellen (Haarlos), Wunde, Schwellung, Verschmutzung im oberen hinteren Bereich (Verschm\_OH) und Verschmutzung im unteren hinteren Bereich (Verschm\_OH) und Verschmutzung im unteren hinteren Bereich (Verschm\_UH) in Durchgang 1; x=Mittelwert

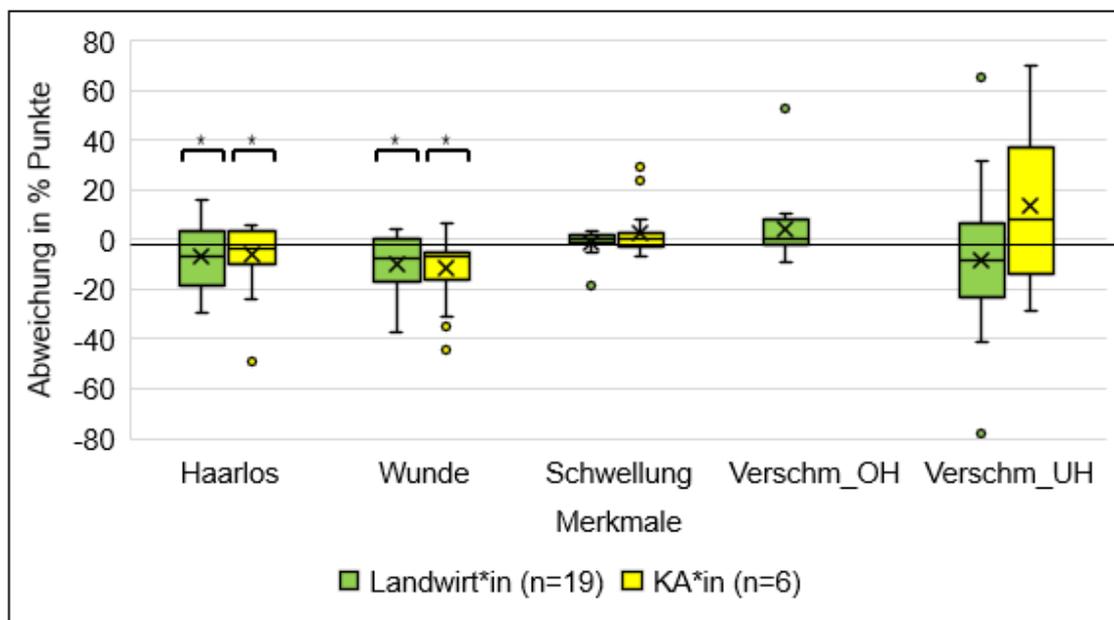


Abb. 23: Abweichung in %-Punkte der Landwirt\*innen/Kontrollassistent\*innen zur externen Beurteilerin für die Merkmale haarlose Stellen (Haarlos), Wunde, Schwellung, Verschmutzung im oberen hinteren Bereich (Verschm\_OH) und Verschmutzung im unteren hinteren Bereich (Verschm\_UH) in Durchgang 2; x=Mittelwert

Die Merkmale Verschmutzung im unteren hinteren Bereich (-6,5 %/-8,3 %) und haarlose Stellen (-3,8 %/-7 %) wurden von den *LW\*innen* in beiden Durchgängen im Mittel seltener als durch die externe Beurteilerin (EB) erhoben. In DG 2 wurde das Merkmal Haarlos von den *LW\*innen* unterschätzt ( $p=0,019$ ). Das Merkmal Wunde wurde im Mittel in beiden Durchgängen von den *LW\*innen* weniger häufig erhoben (-6,7 %/-10 %,  $p=0,003/0,004$ ). Die Abweichungen von den Ergebnissen der EB beim Merkmal Schwellung sind im Mittel bei den *LW\*innen* 0 (0,3 %/-1 %). Das Merkmal Verschmutzung im oberen hinteren Bereich wurde von den *LW\*innen* im Vergleich zur externen Beurteilerin im Mittel eher überschätzt (4,1 %/3,9 %), allerdings nicht signifikant; (siehe Abbildung 22 und 23).

Im Vergleich zu EB wurde das Merkmale Haarlos von den *KA\*innen* in beiden Durchgängen im Mittel seltener erhoben (-4,1 %/-5,9 %). Hier lag kein signifikanter Unterschied vor. Das Merkmal Wunde wurde von den *KA\*innen* in beiden Durchgängen im Mittel signifikant weniger häufig erhoben (-7,3 %/-11,7 %,  $p=0,018/p=0,001$ ). Die Merkmale Verschmutzung im unteren Bereich (10,2 %/13,6 %) und Schwellung (3,9 %/2,3 %) wurden von den *KA\*innen* in beiden Durchgängen im Mittel häufiger erhoben, allerdings nicht signifikant (siehe Abbildung 22 und 23).

Die mittleren Abweichungen in %-Punkten aller *KA\*innen* und *LW\*innen* von der externen Beurteilerin werden zusammenfassend in Tabelle 12 dargestellt.

Tab. 12: Zusammenfassung der mittleren Abweichung in %-Punkten der *LW\*innen* und *KA\*innen* zur externen Beurteilerin von den Merkmalen LH\_Gang, LH\_Hochgradig, LH\_Stand, abw. Fußst., Haarlos, Wunde, Schwellung, Verschm\_OH und Verschm\_UH; (+): positive Abweichung = höhere Prävalenz, (-): negative Abweichung = geringere Prävalenz, (\*\*): signifikante Abweichung ( $p<0,05/0,01$ )

	DG 1		DG 2	
	LW Mittelwert (%-P. Abw.)	KA Mittelwert (%-P. Abw.)	LW Mittelwert (%-P. Abw.)	KA Mittelwert (%-P. Abw.)
Alle <i>LW*innen</i> Alle <i>KA*innen</i>				
LH_Gang	-*		-*	
LH_Hochgradig	-*		-*	
LH_Stand	+*	+*	+*	+**
Abw. Fußst.	-	-**	-	-*
Haarlos	-	-	-*	-
Wunde	-*	-*	-*	-*
Schwellung	+	+	-	+
Verschm_OH	+		+	
Verschm_UH	-	+	-	+

### **5.3.3.1 Mittlere Abweichungen zur externen Beurteilerin in %-Punkten für individuelle KA\*innen**

In Tab. 13 wurden die mittleren Abweichungen zur externen Beurteilerin in %-Punkten von KA\*in 1, 2, 3 und 4 in DG 1 und 2 mit p-Werten für die Merkmale LH\_Stand, abw. Fußst., Haarlos, Wunde, Schwellung und Verschm\_UH dargestellt. Beim Merkmal LH\_Stand wurde in den meisten Fällen überschätzt (außer bei KA\*in 2 in DG 1). Die Abweichungen gehen bis zu 20 % und sind teilw. signifikant bis hochsignifikant (siehe p-Werte in Tab. 13). Die KA\*innen haben beim Merkmal abw. Fußst. mit bis zu -32 % größtenteils unterschätzt. Das Merkmal Haarlos wurde in den meisten Fällen unterschätzt. Bei dem Merkmal Schwellung gab es die geringsten Abweichungen. Das Merkmal Verschm\_UH wurde sowohl unterschätzt (bis zu -49 %, siehe p-Werte in Tab. 13) als auch überschätzt.

In Abb. 24 und 25 wurden die mittleren Abweichungen in %-Punkten zur externen Beurteilerin noch einmal graphisch verdeutlicht.

Tab. 13: Mittlere Abweichungen (%-Punkte) zur externen Beurteilerin für die Kontrollassistent\*innen 1, 2, 3 und 4 (KA\*in 1,2, 3, 4) in DG 1 und 2 für die Merkmale Lahmheitsbeurteilung im Stand (LH\_Stand), abweichende Fußstellung (abw. Fußst.), haarlose Stellen (Haarlos), Wunde, Schwellung und Verschmutzungen im unteren hinteren Bereich (Verschm\_UH)

DG 1												
	KA*in 1			KA*in 2			KA*in 3			KA*in 4		
	N	Mittelwert (%-P. Abw.)	p-Wert	N	Mittelwert (%-P. Abw.)	p-Wert <sup>1</sup>	N	Mittelwert (%-P. Abw.)	p-Wert <sup>1</sup>	N	Mittelwert (%-P. Abw.)	p-Wert
LH_Stand	8	12,7	0,021	3	7,4	-	3	-4,5	-	6	7,5	0,244
Abw. Fußst.	8	-34,0	<0,001	3	-12,6	-	3	-10,4	-	5	-5,7	0,274
Haarlos	7	-8,5	0,119	2	-3,0	-	3	-19,3	-	6	0,7	0,879
Wunde	7	-18,2	0,009	2	3,0	-	3	-5,1	-	6	1,2	0,783
Schwellung	7	-0,5	0,736	2	5,2	-	3	-8,3	-	6	11,4	0,087
Verschm_UH	7	28,7	0,008	2	8,1	-	3	-1,3	-	6	13,0	0,222
DG 2												
	KA*in 1			KA*in 2			KA*in 3			KA*in 4		
	N	Mittelwert (%-P. Abw.)	p-Wert	N	Mittelwert (%-P. Abw.)	p-Wert <sup>1</sup>	N	Mittelwert (%-P. Abw.)	p-Wert <sup>1</sup>	N	Mittelwert (%-P. Abw.)	p-Wert
LH_Stand	8	20,5	<0,001	3	13,7	-	3	0,2	-	6	8,2	0,094
Abw. Fußst.	8	-32,4	<0,001	3	5,8	-	3	-6,3	-	6	0,4	0,935
Haarlos	7	-5,7	0,206	2	-4,2	-	3	5,9	-	6	-12,9	0,202
Wunde	7	-16,3	0,012	2	-0,2	-	3	0,4	-	6	-10,7	0,181
Schwellung	7	-1,5	0,055	2	-0,9	-	3	-5,3	-	6	11,7	0,060
Verschm_UH	7	49,8	<0,001	2	-4,7	-	3	-1,9	-	6	-6,9	0,439

<sup>1</sup>: keine Prüfung auf Signifikanz aufgrund zu geringer Stichprobengröße

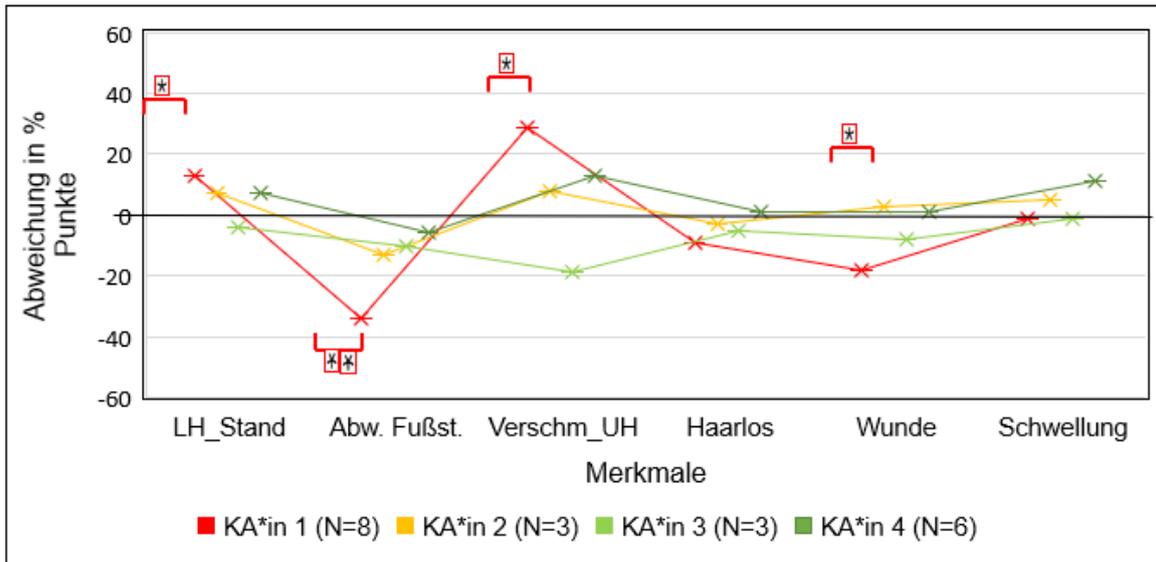


Abb. 24: Mittlere Abweichungen in %-Punkte von KA\*in 1, KA\*in 2, KA\*in 3 und KA\*in 4 zur externen Beurteilerin der Merkmale LH\_Stand, abw. Fußst., Verschm\_UH, Haarlos, Wunde und Schwellung in Durchgang 1

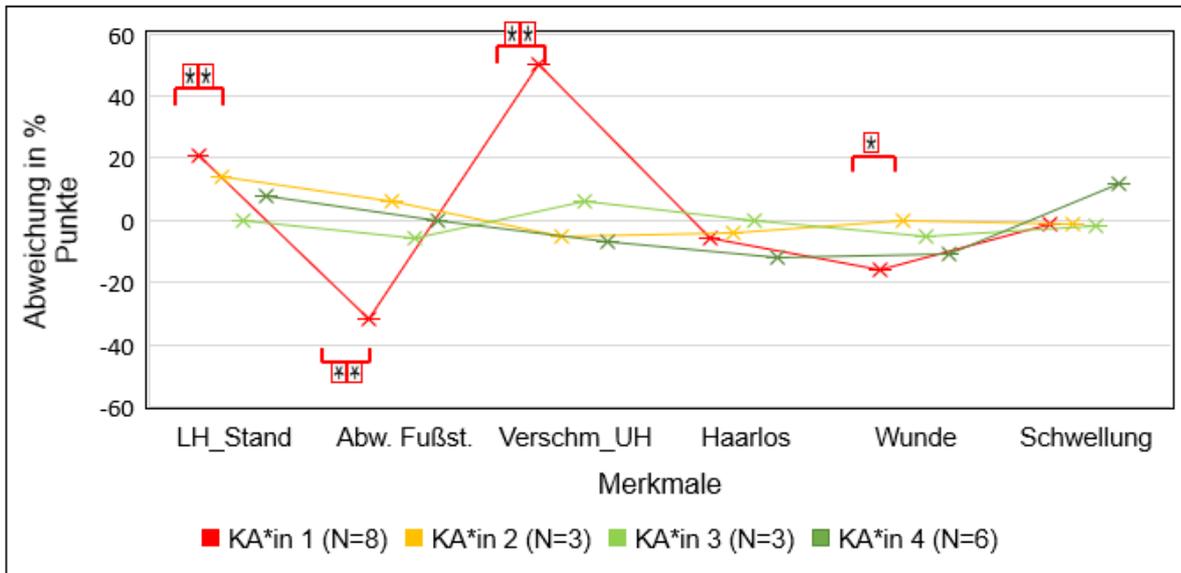


Abb. 25: Mittlere Abweichungen in %-Punkte von KA\*in 1, KA\*in 2, KA\*in 3 und KA\*in 4 zur externen Beurteilerin der Merkmale LH\_Stand, abw. Fußst., Verschm\_UH, Haarlos, Wunde und Schwellung in Durchgang 2

## 5.4 Erfassung ‚hochgradig lahmer‘ Tiere: Beurteilung im Stand (KA\*in bzw. LW\*in) vs. Gangbeurteilung (externe Beurteilerin)

Tab. 14: Spezifität, Sensitivität sowie positive predictive value (PPV) und negative predictive value (NPV) für die Erfassung hochgradiger Lahmheit mittels Standbeurteilung (LH\_Stand) durch KA\*in/LW\*in gegenüber der Gangbeurteilung (LH\_Hochgradig) durch die externe Beurteilerin

	Spezifität = TN/(TN+FP)			Sensitivität = TP/(TP+FN)			PPV = TP/(TP+FP)			NPV = TN/(TN+FN)		
	Mittel	Min	Max	Mittel	Min	Max	Mittel	Min	Max	Mittel	Min	Max
<b>DG 1</b>												
N	LW=17/KA=22			LW=14/KA=17			LW/KA=16			LW=18/KA=22		
LW*in	<b>93%</b>	74%	100 %	<b>39%</b>	0	67%	<b>39%</b>	0%	100 %	<b>92%</b>	69%	100 %
KA*in	<b>89%</b>	61%	100 %	<b>30%</b>	0	100 %	<b>27%</b>	0%	80%	<b>91%</b>	74%	100 %
<b>DG 2</b>												
N	LW18/KA22			LW13/KA16			LW13/KA19			LW18/KA22		
LW*in	<b>96%</b>	84%	100 %	<b>33%</b>	0	75%	<b>46%</b>	0%	100 %	<b>94%</b>	70%	100 %
KA*in	<b>88%</b>	59%	100 %	<b>45%</b>	0	100 %	<b>31%</b>	0%	100 %	<b>94%</b>	70%	100 %

Die Beurteilung durch die LW\*innen wies im DG 1 im Mittel eine Spezifität von 93 % (74-100 %) und eine Sensitivität von 39 % (0-67 %) auf (siehe Tab. 14). Das heißt, 93 % der Tiere, die auch laut EB als nicht hochgradig lahm eingestuft wurden, wurden von den LW\*innen als nicht lahm im Stehen eingestuft (Spezifität). Jedoch wurden von den LW\*innen im Mittel nur 39 % der Tiere als lahm im Stand erhoben, die von der externen Beurteilerin als hochgradig lahm eingestuft wurden (Sensitivität). In DG 2 ergab sich für die LW\*innen im Mittel eine Spezifität von 96 % (84-100 %) und eine Sensitivität von 33 % (0-75 %).

Im Mittel lag der positive predictive value (PPV) bei den LW\*innen bei 39 % in DG 1 und 46 % in DG 2 (0-100 %) (siehe Tab. 14). Die Wahrscheinlichkeit, dass LW\*innen hochgradig lahme Tiere (erhoben durch externe Beurteilerin) auch als lahm im Stand identifizierten, lag bei 39 bzw. 46 %. Der negative predictive value (NPV) lag im Mittel bei 92 % in DG 1 und 94 % in DG 2 (69-100 %). Damit lag die Wahrscheinlichkeit, dass LW\*innen Tiere als nicht lahm im Stand identifizierten, welche auch durch die externe Beurteilerin als nicht hochgradig lahm bei der Gangbeurteilung erhoben wurde, bei 92 bzw. 94 %.

Für die Beurteilung durch die KA\*innen ergaben sich im DG 1 im Mittel eine Spezifität von 89 % (61-100 %) und eine Sensitivität von 30 % (0-100 %). Das heißt 89 % der Tiere, die auch laut EB als nicht hochgradig lahm eingestuft wurden, wurden von den KA\*innen als nicht lahm im Stehen eingestuft (Spezifität). Jedoch wurden von den KA\*innen im Mittel nur 30 % der Tiere als lahm im Stand erhoben, die von der

externen Beurteilerin als hochgradig lahm eingestuft wurden (Sensitivität). Im DG 2 lag für die KA\*innen im Mittel eine Spezifität von 88 % (59-100 %) und eine Sensitivität von 45 % (0-100 %) vor.

Im Mittel lag der positive predictive value (PPV) bei den KA\*innen bei 27 % in DG 1 und 31 % in DG 2 (0-100 %). Die Wahrscheinlichkeit, dass KA\*innen hochgradig lahme Tiere (erhoben durch externe Beurteilerin) auch als lahm im Stand identifizierten, lag bei 27 % bzw. 31 %. Der negative predictive value (NPV) lag im Mittel bei 91 % in DG 1 und 94 % in DG 2 (70-100 %). Die Wahrscheinlichkeit, dass KA\*innen Tiere als nicht lahm im Stand identifizierten, welche auch durch die externe Beurteilerin als nicht hochgradig lahm bei der Gangbeurteilung erhoben wurde, lag bei 91 bzw. 94 %.

In der Abb. 26 wird ersichtlich, dass die Spezifität (nicht hochgradig lahme Tiere wurden als nicht lahm im Stand identifiziert) sowohl bei LW\*in wie auch bei KA\*in weniger variierte als die Sensitivität (hochgradig lahme Tiere wurden auch als lahm im Stand identifiziert). Im 2. Durchgang ergab sich ein ähnliches Bild. Eine Berechnung der Sensitivität konnte nicht bei allen Betrieben durchgeführt werden, da bei einigen Betrieben keine hochgradig lahmen Tiere vorkamen.

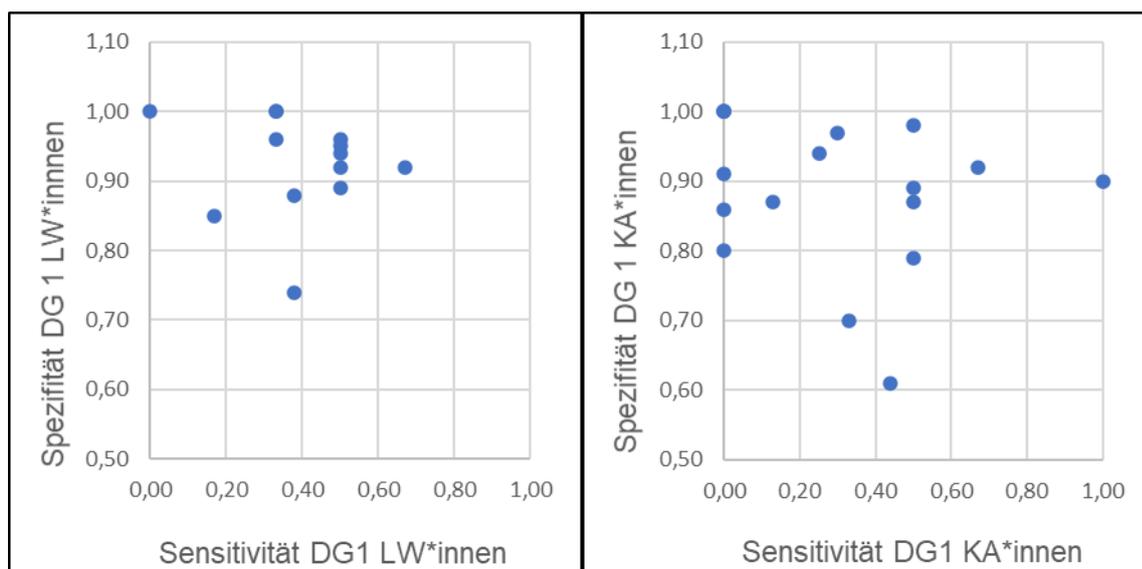


Abb. 26: Gegenüberstellung Sensitivität und Spezifität der Übereinstimmung der Beurteilung hochgradiger Lahmheit zwischen Gangbeurteilung (LH\_Hochgradig) externer Beurteilerin und Standbeurteilung (LH\_Stand) durch LW\*in (N=18)/KA\*in (N=6), die Übereinstimmung wird als Spezifität und Sensitivität abgebildet

## **5.5 Aussagen der Teilnehmer\*innen**

### **5.5.1 Aussagen Landwirt\*innen und Kontrollassistent\*innen beim leitfadengestützten Interview**

Generell war die Einstellung der Landwirt\*innen gegenüber dem Projekt sehr positiv. Die befragten Landwirt\*innen gaben in den meisten Fällen als erstes an, dass die Erhebung der Körperkondition (BCS) als sehr schwierig bzw. aufwendig empfunden wurde und diese unbedingt Übung voraussetzt. Beim zweiten Durchgang gaben die Landwirte\*innen an, dass der BCS leichter zu ermitteln gewesen ist. Insgesamt wurde festgestellt, dass durch die BCS-Beurteilung einzelne Körperregionen genauer beobachtet werden und eine genauere Information über die Kondition der Rinder erhalten wird.

Es wurde von einigen Landwirt\*innen erwähnt, dass die dreistufige Einteilung des Merkmals Lahmheit (Gangbeurteilung) eine gute Erhebung von Lahmheit ermöglicht. Anderen erschien die 3-stufige Lahmheitserkennung als zu wenig differenziert, um Lahmheit frühzeitig zu diagnostizieren, und ein 5-stufiges Lahmheitsscoring wurde bevorzugt, da die zweite Stufe in einem solchen Schema bereits wichtig erscheint, z. B. im Hinblick auf notwendige Klauenpflege. Ein\*e andere\*r Landwirt\*in traf die Aussage, dass er\*sie die Lahmheitsbeurteilung vom Melkstand aus durchführt, und diese Methode zur Bewertung der Lahmheit ausreichend ist. Des Weiteren wurde erwähnt, dass der\*die Landwirt\*in bei der Erhebung voreingenommen wäre, da er\*sie seine\*ihre Tiere kennt und daher eine objektive Beurteilung nur schwer durchgeführt werden könne.

Bei den berücksichtigten Merkmalen, wie Lahmheit Stand- und Gangbeurteilung, abweichende Fußstellung, Integumentschäden, Verschmutzung und Körperkondition, waren sich die Landwirt\*innen einig, dass es einen Zusammenhang zw. den Merkmalen und Klauengesundheit gibt. Durch das Erheben der Merkmale erfolgte eine bewusste Beobachtung der Tiere sowie eine Fokussierung auf Einzeltiere. Aus Sicht der Landwirt\*innen können durch die Merkmale Mängel im Stall aufgedeckt und einer Betriebsblindheit entgegengewirkt werden.

Die Kontrollassistent\*innen erwähnten die Wichtigkeit von Erfahrung und Routine, um eine zuverlässige Merkmalerhebung zu gewährleisten. Die Erhebung an sich wurde nicht als aufwendig beschrieben; eine Beurteilung im Stall könne jedoch vermutlich genauer als im Melkstand durchgeführt werden. Die KA\*innen waren der Meinung, dass die Erhebung von Lahmheit im (Melk-)Stand während der Milchmessung machbar ist, andere Merkmalerhebungen, wie z. B. abweichende Fußstellung, Integumentschäden und Verschmutzung am unteren Hinterbein aber zu aufwendig sind und ihre Arbeit behindern würden. Weiters wurde zum Merkmal Lahmheit im (Melk-)Stand von den Kontrollassistent\*in angeführt, dass die Erhebung teilweise aufgrund des Verhaltens der Tiere während des Melkens verzerrt sein könnte. Ein Herumtrippeln bzw. wiederholte Entlastung der Beine könnte nicht nur auf Lahmheit im Stand sondern auch auf Nervosität des Tieres im Melkstand (z. B. veränderte Situation im Melkstand durch Anwesenheit des\*der Kontrollassistent\*in etc.) zurückzuführen sein.

Zum Merkmal abweichende Fußstellung wurde von den Kontrollassistent\*innen angemerkt, dass eine zuverlässige Erhebung vom Melkstandtyp abhängig ist. Beim Fischgrätenmelkstand stehen die Rinder in einem Winkel von rund 30° von der

Melkgrube weg. Aufgrund der ‚schiefen‘ Stellung des Tieres im Melkstand ist es schwieriger, eine Position direkt hinter dem Tier einzunehmen und damit eine klare Beurteilung zu erhalten. Dagegen wurde angegeben, dass im Side-by-Side-Melkstand, wo die Tiere in einem Winkel von 90° zur Melkgrube nebeneinander stehen, die abweichende Fußstellung gut erhoben werden konnte. Bei diesem Melkstandtyp war auch die Erhebung des Merkmals Lahmheit im Stand am einfachsten möglich.

Beim Merkmal Verschmutzung am unteren Hinterbein wurde von den Kontrollassistent\*innen diskutiert, dass vermehrter Kotabsatz aufgrund von Stresssituationen im Melkstand zu vermehrter Verschmutzung führen kann.

Für die Weiterarbeit mit den Merkmalen wäre es für zwei der KA\*innen interessant, ob und welche Korrelationen es zu den Klauenpflagedaten gibt.

### **5.5.2 Aussagen bei den Abschlussveranstaltungen zur den Piloterhebungen im Projekt Klauen-Q-Wohl**

Die Aussagen der Teilnehmer\*innen können wie folgt zusammengefasst werden:

Die Lahmheitserhebung durch die KA\*innen im Melkstand (Lahmheit im Stand) wurde von mehreren Landwirt\*innen als positiv gesehen; allerdings wurde auch darauf hingewiesen, dass die Erhebung in manchen Melksystemen erschwert ist. Lahmheit im Stand könnte als Hilfsmerkmal für die Erkennung von Lahmheit herangezogen werden. Ein\*e Landwirt\*in traf die Aussage, dass dieses Hilfsmerkmal nicht ausreichend ist, da die LH\_Stand-Erhebung eher eine Momentaufnahme sei. Hier sollte ev. moderne Techniken verwendet werden.

Die Landwirt\*innen befürworteten die Erhebung durch den\*die Kontrollassistent\*in bei der Milchleistungskontrolle. Die Ergebnisse sollten nachher am Tagesbericht angeführt werden. Es wurde angenommen, dass dies zu einem verbesserten Problembewusstsein bei den Landwirt\*innen führt. Derartige Projekte schafften Bewusstsein und man schaue wieder einmal genauer hin. Laut Selbsteinschätzung der Landwirt\*innen wird man mit der Zeit betriebsblind und neigt sehr wahrscheinlich dazu gewisse Probleme zu unterschätzen. Ein Input von außen (z. B. durch externe Personen od. KA\*innen) könnte hier Abhilfe schaffen.

Tierwohlerhebungen durch eine externe Person sollten aus Sicht einiger Landwirt\*innen nur auf freiwilliger Basis durchgeführt werden. Dieser Punkt wurde jedoch unterschiedlich gesehen, da unter diesen Bedingungen „schwarze Schafe“ unter den Landwirt\*innen nicht erkannt werden. Ein\*e Landwirt\*in schlug vor, dass TGD-Tierärzt\*innen bzgl. Tierwohlerhebungen in die Pflicht genommen werden sollten. Sie sollten Tiergesundheitserhebungen gewissenhaft durchführen und Tierwohlparameter aufnehmen. Der\*die Landwirt\*in schlug eine Kooperation mit dem TGD, der/dem Betreuungstierärzt\*in und dem\*der LW\*in (z. B. für die Bestandsbetreuung, regelmäßige Betriebserhebungen, Optimierung der Tiergesundheit sowie Tiergesundheitsprogramme) und eine elektronische Datenerfassung (nicht auf Papierzetteln, die verschwinden können) vor.

## 6 Diskussion

Rinderherden in Laufstallhaltung sind oft von Klauen- und Gliedmaßenkrankungen betroffen. Die frühzeitige Erkennung von gesundheitlichen Störungen wie z. B. Klauenerkrankungen ist ein wichtiger Faktor zur Vermeidung von Lahmheiten, welche mit hohen Tierarzt- und Behandlungskosten verbunden sind (Lorenzini et al., 2018). Um eine wirksame Verbesserung durch Managementmaßnahmen zu erzielen, sollen zunächst die gesteckten Ziele definiert werden und eine Zusammenarbeit von Landwirt\*in, Berater\*in, Klauenpfleger\*in sowie Tierärzt\*in angestrebt werden (Kümper, 2008).

Um die Häufigkeit von Klauenerkrankungen und die damit verbundenen Abgänge zu reduzieren, hat die Zucht auf gesunde Klauen das Potential, langfristig und nachhaltig eine Trendwende zu erreichen und in weiterer Folge die Klauengesundheit züchterisch zu verbessern (Egger-Danner, 2015). Dies setzt jedoch eine zuverlässige Merkmalerfassung voraus.

Die Erhebung des Gangs bei Rindern kann anhand subjektiver Methoden, wie dem Locomotion-Scoring-System nach Sprecher et al. (1997) durchgeführt werden. So wurden z. B. Kraftmessplatten zur Gangbeurteilung verwendet (Scott, 1988; Van der Tol et al., 2002). Rajkondawar et al. (2006) empfahlen Kraftmessplatten als Instrument zur Erkennung von Lahmheit. Jedoch kam es zu Unklarheiten bei den Messungen, wenn mehr als eine Gliedmaße von Lahmheit bzw. Verletzung betroffen war. Andere Ansätze zur automatisierten Lahmheitserkennung nutzen Pedometer oder Pansensoren (Lemmens, s.a.). Es gibt derzeit jedoch noch kein zuverlässiges System zur automatisierten Lahmheitserkennung, das Einzug in die breite landwirtschaftliche Praxis gefunden hat. Daher ist die Gangbeurteilung durch direkte Beobachtung derzeit nicht ersetzbar (Lemmens, s.a.).

Fragen der Reliabilität (Zuverlässigkeit von Erhebungen) wurden in der Vergangenheit oft vernachlässigt und bedürfen in Zukunft einer gründlicheren Untersuchung und Diskussion, insbesondere im Hinblick auf geeignete Teststatistiken und Grenzen der Akzeptanz. Mittel zur Verbesserung der Zuverlässigkeit sind die Verfeinerung von Definitionen oder Aufzeichnungsmethoden und Training (Knierim und Winckler, 2009).

### 6.1 *Schulung und Beobachter\*innentraining*

Die Schulung, welche eine Voraussetzung für die Teilnahme an der vorliegenden Studie war, beinhaltete eine theoretische Einführung in die Erhebung von Lahmheit und weiteren tierwohlbezogenen Indikatoren des Wohlergehens, eine praktische Anleitung für die on-farm-Erhebung mittels der Schulungsunterlagen (siehe Anhang) und einen abschließenden Beobachter\*innenabgleich anhand von Videos und Fotos. Insgesamt wurde eine gute bis akzeptable Übereinstimmung erzielt. Bei einigen Teilnehmer\*innen gab es Probleme beim Erkennen von bestimmten Merkmalen, z. B. abweichende Fußstellung (im Mittel 79 %ÜE; 43-100 %ÜE). Daher wurden die Ergebnisse des Beobachter\*innenabgleichs den Teilnehmer\*innen mit einer Erklärung zur Interpretation (prozentuale Übereinstimmung und Abweichung in %

zum SS) zugesandt. Bei einer Beobachter\*innenübereinstimmung <80% wurden die Beobachter\*innen aufgefordert, die Kriterien für das betreffende Merkmal noch einmal anzuschauen. Eine weitere direkte Diskussion am Tier bzw. Beurteilung von Videos und Fotos wäre wünschenswert gewesen – dies hätte aber den zeitlichen Rahmen der Schulung überschritten.

## **6.2 Auftretenshäufigkeiten (Prävalenzen) im ersten und im zweiten Durchgang (Erhebung durch externe Beurteilerin)**

Die mittlere Auftretenshäufigkeit der Merkmale Lahmheit im Gehen bzw. hochgradige Lahmheit ist vergleichbar mit den Ergebnissen von Van Os et al. (2019), Sarova et al. (2011) und Dippel et al. (2009) bzw. Van Os et al. (2019), Sarova et al. (2011) und Cook et al. (2016). Weiters ermittelten die Autor\*innen eine ähnliche Variabilität für die Merkmale Lahmheit im Gehen bzw. hochgradige Lahmheit.

Auffällig ist, dass anhand Gangbeurteilung im Mittel 28 % (DG 1) bzw. 25 % (DG 2) der Tiere Lahmheit zeigten, davon im Mittel 11 % (DG 1) bzw. 10 % (DG 2) hochgradig lahm gingen, aber nur 5 % (DG 1) bzw. 3 % (DG 2) im Mittel der Tiere Lahmheit im Stehen zeigten. Wenn man davon ausgeht, dass hochgradig lahme Tiere auch Lahmheit im Stand zeigen sollten, dann werden viele Tiere bei der Standbeurteilung als nicht lahm identifiziert. Demnach werden viele hochgradig lahme Tiere im Stand nicht erfasst (vgl. Kapitel 6.5).

Die Prävalenzen der Merkmale Verletzungen bzw. Wunden/Krusten, Schwellungen sind ebenfalls mit den von Van Os et al. (2019) angegebenen Ergebnissen vergleichbar. Auch bei Piloterhebungen im Projekt EIP lagen bezüglich der Merkmale Verschmutzung am oberen bzw. am unteren Hinterbein ähnliche Auftretenshäufigkeiten vor (persönliche Mitteilung C. Winckler). In der Studie von de Boyer des Roches et al. (2014) zeigten eine deutliche Mehrzahl der Rinder Verschmutzungen am oberen bzw. am unteren Hinterbein. Die Ergebnisse von de Boyer des Roches et al. (2014) sind nicht direkt mit den Ergebnissen der Masterarbeit vergleichbar, da die Autor\*innen die Kriterien von Welfare Quality® (Kriterium für Verschmutzung: handtellergröße Kotauflage am oberen Hinterbein vs. 30cm durchgehende Kotauflagerung in der vorliegenden Studie) verwendeten. In der Literatur wurden neben Winckler und Kofler (2018) keine weiteren Studien gefunden, welche die hier verwendete Definition des Merkmals Verschmutzung am oberen Hinterbein verwendeten.

Beim Merkmal haarlose Stellen am Sprunggelenk ist in der Studie von Why et al. (2003a) eine ähnliche Prävalenz zu den Ergebnissen der Masterarbeit vorhanden.

Das Merkmal abweichende Fußstellung wird in der Literatur kaum beschrieben. Metzner (2009) nennt die kuhhessige Stellung als das Nachaußenrichten der Klauen. Dies wird häufig bei schmerzhaften Prozessen (z. B. Klauensohlengeschwür) unter den Außenklauen der Hinterbeine beobachtet. Mit etwa 30 % lag die Prävalenz in der vorliegenden Studie deutlich über den Angaben von Distl (1995) für kuhhessige Gliedmaßenstellung bei erstlaktierenden Rindern von 8,4 %.

### **6.3 Übereinstimmung externe Beurteilerin mit erfahrenem Beurteiler**

Damit die externe Beurteilerin als Silberstand für die weiteren Erhebungen eingesetzt werden konnte, wurde die Beobachter\*innenübereinstimmung mit einem Experten überprüft. Die Übereinstimmung war dabei unter Berücksichtigung einer prozentualen Übereinstimmung von mind. 80 % (Burn et al. 2009; AssureWel, 2016, zit. n. Schenkenfelder, 2016) und eines PABAK von  $>0,6$  mit Ausnahme der Beurteilung der Körperkondition (Landis und Koch, 1977) gut bis sehr gut.

Sowohl im DG 1 (47 %ÜE, 0,33 PABAK) als auch im DG 2 (56 %ÜE, 0,45 PABAK) zeigte die Beurteilung der Körperkondition die schlechteste Übereinstimmung. Diese unbefriedigende Übereinstimmung lag sowohl bei der externen Beurteilerin als auch bei den LW\*innen vor (siehe Kapitel 6.4). Hier muss erwähnt werden, dass nicht gerundet wurde und keine Abstufungen zugelassen wurden (siehe Kapitel 5.1). Ferguson et al. (1994) erreichte bei der Beurteilung der Körperkondition mit einer 0,25-Punkte-Skala eine Beobachter\*innenüberstimmung von 58 – 67% unter Anwendung des Modalwerts. Beim Zulassen einer Abweichung von 0,5 bzw. 1 Punkte stieg die prozentuale Übereinstimmung zwischen den Beobachter\*innen. Aufgrund der unzureichenden Übereinstimmung wurde das Merkmal Körperkondition in dieser Studie nicht weiter berücksichtigt.

Beim Merkmal Lahmheit lag im DG 2 eine niedrigere Übereinstimmung vor (%ÜE/PABAK DG 1 vs. DG 2: LH\_Gang 81 %/0,71 vs. 72 %/0,58; LH\_Stand 94 %/0,88 vs. 75 %/0,50). In DG 2 erfasste die externe Beurteilerin weniger lahme Tiere als der erfahrene Beurteiler. Auffällig war im DG 2 die insgesamt hohe Anzahl an lahmen bzw. hochgradig lahmen Tieren (75 % lahm, davon 42 % hochgradig lahm). Mullan et al. (2011) wies auf die Tendenz von Beobachter\*innen hin, im Vergleich zum\*r Trainer\*in (in diesem Fall der erfahrene Beurteiler) eine niedrigere Auftretenshäufigkeit eines Merkmals zu erfassen. Trainer\*innen sind häufig sensibler beim Erfassen der Merkmale – gerade bei ‚Grenzfällen‘. In diesem Fall könnte im DG 2 eine Beobachter\*innenverzerrung (Observer bias) dahingehend vorgelegen haben, dass die hohe Prävalenz im Betrieb dazu geführt hat, dass die externe Beurteilerin weniger konsequent die Merkmalsdefinitionen angewendet hat. Durch weiteres Training sowie ein klar definiertes Testprotokoll kann einer Beobachter\*innenverzerrung entgegengewirkt werden (Mullan et al., 2011).

### **6.4 Übereinstimmung der Stakeholdergruppen mit der externen Beurteilerin**

Die mittlere Übereinstimmung mit der externen Beurteilerin erreichte bei den Kontrollassistent\*innen einen PABAK von 0,24 bis 0,87 und bei den Landwirt\*innen von 0,32 bis 0,93. Es lag eine gute bis sehr gute Übereinstimmung ( $>0,6$ ) für Lahmheit, Verletzungen und Verschmutzung der oberen Hinterbeine, eine noch akzeptable Übereinstimmung (0,4-0,6) für abweichende Fußstellung und haarlose Stellen und eine unbefriedigende Übereinstimmung ( $<0,4$ ) für Verschmutzung am unteren Hinterbein und BCS vor. Bis auf das Merkmal LH\_Stand wurde von den LW\*innen als auch von den KA\*innen im Mittel immer unterschätzt – d.h. es wurden

für die Merkmale LH\_Gang, LH\_Hochgradig, Abw. Fußst., Haarlos, Wunde, Verschm\_UH geringere Prävalenzen erhoben als durch die externe Beurteilerin.

#### **6.4.1 Landwirt\*innen**

Auffällig ist, dass bei Berücksichtigung der Prozent-Übereinstimmung eine höhere Beobachter\*innenübereinstimmung bei den LW\*innen vorlag als bei den KA\*innen. Der Anteil an Landwirt\*innen mit zumindest noch akzeptabler Übereinstimmung ( $PABAK > 0,4$ ) schwankte zwischen 58 % (abweichende Fußstellung, DG 2) und 100 %. Als unproblematisch stellten sich die Beurteilung von LH-Stand und Schwellungen, sowie mit Einschränkungen LH\_Hochgradig ( $PABAK 0,32-1$ ) heraus. Bei allen anderen Merkmalen unterschritten zumindest einige Teilnehmer\*innen den angestrebten Grenzwert. Beim Merkmal Verschmutzung im unteren hinteren Bereich war die größte Abweichung in der Beurteilung zu finden (minimaler  $PABAK -0,5$ ). Im Mittel unterschätzten hier die Landwirt\*innen – das heißt, es wurden weniger Tiere als verschmutzt im unteren hinteren Bereich erhoben im Vergleich zur externen Beurteilerin. Die Merkmale LH\_Gang sowie LH\_Hochgradig wurden im Mittel signifikant weniger häufig (Unterschätzung) und LH\_Stand signifikant häufiger erhoben als durch die externe Beurteilerin (Überschätzung). Dieser Befund bestätigt die Untersuchung von Sarova et al. (2011), bei der geschulte Landwirt\*innen Lahmheit im Mittel um 6 % unterschätzten (Variabilität von 0-20 %). Diese Abweichungen waren in den Studien von Whay et al. (2003a) und Wells et al. (1993) sogar noch größer; die Landwirt\*innen identifizierten weniger als die Hälfte (25-50 % bzw. 40-45 %) der Milchkühe als lahm.

Zum einen hängt die Übereinstimmung von Übung, Erfahrung, Training, verfügbarer Zeit für das zu erhebende Merkmal und Aufmerksamkeitsniveau des\*der Beobachter\*in ab. Zum anderen ist die Klarheit der Definition von Merkmalen ein wichtiger Faktor (Martin und Bateson, 1986). Bei der Erhebung des Merkmals Verschmutzung am unteren hinteren Bereich (das untere Hinterbein inkl. Innenseite des gegenüberliegenden Hinterbeins weist in Summe eine mind. handtellergroße Kotauflagen auf), könnte es sein, dass Landwirt\*innen dieses Merkmal falsch interpretierten und nach Definition verschmutzte Tiere als sauber angesehen haben, weil sie dieses Ausmaß an Verschmutzung vermutlich als nicht relevant erachteten. Beim Merkmal Schwellungen lag eine sehr gute Übereinstimmung vor. Hier muss aber erwähnt werden, dass Schwellungen auf den Betrieben eher weniger vorkamen - d.h. es gibt eine gute Beobachter\*innenübereinstimmung bei den ‚Nicht-Schwellungen‘. Main et al. (2000) wies darauf hin, dass bei geringem Auftreten eines Merkmals die Interpretation der Prävalenzen erschwert, die Berechnung des Kappa-Werts nicht möglich (bei Nicht-Auftreten) oder das Ergebnis weniger relevant ist.

#### **6.4.2 Kontrollassistent\*innen**

Der Anteil an Kontrollassistent\*innen mit zumindest noch akzeptabler Übereinstimmung ( $>0,4$ ) schwankte zwischen 30 % (Verschm\_UH; DG 1) und 100 %.

Mit Einschränkung der Beurteilung von Schwellungen ( $PABAK 0,33-1$ ) unterschritten zumindest einige Kontrollassistent\*innen bei allen anderen Merkmalen den angestrebten Grenzwert. Das Merkmal LH\_Stand wurde im Mittel signifikant häufiger (Überschätzung) und die Merkmale Abw. Fußstellung, Wunde und Haarlos (DG 2) signifikant weniger (Unterschätzung) erhoben als durch die externe Beurteilerin.

Beim Merkmal Verschmutzung im unteren hinteren Bereich war die größte Abweichung in der Beurteilung zu finden (minimaler PABAK -0,49). Im Mittel überschätzten hier die Kontrollassistent\*innen – das heißt, es wurden mehr Tiere als verschmutzt im unteren hinteren Bereich erhoben verglichen zur externen Beurteilerin. Bei der Erhebung von Verschmutzung im unteren hinteren Bereich im Melkstand ist anzumerken, dass aufgrund des Melkprozesses und der zweiten Person im Melkstand (Kontrollassistent\*in) die Tiere nervöser sein könnten und dies zu einem vermehrten Koten führen und die Verschmutzung erhöhen kann. Pelzer und Kaufmann (2016) nennen Lautstärke, Stress und Hektik als Ursachen für Abkoten im Melkstand.

Grund für die systematische Überschätzung des Merkmals LH\_Stand könnte Nervosität oder die veränderte Bedingung durch die Anwesenheit des\*der KA\*in im Melkstand sein. Häußermann (2020) weist darauf hin, dass Rinder durch Trippeln, Treten oder Schlagen auf für sie unangenehme bis schmerzhaft Bedingungen aufmerksam machen. Diese mit Entlastung einer Gliedmaße einhergehenden Verhaltensweisen könnten irrtümlicherweise als Lahmheit interpretiert worden sein.

Auffällig ist, dass bei KA\*in 1 häufig starke Abweichungen vorlagen. Gleichzeitig führte diese KA\*in 1 die meisten Erhebungen durch. Dies verdeutlicht die Bedeutung der Erhebung der Zuverlässigkeit, da sonst ein\*e einzelne\*r Beobachter\*in konsistent Merkmale falsch erheben kann (Martin und Bateson 1986).

Zusammenfassend lässt sich festhalten, dass für die meisten untersuchten Parameter eine zumindest zufriedenstellende Übereinstimmung erreicht werden konnte. Allerdings lagen für einzelne Beobachter\*innen eingeschränkte Wiederholbarkeiten vor. Damit eine akzeptable Zuverlässigkeit bei Beurteilungen von Lahmheit und anderen Merkmalen über die Zeit hinweg besteht, benötigen Beobachter\*innen wiederholte Schulungen und Training sowie eine regelmäßige Überprüfung der Übereinstimmung.

In der Praxis sind regelmäßige Trainingseinheiten aufgrund von Kosten oder geographischen Distanzen nicht immer durchführbar. Deshalb wird von Schlageter-Tello et al. (2014) im Zusammenhang mit Label-Programmen vorgeschlagen, dass erfahrene Beurteiler\*innen, für die kein weiteres Training erforderlich bzw. nicht möglich ist, die Lahmheitserhebung durchführen. Hierbei stellt sich jedoch die Frage, ob solche Beurteiler\*innen bei Zertifizierungen eingesetzt werden sollten, die ökonomische Konsequenzen für die landwirtschaftlichen Betriebe haben können.

## **6.5 Erfassung ‚hochgradig lahmer‘ Tiere: Beurteilung im Stand (KA\*in bzw. LW\*in) vs. Gangbeurteilung (externe Beurteilerin)**

Bei der Übereinstimmung von durch die externe Beurteilerin im Gehen als ‚hochgradig lahm‘ erfassten Kühen und der Beurteilung im Stand (durch den KA\*in bzw. LW\*in) lag auf Einzeltierebene eine hohe Spezifität vor (Mittel 96-88 %). Tiere, die auch laut EB als nicht hochgradig lahm eingestuft wurden, wurden von den LW\*innen/KA\*innen als nicht lahm im Stehen eingestuft. Jedoch war die Sensitivität gering (Mittel 30-45 %), d.h. der Anteil jener Tiere, die als lahm im Stand erhoben wurden und gleichzeitig von der externen Beurteilerin als hochgradig lahm eingestuft wurden, war niedrig. Eine niedrigere Sensitivität als Spezifität wurde auch von Mullan

et al. (2011) für tierbezogene Merkmale des Wohlergehens bei Schweinen beschrieben.

Eine hohe Sensitivität ist besonders wichtig, wenn Tests zur Erkennung von schweren Krankheiten verwendet werden. Anhand dieser Studie soll das Merkmal LH\_Stand zur Identifizierung lahmer Tiere und in weiterer Folge zur Verwendung der Daten in der Zuchtwertschätzung nicht verwendet werden, da eine zuverlässige Erfassung tatsächlich erkrankter Tiere, die Rückschluss auf eine genetische Prädisposition von Vererbern erlauben würden, durch die niedrige Sensitivität nicht gegeben ist. Demnach müsste eine Lahmheitsbeurteilung im Gehen durchgeführt werden, um sicher (hochgradig) lahme Tiere ermitteln zu können.

## **6.6 Aussagen der Teilnehmer\*innen zum Projekt Klauen-Q-Wohl**

Der Großteil der teilnehmenden Landwirt\*innen war positiv gegenüber der Anwendung des Protokolls eingestellt. Auch Knierim und Winckler (2009) betonen das hohe Interesse an den Ergebnissen aus der Erhebung von tierwohlbasierenden Merkmalen. Mullan et al. (2010) konnten ebenfalls bei Landwirt\*innen ein hohes Interesse und Zustimmung zu Tierwohlerhebungen feststellen. Dieses Interesse verringerte sich aber, sobald Konsequenzen oder Sanktionen integriert werden sollten.

Die in der Masterarbeit angewandten Merkmale bieten eine gute Möglichkeit, um der Betriebsblindheit entgegen zu wirken und das Stallmanagement zu verbessern (Aussage KA\*innen und LW\*innen). Leeb (2011) weist darauf hin, dass Probleme am Tier von den Halter\*innen unterschätzt und als solche nicht wahrgenommen werden. Um solcher Betriebsblindheit entgegenzuwirken, wird die Einführung einer externen Kontrolle empfohlen.

Bezüglich der Anwendung des dreistufigen Lahmheits-Scorings kam es zu unterschiedlichen Aussagen. Einige Teilnehmer\*innen sehen es als gute Methode an, um Lahmheiten zu erkennen, anderen ist dagegen die 3-stufige Lahmheitserkennung zu wenig differenziert. Eine stärkere Abstufung der Lahmheitskategorien geht jedoch häufig mit einer geringeren Übereinstimmung einher. Zum Beispiel beschreibt D'Eath (2012) Schwierigkeiten der Beobachter\*innen, zwischen leichter bzw. geringfügiger Lahmheit zu unterscheiden. Durch die Kombination von Kategorien erhöhte sich wie auch bei Brenninkmeyer et al., 2007; March et al., 2007 und Schlageter-Tello et al., 2014 die Übereinstimmung zwischen den Beobachter\*innen und es wurde vorgeschlagen, für die Anwendung von Tierwohlbeurteilung auf landwirtschaftlichen Betrieben ein einfaches Bewertungssystem (z. B. Welfare Quality®, 2009) vorzuziehen.

Die Kontrollassistent\*innen bezeichneten die Erhebung von Lahmheit im (Melk-) Stand während der Milchmessung als machbar, befürchteten aber, dass das Ergebnis durch das Verhalten der Rinder und die veränderten Situation (zusätzliche Person im Melkstand) verzerrt sein könnte. Weiters erwähnten die Kontrollassistent\*innen, dass die Lahmheitsbeurteilung im Laufgang genauer vorgenommen werden könnte. Diese Einschätzung deckt sich auch mit der geringen Sensitivität der Anwendung von LH\_Stand zur Erfassung hochgradig lahmer Tiere.

## 7 Literaturverzeichnis

1. Tierhaltungsverordnung, Anlage 2, 2.7 (2017): Mindestanforderung für die Haltung von Rindern. Betreuung.  
<https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20003820> (gesehen am 08.07.2021).
- AssureWel: Dairy Cows. [www.assurewel.org/dairy cows.html](http://www.assurewel.org/dairy cows.html) (gesehen am 13.02.2020)
- AssureWel (2016): Improving Farm Animal Welfare through Welfare Outcome Measures. Schenkenfelder (2016): Erstellung und Evaluierung eines Online-Schulungstools für tierbezogene Parameter des Wohlergehens von Rindern. Wien: Masterthesis Universität für Bodenkultur.
- Bahr, C.; Leroy, T.; Song XiangYu; Maertens, W.; Vranken, E.; Nuffel, A. van; Vangeyte, J.; Sonck, B. and Berckmans, D. (2008): Automatic detection of lameness in dairy cattle by vision analysis of cows' gait. Agricultural and Biosystems Engineering for a Sustainable World. International Conference on Agricultural Engineering. 23-25 June 2008, Hersonissos, Greece.
- Bennewitz, J.; Tetens, J. und die Gesellschaft für Tierzuchtwissenschaften e.V.(GfT) (2019): Stellungnahme der Gesellschaft für Tierzuchtwissenschaften e.V. (GfT). Nutztierhaltung im Umbruch: Kernpunkte und Perspektiven der Tierzuchtwissenschaften. Züchtungskunde, 91 (1), Verlag Eugen Ulmer, Stuttgart, S. 5–8. [https://www.zuechtungskunde.de/artikel.dll/zueku-2019-01-gft-stellungnahme\\_NjA4OTcyMQ.PDF?UID=262BB0AC84CCE721C3EBE2B70AB0033803975CF2DBE148](https://www.zuechtungskunde.de/artikel.dll/zueku-2019-01-gft-stellungnahme_NjA4OTcyMQ.PDF?UID=262BB0AC84CCE721C3EBE2B70AB0033803975CF2DBE148) (gesehen am 10.02.2021)
- Bio Austria (2018): Erhebungsbogen Tierwohl Rind. [https://www.bio-austria.at/app/uploads/Erhebungsbogen\\_tierwohl\\_rind.pdf](https://www.bio-austria.at/app/uploads/Erhebungsbogen_tierwohl_rind.pdf) (gesehen am 13.02.2020)
- Bio Austria: Lahmheit. <https://www.bio-austria.at/lahmheit/> (gesehen am 14.02.2020)
- Brade, W. (2013): Die Energiebilanz hochleistender Milchkühe aus der Sicht der Züchtung und des Tierschutzes. Prakt. Tierarzt 94, 536-544.
- Brenninkmeyer, C.; Dippel, S.; March, S.; Brinkann, J.; Winckler, C. and Knierim, U. (2007): Reliability of a subjective lameness scoring system for dairy cows. ANIMAL WELFARE; 16: 127-129
- Brinkmann, J.; Crimer, K., March, S., Ivemeyer; S., Pelzer, A.; Schultheiß, U.; Zapf, R. und Winckler, C. (2020): Tierschutzindikatoren. Leitfaden für die Praxis – Rind. Vorschläge für die Produktionsrichtung Aufzuchtkalb, Mastrind. 2. aktualisierte Auflage. KTBL.

- Brinkmann J., Ivemeyer S., Pelzer A., Winckler C., Zapf R (2016). Tierschutzindikatoren. Leitfaden für die Praxis – Rind. KTBL.
- Byrt, T.; Bishop, J.; Carlin, J.B. (1993): Bias, prevalence and kappa. *Journal of Clinical Epidemiology* 46, 423 – 429.
- Burn, C. C.; Pritchard, J. C. and Whay, H. R. (2009): Observer reliability for working equine welfare assessment: problems with high prevalences of certain results. *Animal Welfare*, 18, 177-187.
- Cha, E.; Hertl, J. A.; Bar, D. and Gröhn, Y. T. (2010): The cost of different types of lameness in dairy cows calculated by dynamic programming. *Prev Vet Med* 97, 1-8.
- Channon, A.; Walker, A. Pfau, T.; Sheldon, I. und Wilson, A. (2009): Variability of Manson and Leaver locomotion scores assigned to dairy cows by different observers. *Veterinary Record* 164, 388-392.
- Christen, A.-M.; Bergsten, C.; Burgstaller, J.; Capion, N.; Charfeddine, N.; Clarke, J.; Daniel, V.; Döpfer, D.; Fiedler, A.; Fjeldaas, T.; Heringstad, B.; Cramer, G.; Kofler, J.; Mueller, K.; Nielsen, P.; Oakes, E.; Ødegard, C.; O'Driscoll, K.; Pryce, J. E.; Steiner, A.; Stock, K. F.; Thomas, G.; Ulvshammar, K.; Holzhauser, M.; Cole, J.; weitere ICAR WGFT Mitglieder und internationale Klauengesundheitsexperten & Egger-Danner, C. (2015): Recording of claw and foot disorders in dairy cattle: current role and prospects of the international harmonization initiative of ICAR. ICAR Technical Meeting, 10.-12. Juni 2015, Krakau, Polen.
- Cohen, J. (1960): A coefficient of agreement for nominal scales. *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT*, 1960, 20, 37-46.
- Cohen, S. R.; Hassan, S. A.; Lapointe, B. J. and Mount, B. M. (1996): Quality of life in HIV disease as measured by the McGill quality of life questionnaire. *Aids* 10 (12), 1421–1427.
- Cook, N. B.; Hess, J. P.; Foy, M. R.; Bennett, T. B. and Brotzman R.L. (2016): Management characteristics, lameness, and body injuries of dairy cattle housed in high-performance dairy herds in Wisconsin. *J. Dairy Sci.*, 99 (2016), pp. 5879-5891.
- D'Eath R. B. (2012): Repeated locomotion scoring of a sow herd to measure lameness: Consistency over time, the effect of sow characteristics and inter-observer reliability. *Animal welfare (South Mimms, England)*. 21(2): 219 – 231.
- de Boyer des Roches, A.; Veissier, I.; Boivin, X.; Gilot-Fromont, E. and Mounier L. (2014): A prospective exploration of farm, farmer, and animal characteristics in human-animal relationships: An epidemiological survey. *J. Dairy Sci.*, 99 (2016), pp. 5573-5585.

- de Passillé, A. M. and Rushen, J. (2005): Can we measure human-animal interactions in on-farm animal welfare assessment? Some unresolved issues. *Applied Animal Behaviour Science* 92: 193-209.
- Dippel, S.; Dolezal, M.; Brenninkmeyer, C.; Brinkmann, J.; March, S.; Knierim, U. and Winckler, C. (2009): Risk factors for lameness in freestall-housed dairy cows across two breeds, farming systems, and countries. *J. Dairy Sci.*, 92 (2009), pp. 5476-5486.
- Distl, O., 1995. Züchterische Verbesserung von Fundamentmerkmalen und Klauengesundheit beim Rind. *Züchtungskunde* 67, 438-448.
- Dolecheck, K. and Bewley, J. (2018): Animal board invited review: Dairy cow lameness expenditures, losses and total cost. *Animal*. Volume 12, Issue 7, 2018, 1462-1474.
- Edmondson, A. J.; Lean, I. J.; Weaver, L. D.; Farver, T. and Webster, G. (1989). A body condition scoring chart for Holstein dairy cows. *J. Dairy Sci.* 72: 68–78.
- Egger-Danner, C. (2015): Züchterische Verbesserung der Klauengesundheit - Internationale Entwicklungen und Situation in Österreich. In: *Rinderzucht Austria. Gesunde Klauen und gute Fundamente Einflussfaktoren und Verbesserungsmaßnahme. Seminar des Ausschusses für Genetik der ZAR, 12. März 2015, Salzburg.* 35 – 45.
- Egger-Danner, C.; Fürst-Waltl, B.; Fürst, C.; Gruber, L.; Ledinek, M.; Steininger, F.; Zollitsch, W. und Zottl Z. (s.a.): Internationale Entwicklungen und Herausforderungen zur Zucht auf die effiziente Kuh. In: *ZAR-SEMINAR 2017. Der effizienten Kuh auf der Spur. Seminar des Ausschusses für Genetik der ZAR am 9. März 2017 in Salzburg.*
- Engel, B.; Bruin, G.; Andre, G. and Buist, W. (2003): Assessment of observer performance in a subjective scoring system: visual classification of the gait of cows. *Journal of Agricultural Science* 140. 317-333.
- EIP-AGRI: Klauen-Q-Wohl: Aufbau einer österreichweiten Infrastruktur zur zentralen standardisierten Erfassung und Auswertung von Daten zu Klauengesundheit, Lahmheit und Tierwohl. <https://ec.europa.eu/eip/agriculture/en/find-connect/projects/klauen-q-wohl-aufbau-einer-österreichweiten> (gesehen am 11.02.2020)
- Ferguson, J. D.; Galligan, D. T. und Thomsen, N. (1994): Principal Descriptors of Body Condition Score in Holstein Cows. Center for Animal Health and Productivity. University of Pennsylvania.
- Fleiss, J. L.; Levin, B. and Paik, M. C. (2003): *Statistical Methods for Rates and Proportions.* pp 598-626. John Wiles & Sons: Hoboken, NJ, USA.
- Flower, F. C and Weary, D. M. (2008): Gait assessment in dairy cattle. *Animal* (2009), 3:1, pp 87–95. The Animal Consortium 2009.

- Flower, F. C.; Sanderson D. J. and Weary, D. M. (2006): Effects of milking on dairy cow gait. *Journal of Dairy Science* 89. 2084-2089.
- Forkman, B. (2009): Investigating possible measures to include in the assessment system. In: Keeling, L. (2009): An overview of the development of the Welfare Quality® assessment systems. Edited by Linda Keeling. Welfare Quality Reports No. 12. SLU Service/Reproenheten. Uppsala. 9 – 14.
- Forrest, P. (1990): Breast cancer: the decision to screen. London. Nuffield Provincial Hospitals Trust.
- Fürst-Waltl, B.; Fürst, C. und Egger-Danner, C. (2015): Gutes Fundament und gesunde Klauen – wie viel sagt die lineare Nachzuchtbeschreibung über Klauengesundheit aus? In: Rinderzucht Austria (2015): Gesunde Klauen und gute Fundamente. Einflussfaktoren und Verbesserungsmaßnahmen. Seminar des Ausschusses für Genetik der ZAR, 12. März 2015, Salzburg. 47 - 56
- Garner, J. P.; Falcone, C.; Wakenell, P.; Martin, M. and Mench J. A. (2002): Reliability and validity of a modified gait scoring system and its use in assessing tibial dyschondroplasia in broilers. *British Poultry Science* 43: 355-363.
- Gieseke, D.; Lambertz, C.; Traulsen, I.; Krieter, J. und Gaulty M. (2014): Beurteilung von Tiergerechtigkeit in der Milchviehhaltung – Evaluierung des Welfare Quality® Protokolls. *Züchtungskunde*, 86, (1). Verlag Eugen Ulmer. Stuttgart. S. 58–70. <https://www.zuechtungskunde.de/Archiv/Beurteilung-von-Tiergerechtigkeit-in-der-Milchviehhaltung-Evaluierung-des-Welfare-Quality-Protokolls,QUIEPTQxNjMwMDEmTUIEPTY5MTU4.html?UID=2185372463CB588A4F56F7D9208B1ED6089B61213C090D> (gesehen am 10.02.2021)
- Grandin, T. (2015): Implementing Effective Standards and Scoring Systems and Scoring Systems for Assessing Animal Welfare on Farms and Slaughter Plants. *Improving Animal Welfare*, 2 Edition: A Practical Approach.
- Grüner Bericht (2021): Grüner Bericht 2021. Die Situation der österreichischen Land- und Forstwirtschaft. <https://gruenerbericht.at/cm4/jdownload/send/2-gr-bericht-terreich/2393-gb2021> (gesehen am 05.01.2021)
- Häusler (s.a.): „Kuhsignale“. Was sagt mir mein Rind- durch Beobachtung zu mehr Erfolg. [www.raumberg-gumpenstein.at/filearchive/fodok\\_1\\_12116\\_12\\_2012\\_haeusler\\_kuhsignale.pdf](http://www.raumberg-gumpenstein.at/filearchive/fodok_1_12116_12_2012_haeusler_kuhsignale.pdf) (gesehen am 15.02.2020)
- Häußermann, A. (2020): Messen und beobachten Wege zur Verbesserung des Melkvorgangs. *Landtechnik* 4.0. Bauernblatt. 6. Juni 2020. [36-38 Haeussermann.pdf \(lksh.de\)](#) (gesehen am 16.02.2022).
- Hoehler, F. K. (2019): Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. *Journal of clinical epidemiology* 53 (2000). 499-503

- Hollenbeck, A. R. (1978): Problems of reliability in observational research. In: Observing behaviour, vol. II. Data collection and analysis methods (ed. GP Sackett), pp. 79–98. University Park Press, Baltimore.
- Hulek, M. (2015): Klauengesundheit und Klauenpflege. Leopold Stockner Verlag – Stuttgart. 44 – 45
- Kaufman, A. and Rosenthal, R. (2009): Can you believe my eyes? The importance of interobserver reliability statistics in observations of animal behaviour. Elsevier.
- Keeling, L. (s.a.): Defining a framework for developing assessment systems. In: Keeling, L. (2009): An overview of the development of the Welfare Quality® assessment systems. Edited by Linda Keeling. Welfare Quality Reports No. 12. SLU Service/Reproenheten. Uppsala. 1 - 7.
- Knierim, U. and Winckler, C. (2009): On-farm welfare assessment in cattle: validity, reliability and feasibility issues and future perspectives with special regard to the Welfare Quality® approach. Animal Welfare 2009, 18, 451 – 458.
- König, S. und Swalve, H. H. (2006): Modellkalkulationen zu züchterischen Möglichkeiten auf Klauengesundheit beim Milchrind. Züchtungskunde 78: 345 – 356.
- Kofler, J (2014): Monitoring der Klauengesundheit in Milchviehherden und Funktionelle Klauenpflege. Lehr- und Forschungszentrum für Landwirtschaft Raumberg-Gumpenstein. 6. Tierärztetagung 2014. 67 – 76.
- Kofler J. und Winckler C. (2018): Piloterhebung zu tierbezogenen Parametern des Tierwohls im Projekt Klauen-Q-Wohl.
- Kümper, H. (2008): Managementmaßnahmen bei Erkrankungen des Bewegungsapparates des Rindes – Was kann der Landwirt selbst tun? Nutztierschutztagung 2008. Lehr und Forschungszentrum für Landwirtschaft. Irdning. 41 – 50.
- Landis, J.R. and Koch, G.G. (1977): The measurement of observer agreement for categorical data. Biometrics 33, 159-174.
- Leeb, C. (2011): The Concept of Animal Welfare at the Interface between Producers and Scientists: The Example of Organic Pig Farming. Acta Biotheoretica, 59, 173-183.
- Lemmens, L. (s.a.): Neue Technologien für die Lahmheitserkennung. In: LFI-Ländliches Fortbildungsinstitut Österreich (2021): Klauengesundheit im Griff mit System und Voraussicht. [https://blaetterkatalog.lko.at/oe/?catalog=Tiergesundheit\\_Klauengesundheit](https://blaetterkatalog.lko.at/oe/?catalog=Tiergesundheit_Klauengesundheit) (gesehen am 26.10.2021)
- Lorenzini, I.; Grimm, K.; Haidn, B. und Misha E. (2018): Development of a prediction model for automatic lameness detection in dairy cows. Bayerische Landesanstalt für Landwirtschaft, Institut für Landtechnik und Tierhaltung,

- Deutschland. 2ENGS Dairy Solutions, Israel. [https://www.lfl.bayern.de/mam/cms07/ilt/dateien/ilt3b\\_lahmheit\\_akal\\_entwicklung\\_vorhersagemodell.pdf](https://www.lfl.bayern.de/mam/cms07/ilt/dateien/ilt3b_lahmheit_akal_entwicklung_vorhersagemodell.pdf) (gesehen am 29.06.2021).
- Main, D. C. J.; Barker, Z. E.; Leach, K. A.; Bell N. J.; Whay H. R. and Browne W. J. (2010): Sampling strategies for monitoring lameness in dairy cattle. *J. Dairy Sci.* 93: 1970 – 1978.
- Main, D. C. J; Clegg, J.; Spatz, A. and Green, L. E. (2000): Repeatability of a lameness scoring system for finishing pigs. *The Veterinary Record* 147, 574–576.
- March, S.; Brinkmann, J. and Winkler, C. (2007): Effect of training on the inter-observer reliability of lameness scoring in dairy cattle. *ANIMAL WELFARE*; 16: 131-133.
- March, S.; Bergschmidt, A., Renziehausen, C. und Brinkmann, J. (2017) Indikatoren für eine ergebnisorientierte Honorierung von Tierschutzleistungen. Bonn: BÖLN, 280 p.
- Martin, P. and Bateson, P. P. G. (1986). *Measuring behaviour: An introductory guide.* Cambridge University Press.
- Martin, P. and Bateson, P. P. G. (1993). *Measuring behaviour: An introductory guide* (2nd ed.). Cambridge University Press.
- Martin, P. and Bateson, P. P. G. (2007). *Measuring behaviour: An introductory guide* (3rd ed.). Cambridge University Press.
- Metzner, M. (2009): Spezielle Untersuchung: Bewegungsapparat. Klinik für Wiederkäuer, Ludwig-Maximilians-Universität München. (gesehen am 15.02.2022: [2.7. Bewegungsapparat \(rinderskript.net\)](https://www.rinderskript.net)).
- Miesenberger, J.; Sölkner, J. und Eßl, A. (1996): Zuchtwertschätzung für Fruchtbarkeit und Kalbeverlauf. AGÖF (1994). In: ZAR – Zentrale Arbeitsgemeinschaft österreichischer Rinderzüchter (1996): Zuchtwertschätzung beim Rind. Grundlagen und aktuelle Entwicklung. Salzburg.
- Mullan, S.; Butterworth, A.; Whay, H. R.; Edwards, S. and Main, D. C. J. (2010): Consultation of pig farmers on the inclusion of some welfare outcome assessments within UK farm assurance. *Veterinary Record*, 166, 678-680.
- Mullan, S.; Edwards S. A.; Butterworth, A.; Whay, H.R and Main, D. C. J (2011): Inter-observer reliability testing of pig welfare outcome measures proposed for inclusion within farm assurance schemes. *The Veterinary Journal* 190. 100 – 109.
- Nalon, E.; Maes, D.; Van Dongen, S.; van Riet, M. M. J.; Janssens, G. P. J.; Millet, S. and Tuytens, F. A. M. (2014): Comparison of the inter- and intra-observer

- repeatability of three gait-scoring scales for sows. Cambridge University Press Animal (Cambridge, England), 2014-04, Vol.8 (4), p.650-659.
- Pelzer, A. und Kaufmann, O. (2016): Das Tier im Blick – Milchkühe. Hilfen zur systematischen Erfassung von Verhalten und Erscheinungsmerkmalen bei Milchkühen im. Milchviehstall. DLG Merkblatt 381. 5. Auflage, Stand: 10/2016. DLG e.V. Fachzentrum Landwirtschaft. Frankfurt am Main.
- Pritchard, J.C.; Barr, A.R.S. und Whay, H.R. (2007): Repeatability of a skin tent test for dehydration in working horses and donkeys. *Animal Welfare* 16, 181–183.
- Projektsitzung Klauen-Q-Wohl (2018): Klauen-Q-Wohl. Projekt zur österreichweiten Erfassung von Daten zu Klauengesundheit, Lahmheit, und Tierwohl und zur Ableitung von betriebsspezifischen Empfehlungen.
- Rajkondawar P. G.; Liu M.; Dyer R. M.; Neerchal N. K.; Tasch U.; Lefcourt A. M.; Erez B. and Varner M. A. (2006): Comparison of models to identify lame cows based on gait and lesion scores, and limb movement variables. *Journal of Dairy Science* 89, 4267–4275.
- Rinderzucht Austria (2017): Klauen-Q-Wohl. Einrichtung einer österreichweiten zentralen und standardisierten Erfassung und Auswertung von Daten zu Klauengesundheit, Lahmheit und Tierwohl. <https://zar.at/Projekte/Klauen-Q-Wohl.html> (gesehen am 11.02.2020).
- Rosenthal, R. (1976). *Experimenter effects in behavioral research*, (enlarged ed.). New York: Irvington Publishers.
- Rouha-Mülleder, C.; Iben C.; Wagner E.; Laaha, G.; Troxler, J. and Waiblinger, S. (2009): Relative importance of factors influencing the prevalence of lameness in Austrian cubicle loose-housed dairy cows. *Prev Vet Med* 92:123–133.
- Sarova, R.; Stehulova, I.; Kratinova, P.; Firla, P. and Spinka, M. (2011): Farm managers underestimate lameness prevalence in Czech dairy herds. *Animal Welfare*, 20, 201-204.
- Schlageter-Tello, A.; Bokkers, E. A. M.; Groot Koerkamp, P. W. G.; Hertemd, T. V.; Viazzi, S.; Romanini, C. E. B.; Halachmi, I.; Bahr, C.; Berckmans, D. and Lokhorst, K. (2014): Manual and automatic locomotion scoring systems in dairy cows: A review. *Preventive Veterinary Medicine* 116. 12 -15.
- Schlageter-Tello, A.; Van Hertem, T.; Bokkers, E. A. M.; Viazzi, S; Bahr, C. and Lokhorst, K. (2018): Performance of human observers and an automatic 3-dimensional computer-vision-based locomotion scoring method to detect lameness and hoof lesions in dairy cows. *J Dairy Sci*.
- Schober, P.; Boer, C. and Schwarte L. (2018): Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia*. May 2018. Volume 126, Issue 5, p. 1763-1768.

- Scott, G. B. (1988): Studies of the gait of Friesian heifer cattle. *The Veterinary Record* 123, 245–248.
- Scott, E. M.; Nolan, A. M. and Fitzpatrick, J. L. (2001): Conceptual and methodological issues related to welfare assessment: A framework for measurement. *Acta Agriculturae Scandinavica Section A Animal Science Supplementum* 30: 5-10
- Siegel, S. and Castellan, N. J. (1988): *Non-parametric Statistics for the Behavioural Sciences*. McGraw Hill College. New York.
- Sørensen, J.T.; Rousing, T.; Møller, S. H.; Bonde, M. and Hegel and. L. (2007): On-farm welfare assessment systems: what are the recording costs? *Animal Welfare* 2007, 16: 237-239.
- Sprecher, D. J.; Hostetler, D. E. and Kaneene, J. B. (1997): lameness scoring system that uses posture and gait to predict dairy cattle reproductive performance.
- Statistik Austria (2020): Betriebsstruktur. [https://www.statistik.at/web\\_de/statistiken/wirtschaft/land\\_und\\_forstwirtschaft/agrarstruktur\\_flaechen\\_ertraege/betriebsstruktur/index.html](https://www.statistik.at/web_de/statistiken/wirtschaft/land_und_forstwirtschaft/agrarstruktur_flaechen_ertraege/betriebsstruktur/index.html) (gesehen am 29.06.2021).
- Steininger, F.; Fuerst-Waltl, B.; Pfeiffer, C.; Fuerst, C.; Schwarzenbacher, H. und Egger-Danner, C. (2012): Participatory development of breeding goals in Austrian dairy cattle. *Acta Agriculturae Slovenica, Supplement* 3, 143-147.
- Streiner D. L. and Norman, G. R. (1995): *Health measurement scales. A practical guide to their development and use*. Oxford Medical Publications.
- Stygar, A. H.; Gómez, Y.; Berteselli, G. V.; Costa, E. D.; Canali, E.; Niemi, J. K.; Llonch, P.; Pastell, M. A. (2021): Systematic Review on Commercially Available and Validated Sensor Technologies for Welfare Assessment of Dairy Cattle. *Front. Veter. Sci.* 2021, 8, 177.
- Tetens J. (2019): Zucht auf Funktionalität und Gesundheit – von Datenflut und neuen Phänotypen. *Tierärztliche Praxis Ausgabe Nutztier* 2019. Georg Thieme Verlag KG Stuttgart New York. 47(04). 266. <https://www.thieme-connect.com/products/ejournals/abstract/10.1055/s-0039-1692753> (gesehen am 10.02.2021)
- Thomsen, P. T.; Munksgaard, L. and Tøgersen, F. A. (2008): Evaluation of a lameness scoring system for dairy cows. *Journal of Dairy Science* 91: 119-126.
- Tuytens, F. A. M.; Sprenger, M.; Van Nuffel, A.; Maertens, W. and Van Dongen, S. (2009): Reliability of categorical versus continuous scoring of welfare indicators: lameness in cows as a case study. *Animal Welfare* 18, 399–405.
- Tierschutzgesetz (TschG), Fassung vom 19.10.2017, BGBl Nr. 118/2004. Kontrollen. <https://www.ris.bka.gv.at/NormDokument.wxe?Abfrage=Bundesnormen&Gese>

[tzesnummer=20003541&FassungVom=2017-10-19&Artikel=&Paragraf=20&Anlage=&Uebergangsrecht=](#) (gesehen am 08.07.2021).

- Van der Tol, P. P. J.; Metz J. H. M.; Noordhuizen-Stassen E. N.; Back W.; Braam C. R. and Weijs W. A. (2002): The pressure distribution under the bovine claw during square standing on a flat substrate. *Journal of Dairy Science* 85, 1476–1481.
- Van Os, J. M. C.; Weary, D. M.; Costa, J. H. C.; Hötzel, M. J. and von Keyserlingk M. A. G. (2019): Sampling strategies for assessing lameness, injuries, and body condition score on dairy farms. *J. Dairy Sci.*, 102 (2019), pp. 8290-8304.
- Van der Waaij, E.; Holzhauer, M.; Ellen, E.; Kamphuis, C. and De Jong, G. (2005): Genetic Parameters for Claw Disorders in Dutch Dairy Cattle and Correlations with Conformation Traits. *J. Dairy Sci.* 88: 3672-3678.
- Van Reenen, K. and Engel, B. (2009). Testing the assessment system: refinement and definition. In: Keeling, L. (2009): An overview of the development of the Welfare Quality® project assessment systems. Cardiff University, Cardiff, UK. 33–42.
- Waiblinger, S. and Menke, C. (2003): Influence of sample size and experimenter on reliability of measures of avoidance distance in dairy cows. *Animal Welfare* 12: 585-590.
- Waiblinger, S.; Knierim, U. and Winckler, C. (2001): The development of an epidemiologically based on-farm welfare assessmentsystem for use with dairy cows. *Acta Agriculturae Scandinavica Section. Animal Science Supplementum* 30. 73-77.
- Welfare Quality® Consortium (2009): Welfare Quality®Assesment. Protocol for Cattle. Lelystad: The Netherlands.
- Wells, S. J.; Trent, A. M.; Marsh W. E. and Robinson R. A. (1993): Prevalence and severity of lameness in lactating dairy cows in a sample of Minnesota and Wisconsin herds. *J. Am. Vet. Med. Assoc.* 202:78–82.
- Whay, H. R.; Main, D. C. J.; Green, L. E. and Webster A. J. F. (2003a): Assessment of the welfare of dairy cattle using animal-based measurements: Observations and investigation of farm records. *Vet. Rec.*, 153 (2003), pp. 197-202.
- Whay, H. R.; Waterman, A. E.; Webster, A. J. F. and O'Brien, J. K. (1998): The influence of lesion type on the duration of hyperalgesia associated with hindlimb lameness in dairy cattle. *Vet. J.* 156, 23–29.
- Winckler C. and Willen, S. (2001): The reliability and repeatability of a lamenessscoring system for use as an indicator of welfare in dairy cattle. *Acta Agriculturae Scandinavica A (Supplement)* 30, 103–107. Gait assessment in dairy cattle 95.

Winckler, C.; Capdeville, J.; Gebresenbet, G.; Hörning, B.; Roiha, U.; Tosi, M. and Waiblinger, S. (2003): Selection of parameters for on-farm welfare assessment protocols in cattle and buffalo. *Animal Welfare* 12: 619.

ZuchtData EDV-Dienstleistungen GmbH (2019): Zuchtwertschätzung beim Rind. Grundlagen, Methoden und Interpretation. <http://www.zar.at/download/ZWS/ZWS.pdf> (gesehen am 11.02.2021)

ZuchtData EDV-Dienstleistungen GmbH (2020): ZuchtData Jahresbericht 2020. <https://zar.at/Downloads/Jahresberichte/ZuchtData-Jahresberichte.html> (gesehen am 11.02.2021)

ZuchtData EDV-Dienstleistungen GmbH (2021): Zuchtwertschätzung beim Rind. Grundlagen, Methoden und Interpretationen. <https://zar.at/dam/zar/downloads/ZWS.pdf> (gesehen am 26.10.2021)

## Anhang

Tab. 15: Unterschiede der Auftretenshäufigkeit in Prozent der Merkmale vom 1. auf den 2. Durchgang vom Ergebnis der externen Beurteilerin

Merkmale N 22	DG 1 – DG 2 Mittelwert (%)	DG 1- DG 2 p-Wert
LH_Gang	2,87	0,272
LH_Hochgradig	0,70	0,698
LH_Stand	1,70	0,161
Fußstellung	2,04	0,439
Haarlos	<b>-5,78*</b>	0,037
Wunde	-4,57	0,078
Schwellung	-0,65	0,328
Verschm_OH	0,91	0,488
Verschm_UH	2,13	0,553

Tab. 16: Mittlere Abweichung in %-Punkte aller LW\*innen zur externen Beurteilerin mit p-Werte und Prävalenzen (Auftretenshäufigkeiten)

Alle LW*innen	DG 1			DG 2		
	Prävalenzen Mittel (%)	Mittelwert (%-P. Abw.)	Abw. p-Wert	Prävalenzen Mittel (%)	Mittelwert (%-P. Abw.)	Abw. p-Wert
LH_Gang (N=18/19)	28	-9,0	0,002	25	-6,1	0,014
LH_Hochgradig (N=18/19)	11	-9,0	0,002	10	-7,0	0,008
LH_Stand (N=18/18)	5	4,1	0,018	3	4,4	0,014
Fußstellung (N=18/19)	30	-6,9	0,103	28	-7,9	0,11
Haarlos (N=18/19)	28	-3,8	0,231	34	-7,0	0,019
Wunde (N=17/18)	15	-6,7	0,003	20	-10	0,004
Schwellung (N=17/18)	3	0,3	0,749	3	-1,0	0,377
Verschm_OH (N=18/19)	9	4,1	0,171	8	3,9	0,209
Verschm_UH (N=18/19)	38	-6,5	0,323	33	-8,3	0,246

Tab. 17: Mittlere Abweichung in %-Punkte aller KA\*innen zur externen Beurteilerin mit p-Werte und Prävalenzen (Auftrittshäufigkeiten)

Alle KA*innen	DG 1			DG 2		
	Prävalenzen Mittel (%)	Mittelwert (%-P. Abw.)	Abw. p-Wert	Prävalenzen Mittel (%)	Mittelwert (%-P. Abw.)	Abw. p-Wert
LH_Stand (N=22)	5	9,3	0,002	3	11,9	<,001
Abw. Fußst. (N=21/22)	30	-18,6	<,001	28	-13,3	0,003
Haarlos (N=20)	28	-4,1	0,106	34	-5,9	0,063
Wunde (N=20)	15	-7,3	0,018	20	-11,7	0,001
Schwellung (N=20)	3	3,9	0,073	3	2,3	0,267
Verschm_UH (N=20)	38	10,2	0,086	33	13,6	0,075

## Schäden am Sprunggelenk

### 1 Haarlose Stellen

- Außenseite im Bereich Sprunggelenk + gegenüberliegende Innenseite
- haarlose Stellen ab einem Durchmesser von 2,5 cm (1 € Münze) und größer

### 2 Wunden/Krusten

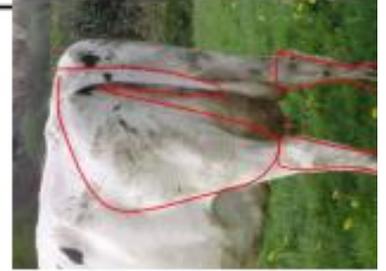
- Außenseite im Bereich Sprunggelenk + gegenüberliegende Innenseite
- Wunde frisch oder verkrustet, jeweils > 2,5 cm (1 € Münze)

### 3 Schwellungen

- Im Bereich Sprunggelenk
- Eindeutig mit bloßem Auge erkennbare Umfangsvermehrung im Vergleich zum Normalzustand

### 4 Verschmutzung

- **Sauber**: Keine Verschmutzungen bzw. nur nasses Fell oder Verfärbung ohne Kotauflagerungen
- **„Verschmutzt“**: Oberes Hinterbein inkl. Schwanz ohne Quaste (mind. 30 cm – Unterarmlänge durchgehende Kotauflagerungen  
Unteres Hinterbein inkl. Innenseite des gegenüberliegenden Hinterbeins (in Summe mind. handtellergröße Kotauflagerungen)



## BCS - Körperkondition

### BCS - Optische Beurteilung

BODY CONDITION SCORE	Verbindungs- linie Dorn- zu Querfort- sätzen	Hinteransicht Hüftbeinhöcker	Seitenansicht der Verbindungs- linie zw. Hüft- u. Sitzbeinhöcker	Höhle zwischen Schwanzansatz u. Sitzbeinhöcker
1	hochgradig abgemagert			
2	Knochenvorsprünge sichtbar			
3	Knochenvorsprünge gut abgedeckt			
4	Knochenvorsprünge angedeutet			
5	hochgradig verfettet			

Body condition scores (nach A.J. Edmondson et al., 1989)

**Punktevergabe von  
1,00 – 5,00  
(1/2 Punkte-Abstände  
z.B. 2,5)**

Abb. 27: Schulungsunterlagen; laminierte Erhebungskarte von den Merkmalen Haarlos, Wunden, Schwellung und Verschmutzung für die Erhebung direkt im Stall; Seite 1

## Lahmheit - Gangbeurteilung

**0 Nicht lahm:** elastischer Gang, raumgreifende Schritte, keine Unregelmäßigkeiten in der Schrittfolge  
Einschließlich **unebener Gang:** z.B.: steif, vorsichtiges Gehen, geht möglicherweise langsamer

### Kriterien für Lahmheit:

**1 Lahm:** unregelmäßige Schrittfolge durch Entlastung eines Beines kann mit aufgekürmtem Rücken verbunden sein

**2 Hochgradig lahm:** deutliches Widerstreben ein Bein zu belasten oder Entlastung von mehr als einem Bein

## Lahmheit - Beurteilung im Stand (Melkstand)

**0 Nicht lahm**

### Kriterien für Lahmheit:

**1 Lahm,** falls eines der drei folgenden Kriterien zutrifft:

- Herumtrippeln
- wiederholtes Anheben der Gliedmaße
- Belastung nur eines Teils der Klaue, z.B. Aufsetzen der Klauenspitze



## Fußstellung

**0 Gut/akzeptabel**

- die Hinterbeine stehen (**fast**) parallel
- die Außenklauen sind (**fast**) gleich hoch wie die Innenklauen
- innen an den Fersenhöcker angelegte Senkrechte trifft auf die Klauen

**1 Zu weit**

- deutlich zu hohe Außenklauen,
- Hinterbeine nach außen gestellt
- Klauen nach außen gedreht
- innen an den Fersenhöcker angelegte Senkrechte trifft **nicht** auf die Klauen

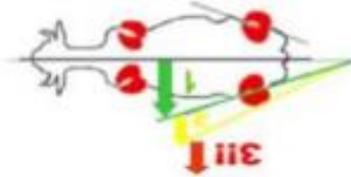


Abb. 28: Schulungsunterlagen; laminierte Erhebungskarte von den Merkmalen Lahmheit Gangbeurteilung, Lahmheit Beurteilung im Stand und abweichende Fußstellung für die Erhebung direkt im Stall; Seite 2