

Tierärztliche Hochschule Hannover

**Tierwohlbewertung von Mastschweinen  
durch übergreifende Analyse vorhandener Produktionsdaten  
und aktiv erfasster Variablen im Betrieb**

INAUGURAL – DISSERTATION

Zur Erlangung des Grades einer

Doktorin der Veterinärmedizin

– Doctor medicinae veterinariae –

(Dr. med. vet.)

vorgelegt von

Julia Große-Kleimann

Göttingen

Hannover 2021

**Wissenschaftliche Betreuung** Prof. Dr. Lothar Kreienbrock  
Institut für Biometrie, Epidemiologie und  
Informationsverarbeitung  
Stiftung Tierärztliche Hochschule Hannover

**1. Gutachter** Prof. Dr. Lothar Kreienbrock  
Institut für Biometrie, Epidemiologie und  
Informationsverarbeitung  
Stiftung Tierärztliche Hochschule Hannover

**2. Gutachterin** Prof. Dr. Nicole Kemper  
Institut für Tierhygiene, Tierschutz und  
Nutztierethologie  
Stiftung Tierärztliche Hochschule Hannover

**Tag der mündlichen Prüfung** 4. November 2021

Die Forschungen zu der vorliegenden Arbeit erfolgten im Rahmen des Projektes "MulTiViS – Multivariate Bewertung des Tierwohls durch integrative Datenerfassung und Validierung von Tierwohlindikatoren in Schweinebeständen", welches von der Bundesanstalt für Landwirtschaft und Ernährung finanziert wurde (Förderkennzeichen 2817905315).

Meiner Familie

## **Dieser Dissertation liegen zwei Publikationen zugrunde:**

Grosse-Kleimann J, Plate H, Meyer H, Gerhardy H, Kreienbrock L.

### **Health Monitoring of Finishing Pigs by Secondary Data Use – A Longitudinal Analysis.**

Porcine Health Management 2021; 7(20).

Grosse-Kleimann J, Wegner B, Spiekermeier I, grosse Beilage E, Kemper N, Nienhoff H,  
Plate H, Meyer H, Gerhardy H, Kreienbrock L.

### **Health Monitoring of Fattening Pigs – Use of Production Data, Farm Characteristics and On-Farm Examination.**

Porcine Health Management 2021; 7(45).

## **Ergebnisse dieser Dissertation wurden auf folgenden Fachkonferenzen durch die Autorin präsentiert:**

Große-Kleimann J, Plate H, Meyer H, Gerhardy H, Kreienbrock L.

### **Erste Bewertung der Nutzbarkeit von Routinedaten als Tierwohlindikatoren in der Schweinemast. (Vortrag)**

DVG DACH-Epidemiologietagung 2018. Greifswald, 6. September 2018.

Grosse-Kleimann J, Wegner B, Spiekermeier I, grosse Beilage E, Kemper N, Nienhoff H, Plate H, Meyer H, Gerhardy H, Kreienbrock L.

### **Welfare Assessment of Fattening Pigs using routinely collected and edited Production Data. (Vortrag)**

ESPHM 2019. Utrecht (Niederlande), 23. Mai 2019.

Grosse-Kleimann J, Wegner B, Spiekermeier I, grosse Beilage E, Kemper N, Nienhoff H, Plate H, Meyer H, Gerhardy H, Kreienbrock L.

### **Health Scores in Finishing Pigs: Do not Forget to Stratify by Initial Body Weight. (Poster)**

SVEPM 2021. Online, 24.-26. März 2021.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Definition und Bewertung von Tierwohl	1
1.2	Tierwohlmonitoring und Sekundärdatennutzung	2
1.3	Das Verbundprojekt MuTiViS	3
1.4	Struktur der vorliegenden Arbeit	5
<b>2</b>	<b>Publikationen</b>	<b>6</b>
2.1	Health Monitoring of Finishing Pigs by Secondary Data Use – A Longitudinal Analysis	6
2.2	Health Monitoring of Fattening Pigs – Use of Production Data, Farm Characteristics and On-Farm Examination	20
<b>3</b>	<b>Übergreifende Diskussion</b>	<b>34</b>
3.1	Planung einer umfassenden Tierwohlstudie	34
3.1.1	Selektion der Projektbetriebe	34
3.1.2	Aufbereitung von Sekundärdaten	35
3.1.3	Erhebungskonzept	43
3.2	Auswertung der Daten und Interpretation der Ergebnisse	47
3.2.1	Festlegung der Tierwohlindikatoren und Transformation zu Tiergesundheits-Scores	48
3.2.2	Untersuchung potentieller Einflussfaktoren aus den Produktionsdaten	51
3.2.3	Gegenüberstellung der Indikatoren aus Produktions- und Erhebungsdaten	52
3.3	Erforderliche Schritte für ein nationales Tierwohlmonitoring	54
3.4	Abschließende Bewertung	56
<b>4</b>	<b>Zusammenfassung</b>	<b>58</b>
<b>5</b>	<b>Summary</b>	<b>60</b>
<b>6</b>	<b>Literaturverzeichnis</b>	<b>62</b>
<b>7</b>	<b>Danksagung</b>	<b>66</b>

## Abkürzungsverzeichnis

ADG	Tiergesundheits-Score Tägliche Zunahmen ( <i>average daily gain</i> )
BMEL	Bundesministerium für Ernährung und Landwirtschaft
BRS	Bundesverband Rind und Schwein e.V.
EXT	Tiergesundheits-Score Äußere Verletzungen ( <i>external injuries</i> )
FAWC	Farm Animal Welfare Council
FCR	Tiergesundheits-Score Futterverwertung ( <i>feed conversion ratio</i> )
MANG	Tiergesundheits-Score Tiermanagement ( <i>animal management</i> )
MOR	Tiergesundheits-Score Verluste ( <i>mortality</i> )
MSG	Marketingservice Gerhardy
MuTiViS	Multivariate Bewertung des Tierwohls durch integrative Datenerfassung und Validierung von Tierwohlintikatoren in Schweinebeständen
QS	Qualität und Sicherheit GmbH
RESP	Tiergesundheits-Score Respirationstrakt ( <i>respiratory tract</i> )
SGD	Schweinegesundheitsdienst der Niedersächsischen Landwirtschaftskammer
TiHo	Stiftung Tierärztliche Hochschule Hannover
TF	Tiergesundheits-Score Therapiehäufigkeit ( <i>treatment frequency</i> )
VzF	Verein zur Förderung der bäuerlichen Veredelungswirtschaft e.V.

# 1 Einleitung

## 1.1 Definition und Bewertung von Tierwohl

Der Begriff Tierwohl kann als direkte Übersetzung des englischen Ausdrucks "Animal Welfare" angesehen werden. Die inhaltliche Definition und die fachliche Auslegung jedoch unterscheiden sich. Während im angelsächsischen Sprachraum "Animal Welfare" als sehr umfassender Terminus verwendet wird, werden im Deutschen verschiedene Teilaspekte, wie zum Beispiel Tiergerechtigkeit, Tierschutz oder Tiergesundheit definiert. Bereits im Jahr 1966 wurde in den USA der Animal Welfare Act zum Schutz von Versuchstieren eingeführt, der aber auch als Mindeststandard für alle anderen die Tierhaltung betreffenden Gesetze gilt (USDA). In Deutschland hingegen wird das Wohlergehen der Tiere über das Tierschutzgesetz bzw. die Tierschutz-Nutztierhaltungsverordnung geregelt – ein "Tierwohlgesetz" gibt es nicht.

Im Jahr 1965 wurde in Großbritannien durch den Brambell Report bzw. 1979 das Farm Animal Welfare Council (FAWC) der Begriff "Animal Welfare" mit internationaler Reichweite geprägt (Brambell 1965; FAWC 1979). Als Grundlage hierfür gilt das Konzept der fünf Freiheiten. Diese umfassen neben körperlichen Aspekten (Freiheit von Hunger und Durst, Freiheit von Schmerz, Verletzungen und Krankheiten) auch das Tierverhalten (Freiheit zum Ausleben normaler Verhaltensmuster) sowie das emotionale Wohlbefinden (Freiheit von Angst und Stress). Die Freiheit von haltungsbedingten Beschwerden kann sowohl der physischen als auch der psychischen Unversehrtheit zugeordnet werden.

Mit den fünf Freiheiten wurde ein erster Grundstein gelegt, um das Thema Tierwohl greifbarer und vor allem messbar zu machen. Viele nationale und internationale Studien haben sich seither damit befasst, geeignete Tierwohlindikatoren für die Nutztierhaltung zu entwickeln (EFSA 2012; IGN 2016; Schrader 2013; Welfare Quality 2009). Generell kann man dabei zwischen tierbezogenen und nicht-tierbezogenen (unterteilt in ressourcen- und managementbezogene) Indikatoren unterscheiden (EFSA 2012). Erstere werden – direkt oder indirekt – am Tier gemessen und letztere beziehen sich auf die Einflussfaktoren aus der Umgebung des Tieres.

Bei der Erstellung der vorliegenden Arbeit hat sich gezeigt, dass vor allem die Tiergesundheit als wichtiger Bestandteil von Tierwohl standardisiert und objektiv messbar ist. Zu diesem Thema wurde bereits viel geforscht und es hat sich eine Anzahl von Indikatoren in der Praxis etabliert (Schrader *et al.* 2016). Unklarer wird es dagegen bei der Operationalisierung von Verhalten und Emotionen wie Leiden oder Freude. Dass diese maßgeblich zum Wohlbefinden beitragen, ist unbestritten (Dawkins 2004; Duncan *et al.* 1983), jedoch haben sich bisher kaum valide Messinstrumente etabliert. Grund hierfür ist, dass die Beobachtung und Beurteilung des Tierverhaltens viel Zeit und ein geschultes Auge erfordern, was meist zu Lasten der Praktikabilität geht. Hinzu kommt die Tatsache, dass Tiere nicht selbst über ihr Wohlbefinden berichten können, weshalb zur Interpretation auf indirekte Methoden zurückgegriffen werden muss. Dazu wurden zum Beispiel Präferenz- oder Verhaltenstests entwickelt und untersucht (Duncan 2005; Wegner *et al.* 2020). Eine andere Möglichkeit ist die Beurteilung gesundheitlicher Folgen von Leiden im psychischen Sinne. So wird zum Beispiel das Schwanzbeißen beim Schwein als atypisches Verhalten definiert, das durch unpassende, nicht der Natur entsprechende Haltungsbedingungen verursacht wird (Sonoda *et al.* 2013). Dieses Verhalten direkt zu beobachten erfordert jedoch eine längere Betrachtungszeit, so dass meist die Unversehrtheit des Schwanzes bzw. dessen Länge als (gesundheitlicher) Indikator herangezogen werden (Valros *et al.* 2020).

Die standardisierte Erfassung von Tierwohl auf allen Ebenen (psychisch und physisch) ist also bisher nicht vollumfänglich etabliert, weshalb der Begriff Tierwohl in dieser Arbeit synonym für Tiergesundheit im weiteren Sinne benutzt wird.

## **1.2 Tierwohlmonitoring und Sekundärdatennutzung**

Die artgerechte Haltung von Tieren, die der Lebensmittelproduktion dienen, ist in den letzten Jahren zunehmend stärker in den Fokus der Öffentlichkeit gerückt (Grunert *et al.* 2018; Tonsor *et al.* 2009). Dabei geht die Forderung nach mehr Tierwohl oft mit der Vorstellung von kleinen, extensiv oder biologisch wirtschaftenden Betrieben einher (Initiative Tierwohl 2015).

Es gibt bereits eine Vielzahl von Bio-Siegeln, sowohl in Deutschland als auch der EU, die eine kontrolliert ökologische Tierhaltung mit entsprechend hohem Tierwohl auszeichnen (Rohleder 2017). Allerdings ist der Anteil biologisch gehaltener Nutztiere in Deutschland – noch – relativ gering. 2016 lag dieser bei 0,7% für die Tierart Schwein (DESTATIS 2018). Um also die Mehrheit der Tiere, die der Lebensmittelproduktion dienen, abzudecken, muss zunächst die konventionelle Landwirtschaft in den Mittelpunkt gestellt werden. Diverse Stakeholder haben bereits Lösungsansätze entwickelt, um diese herkömmliche Nutztierhaltung, die in der öffentlichen Diskussion oft auch als "Massentierhaltung" bezeichnet wird, von ihrem negativen Image zu befreien. So wurde 2015 die Initiative Tierwohl aus einer Kooperation von Landwirtschaft, Fleischwirtschaft, Lebensmittelhandel und Gastronomie ins Leben gerufen (Initiative Tierwohl 2015). Als Reaktion darauf ging im April 2019 offiziell die einheitliche "Haltungsform-Kennzeichnung" verschiedener Discounter auf den Markt und bereits 2016 wurde im Rahmen der Nutztierhaltungsstrategie des Bundesministeriums für Ernährung und Landwirtschaft (BMEL) ein nationales Tierwohllabel angekündigt (BMEL 2016). Zur Sicherstellung einer einheitlichen und flächendeckenden Dokumentation forderte die deutsche Tierärzteschaft 2018 die Etablierung eines nationalen Monitoringprogramms von Tiergesundheit und Tierschutz, mit dessen Hilfe fortlaufend ein Status Quo erhoben und somit frühzeitig Handlungsbedarf erkannt werden kann (Tiedemann 2018).

### **1.3 Das Verbundprojekt MuTiViS**

Um in großem Rahmen Informationen zu Tierwohl bzw. Tiergesundheit von Nutztieren zu bekommen, ist ein immenser Aufwand erforderlich. Eine speziell für diesen Zweck durchgeführte Erhebung würde erhebliche ökonomische und personelle Ressourcen verbrauchen. Eine Alternative stellt die Nutzung bereits vorhandener Daten aus der Wertschöpfungskette dar und auf diesem Ansatz beruht das Verbundprojekt "MuTiViS" (Multivariate Bewertung des Tierwohls durch integrative Datenerfassung und Validierung von Tierwohlintaktoren in Schweinebeständen), in Zusammenarbeit von der Stiftung

Tierärztliche Hochschule Hannover (TiHo), dem Agrardienstleister Verein zur Förderung der bäuerlichen Veredelungswirtschaft e.V. (VzF), dem Schweinegesundheitsdienst der Niedersächsischen Landwirtschaftskammer (SGD) und dem Marketingservice Gerhardy (MSG). Das Projekt befasst sich mit der Fragestellung, ob Sekundärdaten aus der Produktionskette Mastschwein die spezielle tierärztliche Untersuchung im Bestand zum Zwecke eines Tierwohlmonitorings ergänzen oder sogar ersetzen können.

Dazu wurden die Produktionsdaten von ca. 200 Mitgliedsbetrieben des VzF für die Wirtschaftsjahre 2017/18 und 2018/19 geprüft und aufbereitet, um mit Hilfe von speziell entwickelten Tiergesundheits-Scores eine erste Beurteilung des Tierwohlstatus der in den Projektbetrieben gehaltenen Schweine vornehmen zu können. Die Produktionsdaten setzen sich wie folgt zusammen:

- (1) Ergebnisse aus der Betriebszweiganalyse des VzF
  - a) biologische und ökonomische Leistungsdaten
  - b) Betriebseigenschaften
- (2) Informationen zum Antibiotikaeinsatz
- (3) Schlachtbefunde aus der routinemäßigen Fleischuntersuchung

Die Daten zu (2) und (3) werden von der Qualität und Sicherheit GmbH (QS) erfasst und an den VzF übermittelt, wo sie in das firmeneigene Datenverarbeitungssystem eingepflegt und zusammen mit den anderen Produktionsdaten für das Projekt zur Verfügung gestellt wurden. In einem zweiten Schritt erfolgte die tierärztliche Untersuchung der Schweine auf den Projektbetrieben. Dazu wurden diese einmalig im Zeitraum von November 2017 bis November 2018 von zwei speziell geschulten Projekttierärztinnen an einem zufälligen Stichtag besucht und mittels eines vierteiligen Erhebungsbogens Informationen zur Tierumgebung und der Tiergesundheit erfasst.

## 1.4 Struktur der vorliegenden Arbeit

Zunächst wurde in **Publikation 1** die technische Eignung von sowohl routinemäßig als auch speziell erfassten und anschließend aufbereiteten Produktionsdaten von Schweinemastbetrieben untersucht. Dabei handelt es sich um Daten auf Halbjahresbasis, die aus der Betriebszweiganalyse des VzF stammen sowie Informationen zum Antibiotikaeinsatz und den Schlachtbefunden. Es zeigte sich, dass für die Sekundärdatennutzung zum Zwecke eines Monitorings die Schichtung nach Einstallgewichten zur Mast für einige Indikatorvariablen unerlässlich ist. Außerdem erfolgte die Entwicklung und Längsschnittbetrachtung standardisierter Tiergesundheits-Scores, die im Sinne eines Screenings genutzt werden können.

**Publikation 2** beschäftigt sich damit, inwiefern die etablierten Tiergesundheits-Scores durch betriebliche Eigenschaften und Managemententscheidungen beeinflusst werden können. Zudem erfolgte eine Gegenüberstellung der Schlachtbefunde und den darauf beruhenden Tiergesundheits-Scores mit (weitestgehend) äquivalent erfassten Gesundheitsvariablen im Betrieb, um daraus Handlungsempfehlungen für potentielle zusätzliche Indikatoren im Rahmen eines Monitorings abzuleiten.

In der **übergreifenden Diskussion** wird herausgestellt, welche Erkenntnisse aus der Projektarbeit neben den Ergebnissen der oben genannten Veröffentlichungen entstanden sind. Dazu werden Faktoren beleuchtet, die bei der Planung und Auswertung einer solchen Studie berücksichtigt werden sollten. Zur grundlegenden Vorbereitung gehören dabei die Selektion der Projektbetriebe, die Aufbereitung der verwendeten Daten sowie die Konzeption der Erhebung in den einzelnen Betrieben. Weiterhin werden potentielle Fallstricke bei der übergreifenden Auswertung der Daten und fachlichen Interpretation der Ergebnisse beleuchtet. Zum Abschluss werden die Ergebnisse in einen fachlichen Kontext eingeordnet und dargestellt, welche Schritte für ein nationales Tierwohlmonitoring notwendig sind.

## 2 Publikationen

### 2.1 Health Monitoring of Finishing Pigs by Secondary Data Use – A Longitudinal Analysis

Große-Kleimann J, Plate H, Meyer H, Gerhardy H, Kreienbrock L. Health Monitoring of Finishing Pigs by Secondary Data Use – A Longitudinal Analysis. *Porcine Health Management* 2021; 7(20) DOI 10.1186/s40813-021-00197-z.

**Journal** Porcine Health Management

**Editor** Prof Paolo Martelli (Università di Parma, Italy)

Prof Joaquim Segalés (Universitat Autònoma de Barcelona & Centre de Recerca en Sanitat Animal, Spain)

**Eingereicht** 14. Dezember 2020

**Akzeptiert** 28. Januar 2021

**Veröffentlicht** 24. Februar 2021

RESEARCH

Open Access



# Health monitoring of finishing pigs by secondary data use – a longitudinal analysis

Julia Grosse-Kleimann<sup>1\*</sup> , Heiko Plate<sup>2</sup>, Henning Meyer<sup>2</sup>, Hubert Gerhardy<sup>3</sup>, Corinna Elisabeth Heucke<sup>1</sup> and Lothar Krienbrock<sup>1</sup>

## Abstract

**Background:** In Germany, animal welfare has become an increasingly important issue. Since 2006, German legislation demands self-monitoring of animal welfare by farmers, but there is a lack of prescribed indicators for governmental monitoring. Since recording of the health status through examinations on individual farms requires many resources, secondary data use is obvious. Therefore, this study deals with the overall evaluation and utilization of existing production data from the German pork production. Performance data and information on antibiotic usage and meat inspection were used for a benchmarking system of animal health in finishing pigs.

**Results:** Seven health scores and one total score were evaluated for 184 finishing pig herds on semi-annual basis between July 2017 and June 2019, based on the health indicators mortality, average daily gain, feed conversion ratio, treatment frequency, respiratory lesions, exterior lesions and animal management. In preparation, the selected health indicators were brought to the same scale and skewed data were transformed to build scores (MOR, ADG, FCR, TF, RESP, EXT and MANG). A differentiated analysis was carried out for three classes of initial body weight regarding to farmers' fattening management strategies.

**Conclusions:** The present study shows that existing production data of German finishing pigs are usable for welfare monitoring. However, preparatory editing steps are crucial. The total score can only be an estimate of health status because partly bad or good performance could be disguised. It has also been demonstrated, that relative benchmarking is suitable for depicting temporary fluctuations in the investigated collective.

**Keywords:** Animal welfare, Antimicrobial treatment, Meat inspection, Performance data

## Background

Animal welfare is an issue that goes along with livestock farming in Germany for decades. Nevertheless, in the last couple of years, farm animal health and welfare have substantially attracted political and social attention as consumers demand transparency on conditions under

which food-producing animals are kept [1, 2]. Since 2006, the eleventh paragraph of the German Animal Protection Law prescribes the gathering and assessment of suitable animal-related items by all livestock keepers in a programme of self-monitoring [3]. Different stakeholders specified several indicators [4–7], but a mandatory set of suitable items as well as a general scientific evaluation are still missing. Furthermore, the term “animal welfare” was not clearly defined which leads to a lack of harmonisation and validation of data capturing because interpretation and implementation differ. The

\* Correspondence: [julia.grosse-kleimann@tiho-hannover.de](mailto:julia.grosse-kleimann@tiho-hannover.de)

<sup>1</sup>Department of Biometry, Epidemiology and Information Processing, WHO Collaborating Centre for Research and Training for Health in the Human-Animal-Environment Interface, University of Veterinary Medicine Hannover, Foundation, Buenteweg 2, 30559 Hanover, Germany  
Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

EFSA approach to define welfare describes two groups of indicators [8]: those predicting animal welfare directly or indirectly (animal-based measures), and those having an influence on animal welfare (non-animal-based measures or resource- and management-based measures). Another definition from the Brambell Report in 1965 [9] describes welfare with the help of five freedoms: freedom from hunger and thirst (I), discomfort (II), pain, injury and disease (III), fear and distress (IV) and the possibility to express normal behaviour (V). Following this approach, physical health is a major part of welfare (freedoms I and III), in addition to behavioural and mental aspects (freedoms II, IV and V). As there are only few methods to measure latter ones, whereas health measures are well established, this paper focusses on physical and biological animal-based measures as health indicators.

However, establishing the health status for each individual farm is very costly and time consuming. Therefore, this study deals with the use of already existing data from the German pork production chain for health monitoring and a suitable benchmarking system. Herd-specific data of 205 finishing pig farms for a time-span of 2 years from public or private databases were merged and analysed. We had both production data, levied within the frame of quality assurance and consulting, and mandatory data concerning antibiotic usage and meat inspection. For Germany, this is a unique approach because there is no official or governmental monitoring programme dealing with a multiple set of indicators [1].

## Material and methods

### Study design

The study was carried out in the frame of the project “Multivariate Assessment of Animal Welfare through Integrative Data Collection and Validation of Welfare Indicators in Finishing Pigs” (MulTiViS), which was supported by the German Federal Ministry of Food and Agriculture. The project was launched in April 2017 by a consortium of the University of Veterinary Medicine Hannover (“Tierärztliche Hochschule Hannover”, TiHo), the Swine Health Service of the Chamber of Agriculture in Lower Saxony (“Schweinegesundheitsdienst”, SGD), the swine service provider VzF (registered association) (VzF) and Marketing Service Gerhardt (MSG). All pig herds under study were advised by VzF and were participants of the “QS Quality Scheme for Food” (QS). This scheme covers approximately 95% of the German pig husbandry [10] and includes, among others, a monitoring programme for antibiotic usage and a data collection of results of official meat inspection at slaughter.

Out of different production directions, we decided to put the focus of this study on the finishing phase, as there is most reliable data available. In addition to this,

the aim of the project was to create a benchmarking system which is adapted to fluctuation in time. For this reason, the given data was acquired and analysed semi-annual to consider temporal changes in farm management.

### Study collective

Initially, a total of 205 commercial finishing pig herds was included in the study, after the farmers were asked for their written declaration of consent to participate voluntarily. As some farmers have withdrawn their consent while project duration and some data were supplied incompletely, ultimately, 184 pig units remained for further analyses. The farms of the study collective were located in the northeast of Lower Saxony, Germany, and had a mean size of 1132 ( $\pm$  555) finishing pig places (FP). Most of the farmers constantly purchased pigs that were bred in Germany (78.8%), whereas the others were steadily of Danish (4.9%) or Dutch (3.3%) origin or varied over the time (13.0%). The initial body weight (IBW) was between 6.5 kg and 47.6 kg ( $29 \pm 5.4$  kg) and live weight at slaughtering was on an average 122.6 ( $\pm$  2.8) kg. As IBW is based on farmers' management decisions, some performance criteria were impacted by the IBW, e.g. mortality, average daily gain, feed conversion ratio or antibiotic usage (Fig. 2). Hence, we decided to stratify further analyses by three different classes of piglet IBW:

- (1) light: < 24 kg
- (2) medium: 24–33,5 kg
- (3) heavy: > 33,5 kg

A t-test was performed to check statistical significance of the difference between the three classes in important health indicators. Previously, F-test for homogeneity of variances determined whether normal t-test or Satterthwaite-t-test should be used. A *p*-value below 0.05 in F-test indicates heterogeneous variances and leads to Satterthwaite-t-test.

### Data sources

For health monitoring of finishing pigs, three different data sources from several stages of the supply chain were selected:

1. **Farm specific production data (PD)**, including variables of biological and economic performance, i.e., mortality, average daily gain or feed conversion ratio. This information came from routine farm visits that were conducted semi-annually by VzF in the context of advising service. The documentation of the number of farm animals that died or were culled is also regulated by law in Germany [11].

2. **Application and delivery forms of antibiotic usage (AB)**, including information about the number of treated animals, duration of treatment and number of active substances. These data were also acquired via QS within the scope of the antibiotics monitoring programme. In Germany, the documentation of antimicrobial therapy in farmed animals is laid down by legislation in the 16th amendment of the German Medicinal Products Act [12].
3. **Diagnostic data from slaughter (SL)**, containing results of official meat inspection at the abattoirs. These data were provided by QS and came from a specific scheme that looks for 13 carcass and organ lesions of pigs at slaughter [13]. The classic findings (pneumonia, pleurisy, pericarditis and liver milk spots) were supplemented by new ones, initially on a voluntary basis, in 2016 and made mandatory from 2018 [14]. Routine post-mortem meat inspection in Germany is carried out according to EU regulation [15].

All data sets were related to a specific time-span of each half-year from July 1, 2017 to June 30, 2019. Hence, we investigated four data sets for each data source. VzF collected the raw data and incorporated it into the company-owned data processing software “VzF:professional”. VzF and MSG conducted first analysis and editing steps as well as plausibility checks by reinvestigation of missing information and implausible values. Furthermore, new variables were established to build up the “MulTiViS dataset”. These data were transmitted to TiHo, which merged the datasets and checked for integration and statistical plausibility (Fig. 1).

#### Definition of health scores

To assess the health status of pig herds, indicators from the animal-based measures were defined in the given data sets (Table 1), regarding German animal welfare discussions [6, 17, 18] and experiences of the project team. To use those indicators for health monitoring, they were transformed to specific scores that allow for distinguishing between different farms and time-periods. For this, the general approach proposed by Nienhaus et al. [16] was adapted. This method follows normalisation of skewed data via ordinary logarithm to a base of ten and a logit-transformation of prevalence data, respectively, to harmonise the information. Previously, 0-values were set to the half of the lowest value to avoid mathematical errors caused by division by zero. Afterwards, z-standardisation, i.e.,  $z_i = [(x_i - \bar{x})/s]$  puts all variables on an equal scale with a mean of zero and a standard deviation of 1. Following this, negative values generally indicate an over-average good health status

and positive values generally represent a health status inferior to the mean.

From PD, we selected mortality, average daily gain and feed conversion ratio as health indicators and transformed them to the scores MOR, ADG and FCR, respectively. To ensure the same general interpretation, the ADG score was multiplied by  $-1$ , because in contrast to the other indicators, high values are representing good performance.

To measure the use of antibiotics within a pig unit, a herd-specific treatment frequency was calculated by:

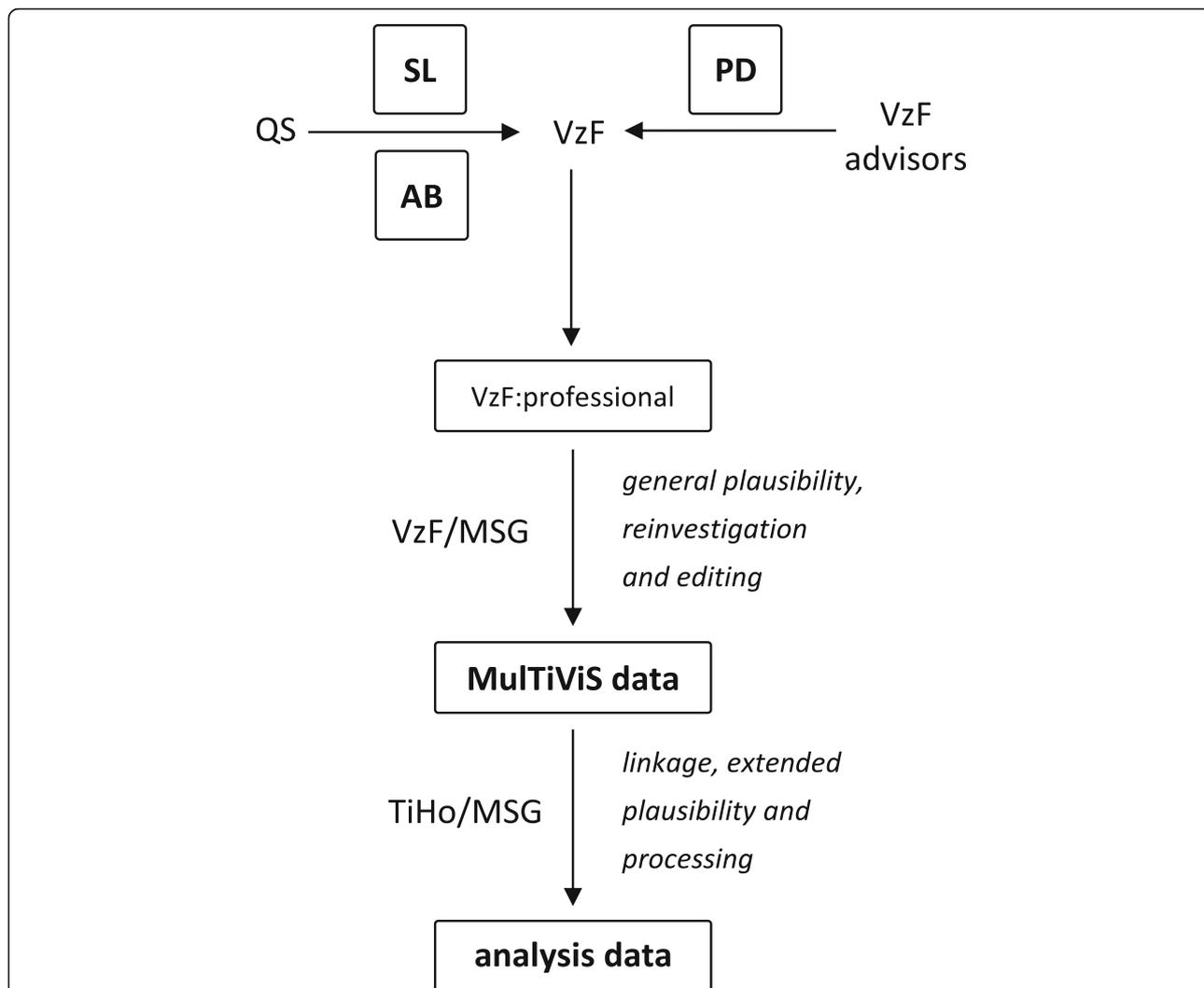
$$\text{Treatment frequency} = \frac{nUDD}{\#FP}$$

with nUDD = number of used daily doses = number of animals treated  $\times$  number of days treated  $\times$  number of active ingredients and # FP = number of finishing pig places. This definition is in line with the definition of the German Medicinal Products Act and recent scientific analyses [19, 20]. A transformation of the treatment frequency via natural logarithm and z-standardization led to TF score.

As SL-indicators, 13 assessments of carcass and organs from the QS meat inspection scheme (Table 1) were chosen. Although, since 2018, all slaughterhouses participating at the QS scheme work according to specific standards [13], it is known that there are still differences in the prevalence levels of recorded pathogenic lesions between the abattoirs [8, 21]. Hence, we defined a specific correction factor  $F_k$  for each abattoir  $k$  that compensates these differences. At the slaughterhouses, lesions were recorded for every single pig, which yields to 11917 different abattoir-herd-date combinations (batches). Batches of less than ten pigs were excluded. Furthermore, some farmers served more than one abattoir in the chosen six-month period. Hence, an adjusted herd-specific prevalence  $P_i$  for each pig unit  $i$  was defined as shown in the following formula:

$$P_i = \sum_k \left( \frac{n_{ik}}{\sum n_{ik}} \times P_{ik} \times F_k \right)$$

with  $P_i$  = herd-specific prevalence of one indicator in half a year,  $n_{ik}$  = number of animals from pig unit  $i$  to abattoir  $k$  and  $P_{ik}$  = prevalence of pig unit  $i$  at abattoir  $k$ . Because the grades of pneumonia and pleurisy are documented in four categories, we decided to combine moderate (10–30%) and high (> 30%) alterations as positive records and merged slight (< 10%) and no alterations as negative records to report a unique prevalence. For the other SL-indicators, the original information of presence or absence were used. Following the approach of Nienhaus et al. [16], 13 SL-indicators were aggregated to scores, weighted by expert opinions (Table 1), to reduce



**Fig. 1** MultiTiViS data management procedures (*MultiTiViS* project on multivariate assessment of pig welfare, *QS* Quality Scheme for Food, *VzF* swine advisory service, *VzF:professional* processing software, *MSG* Marketing Service Gerhardy, *TiHo* University of Veterinary Medicine, Hannover, *SL* diagnostic data from slaughter, *AB* information on antibiotic usage, *PD* performance data)

complexity. A respiratory lesions score (RESP) was composed of pneumonia, pleurisy and pericarditis; an exterior lesions score (EXT) was composed of arthritis, abscess, ear lesions, tail lesions, dermal alterations and bursitis; and animal management score (MANG) was composed of liver milk spots, dermal damage, intestinal alteration and whole carcass condemnation.

**Aggregation to a total score**

Finally, the health scores were aggregated to a total score (TOTAL) that should give a rough estimate of the “average” herd health status. Nienhaus et al. [16] used expert opinions to give the single health scores specific weights. However, we do not have information about *salmonella* status, but ADG and FCR instead. Coming from this, we

both gave ADG and FCR the expert weight of *salmonella* status, which yielded the following formula:

$$TOTAL = ((5 \times MOR) + (2.5 \times ADG) + (2.5 \times FCR) + (3.5 \times TF) + (5 \times RESP) + (4 \times EXT) + (4.5 \times MANG))/27$$

To compare the method of Nienhaus et al. [16] with unweighted score aggregation, we opposed both score-rakings in a scatter plot.

**Temporal development of health scores in the study collective**

As the defined health-scores are on a z-scale, this leads to a relative benchmarking. A z-value provides information about the health status of the respective farm compared to the others in the collective. To assess the variation of the health ranking of the pig herds during

**Table 1** Chosen indicators and health scores (*MOR* mortality score, *ADG* average daily weight gain score, *FCR* feed conversion ratio score, *TF* treatment frequency score, *RESP* respiratory lesions score, *EXT* exterior lesions score, *MANG* animal management score, *UDD* used daily doses, *FP* finishing pig place) with expert weights

Score	Indicator	Unit	Description	Expert weights <sup>a</sup>
<b>MOR</b>	mortality	%	dead and culled animals	. <sub>b</sub>
<b>ADG</b>	average daily gain	g	liveweight gain per day	. <sub>b</sub>
<b>FCR</b>	feed conversion ratio	kg/kg	feed per unit of liveweight gain	. <sub>b</sub>
<b>TF</b>	treatment frequency	UDD/FP	number of used daily doses per finishing pig places	. <sub>b</sub>
<b>RESP</b>	pneumonia	%	alteration of lung > 10%	5
	pleurisy	%	alteration of pleura > 10%	5
	pericarditis	%	alteration of pericardium	4
<b>EXT</b>	arthritis	%	inflammation of joint	3.5
	abscess	%	abscess	3
	ear lesions	%	necrosis/inflammation of ear	5
	tail lesions	%	necrosis/inflammation of tail	5
	dermal alterations	%	inflammation of skin	4
	bursitis	%	bursitis with > 5 cm in diameter	2.5
<b>MANG</b>	liver milk spots	%	alteration of liver by milk spots	4.5
	dermal damage	%	alteration through punch marks	5
	intestinal alteration	%	inflammation of intestines	1
	whole carcass condemnation	%	extensive alteration of carcass	1.5

<sup>a</sup>Adopted by Nienhaus et al. [16]

<sup>b</sup>No expert weights needed because of direct implementation to z-scores

the study period, they were assigned to categories. These were based on classification of the collective in quarters, as follows:

- category 1: z-value  $\leq 0.25$ -quartile,
- category 2: 0.25-quartile < z-value  $\leq 0.5$ -quartile,
- category 3: 0.5-quartile < z-value  $\leq 0.75$ -quartile,
- category 4: z-value > 0.75-quartile.

These categories were calculated for each score, each half-year and the IBW-class, respectively.

Thereupon it was assessed how the individual farms switched between the categories over the four half-years. Four categories were defined, depending on if the herds stayed in the upper, middle or lower 50% of the collective or if higher variations occurred.

All statistical evaluations mentioned were performed with SAS<sup>®</sup>, version 9.4.

## Results

### Stratification by IBW

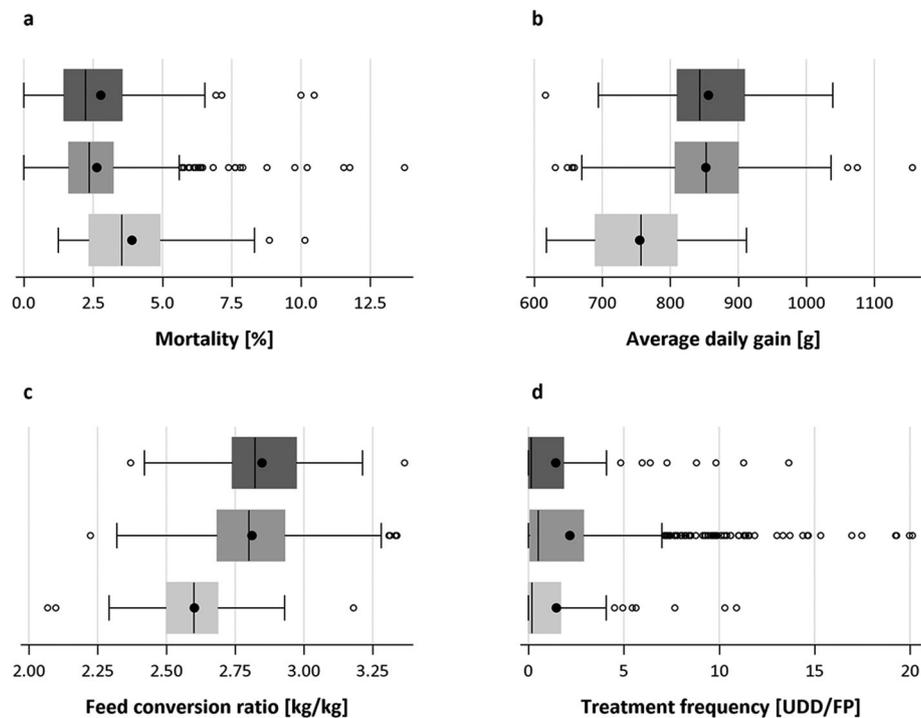
Descriptive statistical evaluations showed that the mean of important indicators varies between classes of light or heavy IBW and the medium class, respectively (Fig. 2, Table 2). The results of the t-tests proved statistical significance of higher mortality ( $p < 0.0001$ ) and lower

average daily gain ( $p < 0.0001$ ) and feed conversion ratio ( $p < 0.0001$ ) for light class against medium class. For heavy class against medium class, treatment frequency was significantly lower ( $p = 0.0367$ ).

### Transformation of original health data

In total, seven health scores (*MOR*, *ADG*, *FCR*, *TF*, *RESP*, *EXT* and *MANG*) and one total score were defined for 184 pig units and four half-years. The original data for *ADG* and *FCR* criteria appeared nearly normally distributed, and no substantial outlier was identified. In contrast, the untransformed data of *MOR*, *TF* and the *SL*-scores showed a skewed distribution. After transformation via natural logarithm to a base of ten for treatment frequency and logit-transformation for mortality and *SL*-indicators, they showed normal distribution and zero-inflation for variables with many zero values (Fig. 3).

The scatterplot comparing the two methods of calculating the total score shows in the four half-years that most pig units are at or near the diagonal and only a few outliers can be seen (Fig. 4). This means that ranking-position on z-scale does not differ substantially between the two scoring methods. Hence, we decided to adopt the expert weights from Nienhaus et al. [16].



**Fig. 2** Variation of performance between different IBW-classes (1 = light (light grey), 2 = medium (grey), 3 = heavy (dark grey)) in mortality (a), average daily gain (b), feed conversion ratio (c) and treatment frequency (d) in the study collective with  $n = 736$  (184 pig units  $\times$  4 half-years) in time-span from July 1, 2017 to June 30, 2019

### Distribution and temporal development of indicators

As most of the data were skewed, median instead of mean as descriptive parameter was used (Table 3). The following results refer to IBW-class 2, since the majority of pig units under study is in this class. Descriptive statistics for class 1 and 3 could be found in Additional file 1. The three PD indicators mortality, daily weight gain and feed conversion ratio varied very little over the time. The median mortality was between 2.26% and 2.40%, the median of average daily gain varied between 849 g and 860 g and feed conversion ratio was between 2.77 kg/kg and 2.81 kg/kg. A clear downward trend can be seen in the treatment frequency. It decreases from 0.72 UDD/FP in the first half-year to 0.46 UDD/FP in the last half-year. Descriptive measures of meat inspection data show that the variation over the four half-years was equally low for all SL-indicators. However, it is striking that classic indicators (pneumonia, pleurisy, pericarditis and liver milk spots) were found significantly more often than the new ones, which are only mandatory since 2018. Their median prevalence was below 1%. Pneumonia was found most frequently in the study period with a median prevalence between 8.52% and 10.74%, followed by pleurisy, pericarditis and liver milk spots. Ear lesions and dermal damage caused by handling

occurred in less than half of the study collective in all four half-years.

### Development of health scores over time

One phenomenon in the temporal development of health benchmarking is that the majority of the collective has a variable health status over the four half-years under study, i.e. farms vary temporarily per score in ranking position. In contrast, very few herds remained in the middle 50% of the collective (category 2 or 3) over time.

To quantify this general effect, Table 4 shows the changes in benchmark categories for each health score and the total score. RESP seems to be the most constant parameter, because more than a quarter of the pig units were always found in the upper or lower half of the collective, respectively. In contrast, FCR is the score with the greatest variation. Approximately two-thirds of the pig units change their ranking position by more than one category during the study period.

### Discussion

The study aimed to connect different data sources from daily production processes along the supply chain for the purpose of a health monitoring programme of finishing pigs. Seven health scores and one total score were

**Table 2** Means of four indicators mortality, average daily gain, feed conversion ratio and treatment frequency for different IBW-classes (1 = light, 2 = medium, 3 = heavy) and *p*-values of t-test

Indicator	Mean			P-Value t-test	
	Class 1	Class 2	Class 3	Class 1 vs. 2	Class 2 vs. 3
Mortality [%]	3.90	2.63	2.77	< <b>0.0001</b>	0.5483
Average daily gain [g]	755.2	852.3	856.1	< <b>0.0001</b>	0.6565
Feed conversion ratio [kg/kg]	2.60	2.81	2.85	< <b>0.0001</b>	0.1067
Treatment Frequency [UDD/FP]	1.46	2.17	1.43	0.0744	<b>0.0367</b>

used to classify pig units by their health status: MOR, ADG, FCR, TF, RESP, EXT and MANG.

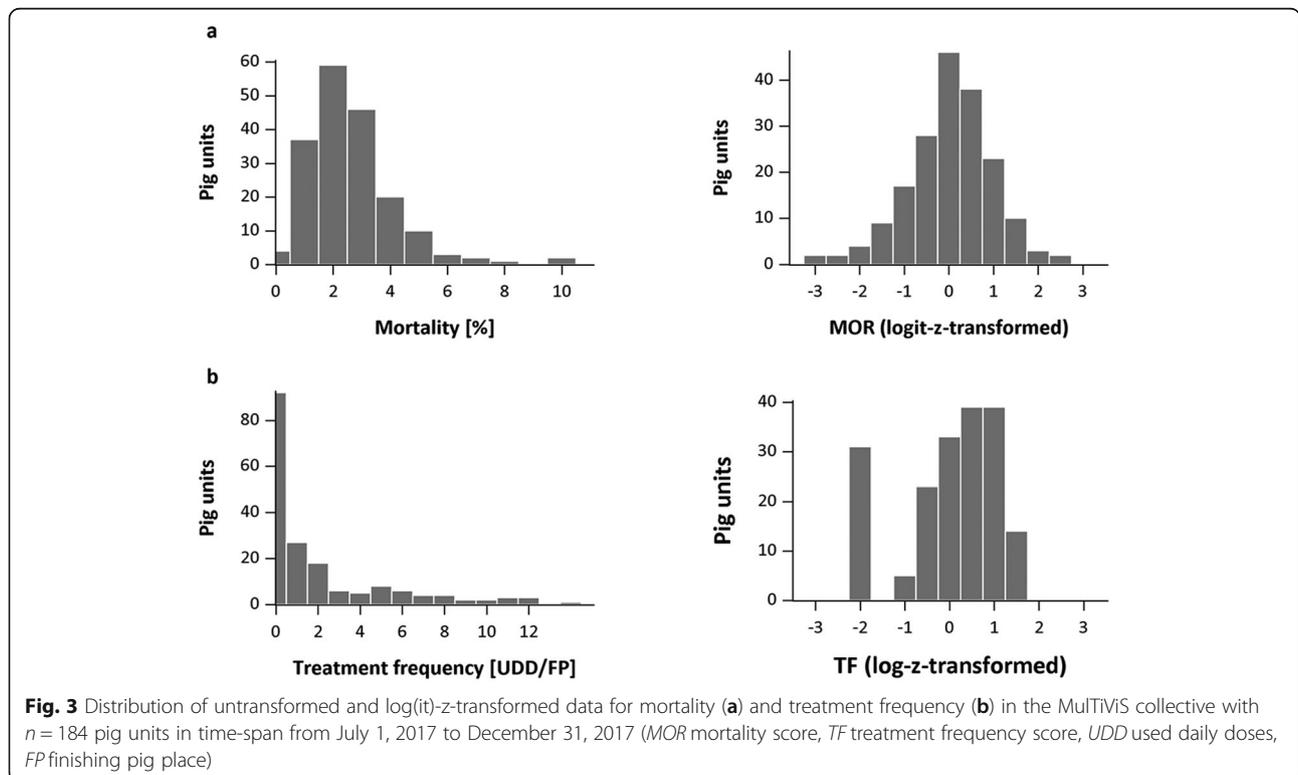
Combining seven individual scores into one total score has the advantage of providing a quick first impression of the health status. Furthermore, such a rough assessment is easier to handle for a monitoring system. However, such an overall score cannot indicate in which areas the individual pig units have potential for improvement. For this reason, it is imperative that the individual health scores are consulted as part of veterinary and agricultural advice.

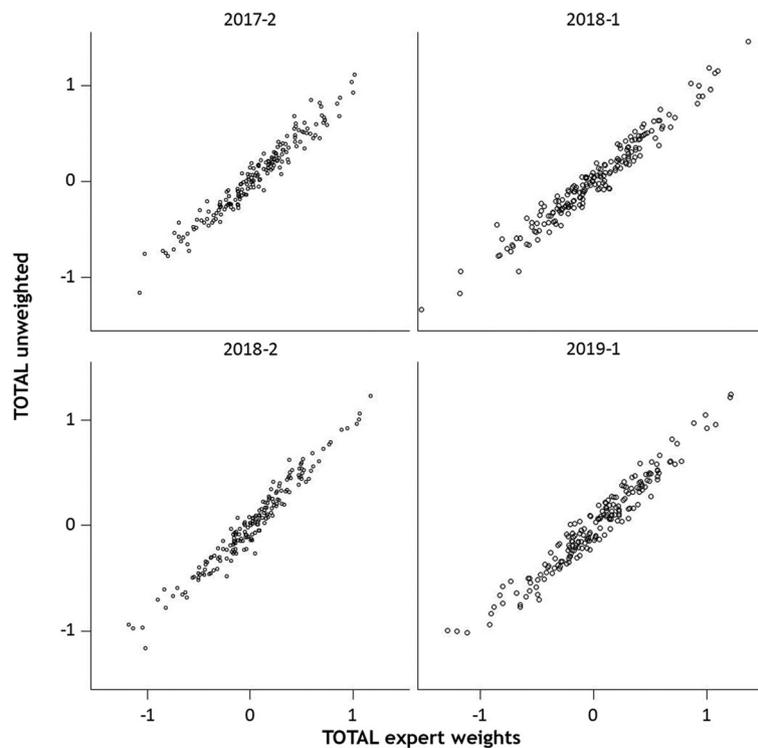
**Sampled collective of pig units**

Because the study is part of the process aiming to find methods for a national health monitoring programme, a high number of approximately 200 participating pig units was needed. However, because the collective of

finishing pig herds in the study is based on both membership of VzF and willingness to participate, it may not be free of a selection bias.

Moreover, the location of the pig units has to be taken into account, as there are great regional differences concerning farm sizes and density of pig husbandry in Germany [22]. All participating herds were located in the northeast of Lower Saxony, which is a region with an extended density of pig husbandry compared with other German regions [23]. In November 2017, 4.3 million finishing pigs on 5100 farms were kept in that region [23]. Since the VzF population was not selected randomly and is located only in Lower Saxony, the results of the study are not purely representative of the population of finishing pig farms in Germany. However, due to VzF as a consultant for typical pig farming in Germany, the sample is seen as a foundation for a study





**Fig. 4** Comparison of ranking for total score (TOTAL) in four half-years, calculated with expert-weights and unweighted

**Table 3** Median (P50) and Interquartile range (IQR) of indicators in four half-years from July 1, 2017 to June 30, 2019 for IBW-class 2 (n = 610, UDD used daily doses, FP finishing pig place)

Indicator	2017-2		2018-1		2018-2		2019-1	
	P50	IQR	P50	IQR	P50	IQR	P50	IQR
Mortality [%]	2.38	1.60	2.26	1.45	2.37	1.59	2.40	1.75
Average daily gain [g]	856	85	850	95	852	69	860	105
Feed conversion ratio [kg/kg]	2.81	0.20	2.80	0.26	2.81	0.24	2.77	0.30
Treatment frequency [UDD/FP]	0.72	2.90	0.49	2.92	0.55	3.12	0.46	2.61
Pneumonia [%]	10.74	8.91	10.34	9.26	8.52	8.90	9.64	11.33
Pleurisy [%]	4.54	7.44	4.15	8.58	3.74	5.52	4.50	8.99
Pericarditis [%]	3.00	2.26	2.97	2.76	3.15	2.69	3.92	2.93
Arthritis [%]	0.55	0.65	0.50	0.60	0.66	0.86	0.40	0.61
Abscess [%]	0.87	0.67	0.79	0.62	0.92	0.82	0.90	0.82
Ear lesions [%]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04
Tail lesions [%]	0.42	0.79	0.37	0.73	0.54	1.03	0.57	1.03
Dermal alterations [%]	0.07	0.20	0.12	0.28	0.15	0.22	0.08	0.20
Bursitis [%]	0.45	0.69	0.29	0.51	0.60	0.72	0.58	0.87
Liver milk spots [%]	3.89	6.08	3.40	4.62	3.51	5.79	3.16	5.28
Dermal damage [%]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Intestinal alteration [%]	0.47	0.58	0.45	0.64	0.38	0.59	0.27	0.62
Whole carcass condemnation [%]	0.09	0.20	0.13	0.25	0.11	0.23	0.08	0.16

design for a national monitoring programme because it provides insight into the general suitability of secondary data usage for welfare scoring.

#### Stratification by IBW-classes

The IBW of pigs when installed has impact on several health parameters. Results of t-test showed that for light class-1-pig herds mortality is higher and average daily gain is lower than in class-2 herds. One possible explanation is that younger piglets are more susceptible to health problems because their immune system is not yet fully developed and that they are reared for a longer period, which leads to higher mortality. Furthermore, younger piglets have relatively better feed conversion ratio than older ones. However, the growth curve of weight gain per kilogram live weight has its peak in the later phase of the finishing period [24]. This is a possible explanation for lower average daily weight gain of pig units in class 1, as the calculation of daily gain takes into account the complete finishing period and therefore, in the case of class-1 herds, also the lower values from the initial period.

The restricted usage of antibiotics in class-3 herds could be caused by the fact, that the piglets are stalled in with higher age because they had a pre finishing phase and therefore a shorter time that is depicted in the study data. These animals have usually overcome the typical “childhood diseases” and therefore need less antibiotic treatment.

#### Indicators for health scoring

This study concentrated on health- and performance-related items as they can be easily quantified and are often harmonised and captured within daily processes along the supply chain. The selection of specific health indicators was based on expertise of the project teams and results of various national and international studies that tried to quantify pig health and welfare.

In Germany, the “Kuratorium für Technik und Bauwesen in der Landwirtschaft e.V.” (KTBL) conducted an expert survey to define welfare indicators with high validity and reliability under the usual conditions of the German production system. To keep the economic effort manageable, they preferred indicators built from existing databases [25]. For the same reason, we took the KTBL indicators mortality, average daily weight gain, feed conversion ratio, treatment frequency and assessments of meat inspection into account [6]. KTBL outlines that these data could be a hint for possible welfare problems [25].

In 2009, Dickhaus et al. [18] tried to quantify the health status of pig herds with a herd health score (HHS), which includes information about the mortality rate, the frequency of pathological assessments in carcasses and organs, an animal-treatment-index and the duration of the finishing period. It was noted that the combination of this information is suitable for benchmarking systems as a tool for improving animal health and welfare. With the exception of the fourth point (duration of finishing period), these variables are the same as those selected for our study. Due to the lack of information concerning the duration of the finishing period, we decided to use average daily gain and feed conversion ratio instead.

Pandolfi et al. [26] merged different data sources to assess the interconnections between biosecurity, health, welfare and performance in commercial pig farms in Great Britain. They had performance data (average daily gain, mortality and feed conversion ratio) from farm visits, welfare information from the Real Welfare programme and lesions recorded at the abattoir from the British Pig Health Scheme. Additionally, they collected information about external and internal biosecurity during farm visits to examine the association between biosecurity and welfare outcomes. The Real Welfare Scheme is a programme of the Agricultural and

**Table 4** Categorisation of changes of the study collective (n = 184) for seven health scores (*MOR* mortality score, *ADG* average daily weight gain score, *FCR* feed conversion ratio score, *TF* treatment frequency score, *RESP* respiratory lesions score, *EXT* exterior lesions score, *MANG* animal management score) and one total score (TOTAL)

	Always category 1 or 2 %	Always category 2 or 3 %	Always category 3 or 4 %	Variable health status %
<b>MOR</b>	24.46	11.96	22.28	41.30
<b>ADG</b>	22.83	11.41	20.65	45.11
<b>FCR</b>	13.59	7.61	11.41	67.39
<b>TF</b>	20.65	14.13	19.02	46.20
<b>RESP</b>	28.80	11.41	26.63	33.15
<b>EXT</b>	14.67	10.33	14.67	60.33
<b>MANG</b>	23.37	13.59	22.83	40.22
<b>TOTAL</b>	21.20	12.50	21.20	45.11

Horticultural Development Board (AHDB) within the frame of quality assurance (Red Tractor) and covers 95% of the production [27]. It contains welfare indicators that are derived from on-farm data related to the prevalence of pigs that would benefit from removal to hospital pens, lame pigs, pigs with tail damage, pigs with body marks and environmental enrichment provision and use [27].

Overall, it may be stated that the variables chosen for scoring are in line with the scientific literature and could be collected in daily processes. However, as the data were not originally gathered for the purpose of health and welfare monitoring, secondary data use has to be addressed in more detail under the conditions of the German production system.

### **Secondary data analysis**

As this investigation is an observational study, it was targeted to examine the usability of already existing information sources for monitoring, and only secondary data drawn from processes in the supply chain was used. The data were collected within the scope of consulting (PD) and quality assurance (AB, SL) but not with the aim of an interdependent analysis or to use it for combined health monitoring.

A striking advantage of secondary data usage is the lack of effort needed to acquire the data. Therefore, it requires fewer economic and personnel resources without any additional burden for the farmers.

However, the use of secondary data is not straightforward as they are collected for another purpose. The data need to be cleaned and checked for both integration and plausibility, which requires extra work. Due to their internal and external plausibility as well as association structures, not all variables can be used. Implausible and missing values might occur. On the one hand, these processes are linked to the data source, and on the other hand, they are connected to the purpose of the analyses, i.e., in our case, the scoring of animal health [28]. Therefore, a source-by-source discussion is necessary.

### **Biological and economic performance data**

In contrast to AB and SL, for which strict documentation is mandatory by law, general performance data (PD) on herd level are not standardised on a national level in Germany. Rules for a harmonisation of these data are rare. Only the documentation of mortality is required by law. However, such data collection is requested or already implemented by various instances, for example, “vit” [29], that offers a harmonised data approach for its members or, much stricter, the chamber of veterinarians, that demands a standardised, routine documentation of performance data [30]. However, pig units connected to consulting services such as VzF work with their own standards based on the suggestions mentioned above.

For MulTiViS, PD was constructed from the internal VzF database, which is merged from data, the pig units documented due to different regulations as well as from different optional data during the production process. For this, agricultural advisors collected basic claims data of the pig units as well as variable data of the production process. The latter was aggregated by time as well as between the pig units' compartments. For categorical data, if expressions vary in different compartments within the pig unit, VzF has defined the subclasses according to the majority principle. For continuous data, usual averages, weighted by the number of animals, were used. Therefore, on the one hand, these data are prone to an information bias, especially if the pig units hold a huge variety of different compartments. On the other hand, continuous data about mortality, average daily gain and feed conversion ratio are usually under strict farmer control, which generally avoids serious bias.

### **Information on antibiotic usage**

Since their collection is related to the mandatory documentation in the German Medicinal Products Act, register data on antibiotic usage are available for all finishing pig units of a certain size. In addition, AB data from pig units participating in the QS quality programme, are available for pig units of all sizes. Because the QS system covers 95% of the entire German pork production system, the information on AB is close to covering the general target population [10].

Despite this advantage, the QS system has pitfalls. First, information about the indication of antibiotic treatment is usually lacking. Therefore, a direct link to animal health aspects is not possible. Furthermore, bias may occur from the veterinary practices because the methods and intensity of antibiotic treatments differ. Another problem is information bias of AB as these data are generally used for farm consulting. Therefore, plausibility checks were only performed if the treatment frequency is above specific cut-off values in the benchmarking system. Furthermore, plausibility checks were only executed for general amounts but not for substances in detail.

The general usability of AB as health indicator should also be critically reviewed, as the association of extended antimicrobial treatment with diseased animals (poor health status) is often broken twice. On one hand, (high) antibiotic usage could also be associated with good animal health because it is used for the purpose of curing. On the other hand, farmers may avoid antibiotic therapy to keep from being benchmarked but that could cause poor animal health.

### **Meat inspection data**

The general usability of slaughterhouse data for animal health and welfare purposes in Germany is well

discussed [21, 31]. Although the gathering of these data is formally harmonised, in daily practice, there is a lack of standardisation between abattoirs and inspectors. This leads to differences of the prevalence levels on the temporal scale as well as between slaughterhouses. Therefore, only some studies found a relationship between respiratory lesions at the abattoir and the respiratory health and performance of living pigs [32, 33]. To compensate for this phenomenon, we have made a specific correction to the individual prevalence data.

As the collection of findings according to the present QS scheme has only been established mandatory since 2018, the reporting of the newly added lesions is therefore correspondingly non-standardised. This is manifested by a prevalence of below 1% and a skewed distribution which hence limits the statistical usability.

However, information of meat inspection from QS is the first nationwide data collection in Germany. It covers approximately 95% of the entire German pork production system, which corresponds to information of approximately 30000 pig units and 200 abattoirs in one system. For that reason, the use of these data has to be stated as crucial.

#### Relative benchmarking of health scores

Because the original data are on different scales, the indicators were standardised via z-transformation to health scores, following the approach from Nienhaus et al. [16]. That implies each pig unit is assigned a certain rank, reflecting its relative position within the distribution of the study collective. If the collective of pig units was changed, i.e. between different time-spans, a shift of the individual z-value would be possible. Hence, health-scores based on z-values only allow conclusions to be drawn about the health status of a pig herd in comparison to the other herds under study. However, there is always an incentive for improvement if a pig unit has a high z-value (which stands for relative poor health status) or if the z-value is worse than in the previous half-year. Either the own health status declined or the average performance in the collective improved. Another advantage of such relative scoring method is that no absolute threshold needs to be defined and that trends over time can be offset.

This strategy is in line with other benchmarking systems. As an example, the monitoring programme of antibiotics due to the 16th amendment of the German Medicinal Products Act is using the 50%- and the 75%-percentile of the entire distribution of pig units as benchmarks. Twice a year, these measures are calculated, so the critical thresholds change every 6 months.

#### Temporal development of health status

All the indicators examined except the TF showed only moderate fluctuation over time. For the TF, however, a

clear downward trend could be identified. The increasing problem of antimicrobial resistances has led to a strong focus on the use of antibiotics in livestock production. Both nationally and internationally, there are calls for a reduction in the use of antibiotics in farm animals [34–36]. This is reflected in the decreasing TF in the study farms.

Another temporal factor is seasonality, which has an influence on many aspects of health, e.g. mortality or feed conversion ratio [37, 38]. However, as the available data in the study was collected on a half-year basis and include parts of the warm and cold seasons each, rough trends can be balanced out. This can be seen in Table 3, where is shown that there are no strong fluctuations between the half-years.

For relative benchmarking with health scores, it can be stated that changing of benchmark categories may be attributed to two different causes: firstly, the health status of a certain farm improved or declined; secondly, the health status remained the same but the other pig herds in the collective performed better or worse.

The results from Table 4 showed, that most pig units have varied strongly in their benchmark position. This again underlines the advantage of a relative benchmarking system, which takes account of such fluctuations.

It can also be noted that the pig herds that did not fluctuate that much were found either in the upper or the lower half of the collective. Constantly moderate farms, that were steadily in category 2 or 3 over time, occurred very seldom. This suggests that the health status of averagely performing farms usually changes in one direction or another. Reasons for this could be, i.e., veterinary advice or special events such as a vaccination breakthrough or disease outbreak.

Monitoring the scores over time can provide an indication of whether an individual farm is experiencing sustained animal health problems. If, for example, it remains in the worst category over all four half-years, it either has long-term difficulties with health management or it cannot keep up with the upward trend of the other farms. Both, however, highlight the need for sustainable veterinary or agricultural advice. If, on the other hand, a pig unit often changes category, this is probably more likely to be due to constant fluctuations in the collective or short-term changes or spontaneous events in health management.

#### Conclusion

The study proves that routine data from existing databases of the German pork production chain are usable for health monitoring. However, it has to be stated, that preparatory editing steps are crucial. Furthermore, we suggest stratification by IBW of piglet, because this has impact on important health indicators. A total score

could give a rough estimate of a herd health status but should be complemented by specific health scores in the frame of veterinary and agricultural advising. It was also shown that relative benchmarking can be used to depict fluctuations in the composition of the target collective and temporal conditions, without having to define absolute thresholds. Additionally, it is also a constant incentive for individual farms to improve their health status. Evaluating the health scores over time can also show which farms have sustainable improvement potential.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40813-021-00197-z>.

**Additional file 1.** Descriptive statistics of indicators for IBW-classes 1, 2 and 3.

## Abbreviations

ADG: Average daily weight gain score; EXT: Exterior lesions score; FCR: Feed conversion ratio score; FP: Finishing pig place; IBW: Initial body weight; IQR: Interquartile range; MANG: Animal management score; MOR: Mortality score; nUDD: Number of used daily doses; PD: Performance data; QS: Quality Scheme for Food; RESP: Respiratory lesions score; SL: Diagnostic data from slaughter; TF: Treatment frequency score

## Acknowledgements

This work was conducted under the MultIVIS project. The MultIVIS consortium would like to thank all participating farmers for their cooperation.

## Authors' contributions

LK, HG and HP designed the study. HP and HM were responsible for raw data acquisition. HG, HP and HM edited the raw data, created the MultIVIS data and prepared it for transfer to JGK and LK. JGK performed further processing steps and analyses of the data with support of CH and HG. JGK wrote the initial draft. LK and HG critically reviewed the manuscript. All authors have read and approved the manuscript.

## Funding

The project is supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support programme with the number 2817905315. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. Open Access funding enabled and organized by Projekt DEAL.

## Availability of data and materials

The data were collected on an individual basis from farmers and slaughterhouses. Each participant gave written consent with the understanding that data would not be transferred to a third party. Therefore, any data transfer to interested persons is not allowed without an additional formal contract. Data are available to qualified researchers who sign a contract with the University of Veterinary Medicine Hannover and VzF. This contract will include guarantees to the obligation to maintain data confidentiality in accordance with the provisions of the German data protection law. Currently, there exists no data access committee or another body that could be contacted for the data. However, for this purpose, a committee will be founded. This future committee will consist of the authors as well as members nominated by the University of Veterinary Medicine Hannover and the VzF. Interested cooperative partners, who are able to sign a contract as described above, may contact: Prof. Dr. Lothar Kreienbrock, Department of Biometry, Epidemiology and Information Processing, University of Veterinary Medicine, Hannover, Buenteweg 2, 30559 Hannover, Email: [lothar.kreienbrock@tiho-hannover.de](mailto:lothar.kreienbrock@tiho-hannover.de)

## Ethics approval and consent to participate

The data used within this study are based on mandatory application and delivery forms, data from official slaughterhouse inspection and performance data accumulated and maintained by VzF for consulting advice for farmers. All information was provided voluntarily by farmers, signing individual written consent data to be used by the study team only. The research does not involve any regulated animals, and there were no scientific procedures performed on animals of any kind. For this reason, formal approval by an ethical committee was not necessary under the provisions of the German regulations.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>Department of Biometry, Epidemiology and Information Processing, WHO Collaborating Centre for Research and Training for Health in the Human-Animal-Environment Interface, University of Veterinary Medicine Hannover, Foundation, Buenteweg 2, 30559 Hanover, Germany. <sup>2</sup>VzF e.V., Association for Promoting Farming Economics, Uelzen, Germany. <sup>3</sup>MSG, Marketing Service Gerhardy, Garbsen, Germany.

Received: 14 December 2020 Accepted: 28 January 2021

Published online: 24 February 2021

## References

- Starosta S, Bergschmidt A. Animal Welfare Reporting in der EU – (was) kann Deutschland von seinen Nachbarn lernen? *Appl Agric Forestry Res.* 2015; 65(1):47–58.
- Broom DM. Animal welfare: an aspect of care, sustainability, and food quality required by the public. *Anim Welfare Educ Res.* 2010;37(1):83–8.
- German Federal Ministry of Justice and Consumer Protection. Tierschutzgesetz in der Fassung der Bekanntmachung vom 18. Mai 2006 (BGBl. I S. 1206, 1313), das zuletzt durch Artikel 1 des Gesetzes vom 17. Dezember 2018 (BGBl. I S. 2586) geändert worden ist. 2018.
- Initiative Tierwohl. <https://initiative-tierwohl.de/>. Accessed 22 Oct 2020.
- Welfare Quality® Consortium. Welfare Quality® assessment protocol for pigs (sows and piglets, growing and finishing pigs). 2009.
- Kuratorium für Technik und Bauwesen in der Landwirtschaft e. V (KTBL). Tierschutzindikatoren: Leitfaden für die Praxis - Schwein. KTBL-Praktikerleitfaden „Tierschutzindikatoren – Schwein“. 2016. p. 37–51.
- Pandolfi F, Kyriazakis I, Stoddart K, Wainwright N, Edwards SA. The 'Real Welfare' scheme: identification of risk and protective factors for welfare outcomes in commercial pig farms in the UK. *Prev Vet Med.* 2017;146:34–43.
- EFSA Panel on Biological Hazards (BIOHAZ). Scientific Opinion on the public health hazards to be covered by inspection of meat (swine). *EFSA J.* 2011; 9(10):2351.
- Brambell FWR. Report of the technical committee to enquire into the welfare of animals kept under intensive livestock husbandry systems. London: Her Majesty's Stationery Office; 1965.
- Qualität QS, Sicherheit GH. Zum Hofe; 2019.
- German Government. Tierschutz-Nutztierhaltungsverordnung in der Fassung der Bekanntmachung vom 22. August 2006 (BGBl. I S. 2043), die zuletzt durch Artikel 3 Absatz 2 des Gesetzes vom 30. Juni 2017 (BGBl. I S. 2147) geändert worden ist. 2006.
- German Federal Parliament. 16. Gesetz zur Änderung zur Änderung des Arzneimittelgesetzes. *Bundesgesetzblatt Teil I Nr. 62. § 58b.* 2013.
- Qualität QS, Sicherheit GH. Leitfaden "Befunddaten in der Schweineschlachtung"; 2018.
- QS Qualität und Sicherheit GmbH. Recording of diagnostic data - Data recording and reporting for pigs adapted to practice. *QS Report Meat and Meat Products.* 2018;01/2018 p. 3.
- Commission delegated regulation (EU) of 15 March 2019 Laying down uniform practical arrangements for the performance of official controls on products of animal origin intended for human consumption in accordance with regulation (EU) 2017/625 of the European Parliament and of the council and amending commission regulation (EC) no 2074/2005 as regards official controls. 2019; L 131:51–100.

16. Nienhaus F, Meemken D, Schoneberg C, Hartmann M, Kornhoff T, May T, et al. Health scores for farmed animals: Screening pig health with register data from public and private databases. *Plos One*. 2020;15(2).
17. Böckel V. Untersuchungen zur quantitativen Bewertung der Tiergesundheit in Schweinebeständen. Field Station for Epidemiology, Bakum; University of Veterinary Medicine, Foundation, Hannover, Germany. 2008.
18. Dickhaus CP, Meemken D, Blaha T. Attempts to quantify the health status of pig herds: developing and validating a herd health score (HHS). *Sustainable Animal Production: the Challenges and Potential Developments for Professional Farming*, Wageningen Academic Publishers, The Netherlands. 2009; pp. 191–201.
19. Hemme M, Ruddat I, Hartmann M, Werner N, van Rennings L, Käsbohrer A, et al. Antibiotic use on German pig farms - A longitudinal analysis for 2011, 2013 and 2014. *Plos One* 2018;13(7).
20. Kasabova S, Hartmann M, Werner N, Kaesbohrer A, Kreienbrock L. Used Daily Dose vs. Defined Daily Dose - Contrasting Two Different Methods to Measure Antibiotic Consumption at the Farm Level. *Front Vet Sci*. 2019;6:116.
21. Hoischen-Tauber S, Blaha T, Werner C, Sundrum A. Zur Reproduzierbarkeit der Befunderfassung am Schlachthof für Merkmale der Tiergesundheit. *Arch Leb*. 2011;62(3):82–7.
22. Merle R, Busse M, Rechter G, Meer U. Regionalisation of Germany by data of agricultural structures. *Berl Munch Tierarztl Wochenschr*. 2012;8(1–2):52–9.
23. Statistisches Bundesamt (Destatis). Land- und Forstwirtschaft, Fischerei. Fachserie 3, Reihe 4 - Viehbestand. 2017.
24. Wellock IJ, Emmans GC, Kyriazakis I. Describing and predicting potential growth in the pig. *Anim Sci*. 2004;78:379–88.
25. Zapf R. Eigenkontrolle Tierwohl. Kuratorium für Technik und Bauwesen in der Landwirtschaft eV (KTBL); 2016.
26. Pandolfi F, Edwards SA, Maes D, Kyriazakis I. Connecting different data sources to assess the interconnections between biosecurity, health, welfare, and performance in commercial pig farms in Great Britain. *Front Vet Sci*. 2018;5:41.
27. Agriculture and Horticulture Development Board (AHDB). Real Welfare Baseline Report: 2013–2016. 2017.
28. Swart E, Ihle P, Gothe H. Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven: Verlag Hans Huber, Bern; 2014.
29. Vereinigte Informationssysteme Tierhaltung w.V. (vit). <https://www.vit.de/>. Accessed 22 Oct 2020.
30. Tiedemann U. Bericht zur Lage anlässlich des 28. Deutschen Tierärzttags in Dresden. *Deutsches Tierärzteblatt*. 2018;10:1380–8.
31. German Government. Antwort auf die Anfrage Bündnis 90/DIE GRÜNEN: Monitoring-Daten zum Tierwohl aus dem Nutztierbereich. 2017.
32. Holt HR, Alarcon P, Velasova M, Pfeiffer DU, Wieland B. BPEX pig health scheme: a useful monitoring system for respiratory disease control in pig farms? *BMC Vet Res*. 2011;7(1):82.
33. Brewster VR, Maiti HC, Tucker AW, Nevel A. Associations between EP-like lesions and pleuritis and post trimming carcass weights of finishing pigs in England. *Livest Sci*. 2017;201:1–4.
34. Wiegel B. DART: deutsche Antibiotika-Resistenzstrategie. Bericht zu Vorgeschichte und Sachstand/DART: the German strategy to fight antimicrobial resistance. Report on history and current state. *J Lab Med*. 2011;35(4):185–94.
35. WHO. Averting the AMR crisis. <https://www.euro.who.int/en/about-us/partners/observatory/publications/policy-briefs-and-summaries/averting-the-amr-crisis>. Accessed 4 Nov 2020.
36. FAO. Antimicrobial resistance. <http://www.fao.org/antimicrobial-resistance/en/>. Accessed 4 Nov 2020. Accessed.
37. Agostini PS, Fahey AG, Manzanilla EG, O'Doherty JV, De Blas C, Gasa J. Management factors affecting mortality, feed intake and feed conversion ratio of grow-finishing pigs. *Animal*. 2014;8(8):1312.
38. Maes DGD, Duchateau L, Larriestra A, Deen J, Morrison RB, de Kruijf A. Risk factors for mortality in grow-finishing pigs in Belgium. *J Veterinary Med Ser B*. 2004;51(7):321–6.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)



## 2.2 Health Monitoring of Fattening Pigs – Use of Production Data, Farm Characteristics and On-Farm Examination

Große-Kleimann J, Wegner B, Spiekermeier I, grosse Beilage E, Kemper N, Nienhoff H, Plate H, Meyer H, Gerhardy H, Kreienbrock L. Health Monitoring of Fattening Pigs – Use of Production Data, Farm Characteristics and On-Farm Examination. *Porcine Health Management* 2021; 7(45) DOI 10.1186/s40813-021-00197-z.

**Journal** Porcine Health Management

**Editor** Prof Paolo Martelli (Università di Parma, Italy)

Prof Joaquim Segalés (Universitat Autònoma de Barcelona & Centre de Recerca en Sanitat Animal, Spain)

**Eingereicht** 12. April 2021

**Akzeptiert** 16. Juli 2021

**Veröffentlicht** 3. August 2021

RESEARCH

Open Access



# Health Monitoring of Fattening Pigs – Use of Production Data, Farm Characteristics and On-Farm Examination

Julia Grosse-Kleimann<sup>1\*</sup>, Birte Wegner<sup>2</sup>, Ines Spiekermeier<sup>3</sup>, Elisabeth grosse Beilage<sup>4</sup>, Nicole Kemper<sup>2</sup>, Hendrik Nienhoff<sup>2</sup>, Heiko Plate<sup>5</sup>, Henning Meyer<sup>5</sup>, Hubert Gerhardt<sup>6</sup> and Lothar Kreienbrock<sup>1</sup>

## Abstract

**Background:** The use of processed secondary data for health monitoring of fattening pigs has been established in various areas, such as the use of antibiotics or in the context of meat inspection. Standardized scores were calculated based on several sources of production data and can be used to describe animal health in a large collective of pig units. In the present study, the extent to which these scores are related to different farm characteristics and management decisions were investigated. In addition, slaughter scores were compared with the results of a veterinary examination on the farms.

**Results:** The comparison of the results of the uni- and multifactorial analyses revealed that almost all of the examined factors play a role in at least one of the scores when considered individually. However, when various significant influencing factors were taken into account at any one time, most of the variables lost their statistical significance due to confounding effects. In particular, production data such as production costs or daily feed intake remained in the final models of the scores on mortality, average daily gain and external lesions. Regarding the second part of the investigation, a basic technical correlation between the slaughter scores and the on-farm indicators could be established via principal component analysis. The modelling of the slaughter scores by the on-farm indicators showed that the score on external lesions could be represented by equivalent variables recorded on the farm (e.g., lesions caused by tail or ear biting).

**Conclusions:** It has been demonstrated that the examined health scores are influenced by various farm and management characteristics. However, when several factors are taken into account, confounding occurs in some cases, which must be considered by consultants. Additionally, it was shown that on-farm examination content is related to the scores based on equivalent findings from slaughter pigs.

**Keywords:** farm management, secondary data use, welfare indicators, swine

\* Correspondence: [julia.grosse-kleimann@tiho-hannover.de](mailto:julia.grosse-kleimann@tiho-hannover.de)

<sup>1</sup>Department of Biometry, Epidemiology and Information Processing, WHO Collaborating Centre for Research and Training for Health in the Human-Animal-Environment Interface, University of Veterinary Medicine Hannover, Foundation, Hanover, Germany

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Measuring animal welfare and animal health has steadily increased in importance recently. Many methods have been developed, and a rough distinction, following the EFSA [1], can be made between animal-based, active indicators and passive, environment-based indicators. The latter can be used if their relation to animal health issues is very close and their recording is easier, which is important for monitoring the health status of a wide range of pig farms. GrosseKleimann et al. [2] have already derived a set of seven health scores and one total score for this purpose, which allows longitudinal screening of a target collective of farms under conventional production conditions in Germany. One-half of the scores is based on findings in slaughtered pigs, whereas the other half consists of information regarding biological performance (average daily gain and feed conversion ratio), mortality and treatment with antibiotics. The data were generated from monitoring and advisory processes along the pig production chain, and both were checked and processed for secondary data use beforehand.

Concerning the aforementioned health scores [2], the question is whether management and farm characteristics could influence them and whether particular factors for the prevention of health incidents can be identified. Several studies have addressed this issue and found production items that can be used as a basis for extended veterinary and agricultural consulting to improve the health status of fattening pigs in the long run. However, these factors may vary by production system or the particular collective of farms. Therefore, this study will describe which farm characteristics can be used as adjusting screws for the health scores in fattening pigs housed in typical German pig production systems .

Additionally, this investigation addresses the degree to which health scores, which are based on findings in slaughtered pigs, can be reflected or complemented by animal welfare indicators that are collected on farms in fattening pigs.

## Methods

### Study design

The joint project “Multivariate Assessment of Animal Welfare through Integrative Data Collection and Validation of Welfare Indicators in Finishing Pigs” (MulTiViS) is conducted by a consortium of the University of Veterinary Medicine Hannover, foundation (“Stiftung Tierärztliche Hochschule Hannover”), the Swine Health Service of the Chamber of Agriculture in Lower Saxony (“Schweinegesundheitsdienst”), the swine service provider VzF (registered association, VzF) and Marketing Service Gerhardy. The owners of

a sample of 207 commercial fattening pig units from northeastern Lower Saxony, Germany, all under the advisory service of VzF, were asked for their written confirmation of participation.

For each pig unit, several sources of production data were available, acquired by VzF in the context of farm branch analysis. These production data contain information on various farm and management characteristics, biological and economic performance data, findings in slaughtered pigs and information on antibiotic usage.

Additionally, two trained veterinarians visited each farm on one day between 3 and 2017 and 12 November 2018. They examined fattening pigs in a maximum of eight randomly selected pens per farm ( $7.8 \pm 1$ ; arithmetic mean  $\pm$  standard deviation), according to the method of Kish, 1949 [3], which led to 24,715 observed animals in 1,201 pens. Thirty welfare and health indicators were investigated as on-farm health information. To achieve a holistic “animal health dataset”, on-farm health information was linked to production data from the corresponding half-year in which the on-farm examination took place.

Since Grosse-Kleimann et al. [2] found the initial body weight (IBW) on day one of the fattening period to be an important stratification criterion, only herds with IBW between 24 and 33.5 kg were chosen for further analyses. Additionally, data from pigs housed in pens with straw bedding or outdoor climate were excluded because the number of pens was too small to be representative (4 farms with exclusively straw bedding or outdoor climate). Furthermore, on one farm, on the day of the examination, unexpectedly, no animals were present. Finally, a collective of  $n = 154$  pig units remained for further analyses. The farm size varied between 160 and 3,360 ( $1,153 \pm 582$ ) fattening pig places (FPs).

### Health scores

Seven standardized health scores and one total score according to Grosse-Kleimann et al. [2] were chosen as the target variables: mortality (MOR), average daily gain (ADG), feed conversion ratio (FCR), treatment frequency (TF) and meat inspection indicators associated with respiratory health (RESP), exterior injuries or alterations (EXT) and animal management (MANG). They are based on selected indicators from production data (Table 1) and were standardized on a zscale, with a mean of 0 and a standard deviation of 1. For respiratory slaughter findings of pneumonia and pleurisy, in Germany, four categories are usually documented at meat inspection. For the study, a unique prevalence was calculated by combining moderate and high alterations as positive records and slight and no alterations as negative records.

**Table 1** Source and descriptive measures of indicators, which are the basis for examined health scores in the study collective of 154 pig units, according to Grosse-Kleimann et al. [2]. (MOR mortality score, ADG average daily gain score, FCR feed conversion ratio score, TF treatment frequency score, RESP respiratory lesions score, EXT exterior lesions score, MANG animal management score, UDD used daily doses, FP fattening pig place)

score	Indicator	source	mean	SD	min	max
MOR	mortality [%]	performance data	2.56	1.77	0.17	13.73
ADG	average daily gain [g]	performance data	849	77	631	1,036
FCR	feed conversion ratio [kg/kg]	performance data	2.83	0.19	2.45	3.34
TF	treatment frequency [UDD/FP]	antibiotic usage data	2.22	3.57	0.00	20.13
RESP	pneumonia [%]	findings in slaughtered pigs	11.99	8.52	0.65	80.82
	pleurisy [%]	findings in slaughtered pigs	6.43	6.18	0.00	29.92
	pericarditis [%]	findings in slaughtered pigs	3.69	3.20	0.00	21.83
EXT	arthritis [%]	findings in slaughtered pigs	0.58	0.48	0.00	2.43
	abscess [%]	findings in slaughtered pigs	0.97	0.68	0.00	3.67
	ear lesions [%]	findings in slaughtered pigs	0.01	0.05	0.00	0.37
	tail lesions [%]	findings in slaughtered pigs	0.76	1.12	0.00	7.42
	dermal alterations [%]	findings in slaughtered pigs	0.16	0.23	0.00	1.54
	bursitis [%]	findings in slaughtered pigs	0.46	0.70	0.00	6.14
MANG	liver milk spots [%]	findings in slaughtered pigs	6.89	10.69	0.00	61.16
	dermal damage (handling) [%]	findings in slaughtered pigs	0.05	0.24	0.00	2.30
	intestinal alteration [%]	findings in slaughtered pigs	0.60	1.18	0.00	13.93
	whole carcass condemnation [%]	findings in slaughtered pigs	0.17	0.21	0.00	1.05

### Factors selected from farm characteristics

Sixteen categorical factors, e.g., type of buying in criteria, boar fattening or feeding techniques, and 14 metric factors, e.g., costs for veterinary service or daily feed intake, were extracted from farm characteristics (for a complete list of all factors see [Additional file 1](#)). Factor levels with less than ten pig units were omitted for statistical analyses to avoid extended variation due to sparse data.

To investigate both the individual influence of farm characteristics on the health scores and the interaction with others, uni- and multifactorial ANOVA and linear regression models were conducted. To find the set of variables in FC that showed the most significant influence on each health score, multifactorial analyses of covariance were conducted with a combination of forward selection (entering criterion:  $p$ -value < 0.1) and manual backward selection (removal criterion:  $p$ -value > 0.05).

In preparation for statistical modelling and to control multicollinearity, association analyses within influencing factors were previously applied to all variables [4].

### Preparation and analysis of on-farm indicators

During investigation on the farms, two different age groups of pigs were distinguished: first to sixth fattening weeks (age group 1) and seventh to twelfth fattening weeks (age group 2). Since regrouping of pigs for fattening and mixing of unfamiliar pigs often cause rank fights

[5, 6], an increased incidence of health- and production-relevant findings may occur in age group 1. In addition, the time until slaughter is longer in age group 1, and thus, the probability is higher that injuries developed in the early fattening period have healed or are no longer visible. Hence,  $t$ -tests were carried out to identify those indicators whose mean prevalence differed significantly between the age groups. If differences were identified, for these characteristics, a correction factor for age group 1 was calculated by the following formula:

$$p_{corr} = p \times \frac{mean_{group2}}{mean_{group1}}$$

where  $p_{corr}$  = corrected prevalence of age group 1 and  $p$  = observed prevalence of age group 1. This approach is in line with a general adjustment by ratio extrapolation [7].

On-farm health information was originally examined as pen-specific prevalence (number of affected animals per pen), but health scores were calculated at the farm level. To bring them to a uniform level, an average farm-specific prevalence for each on-farm indicator and farm was calculated. To meet the requirements for modelling, logit transformation and  $z$ -standardization were applied on all variables following the approach of Nienhaus et al. [8].

To investigate whether the slaughter scores could sufficiently depict pig health or whether investigation of living animals is necessary, several analyses were carried out. A principal component analysis was implemented to provide an overview of the interconnections between on-farm indicators and RESP, EXT and MANG. As the first and second components of the principal component analysis are considered the most discriminating, they were plotted on a factor map. Unifactorial ANOVA for each selected on-farm indicator and three slaughter scores were conducted to examine the degree of predictability and to reveal possible supplemental factors for health monitoring. Previously, the Spearman correlation coefficient was calculated for each on-farm and slaughter indicator.

All statistical evaluations were performed with SAS®, version 9.4 TS level 1M5 (SAS Institute Inc., Cary, NC, United States).

## Results

### Health status (scores)

On average, the mortality was  $2.56\% \pm 1.77\%$ , and the mean average daily gain was  $849\text{ g} \pm 77\text{ g}$ . The food conversion ratio ranged between 2.45 kg/kg and 3.34 kg/kg, and the mean number of used daily doses per fattening place was  $2.22 \pm 3.57$ . Pneumonia, pleurisy, pericarditis and liver milk spots showed a mean prevalence of  $11.99\% \pm 8.52\%$ ,  $6.43\% \pm 6.18\%$ ,  $3.69\% \pm 3.20$  and  $6.89\% \pm 10.69\%$ , respectively, whereas all other slaughter indicators occurred only at a very low level (mean prevalence  $< 1\%$ ). Detailed results are shown in Table 1.

### Farm characteristics and their impact on health scores

Most of the pig units (48.1%) purchased surgically castrated male pigs that were housed in the same pen as females. The majority of the study farms (46.7%) offered  $0.75\text{ m}^2$  per fattening pig place, and 81.8% had pens that were equipped with fully slatted floors. Regarding feeding techniques, the main type was fully automatic liquid feeding (38.0%). On average, the production costs varied between 0.57 € and 0.91 € per kg of live weight gain. The mean costs for veterinary service per sold animal were  $1.00\text{ €} \pm 1.17\text{ €}$ , and the feed intake per pig and day was  $2.40\text{ kg} \pm 0.21\text{ kg}$ . Detailed descriptive results are shown in Additional file 1.

In preparation for modelling, the two variables number of FP and live weight at slaughter were excluded for further analyses to avoid multicollinearity. The results of unifactorial ANOVA revealed that three farm characteristics had no significant influence on any of the observed health scores (feed, phase feeding and feed availability). The factor with the highest number of p-values for entering final models was feed intake per pig and day. Manual backwards selection was performed for the multifactorial

models with a p-value  $< 0.05$  as the selection criterion. It was shown that ADG was the most influential health score, with  $R^2 = 96.1\%$  by observed farm characteristics. The results of uni- and multifactorial analyses for MOR, ADG and EXT are shown in Table 2. Detailed results for all scores are listed in Additional file 2.

The results of multifactorial analyses usually revealed no substantial differences in parameter estimates for categorical characteristics in comparison to unifactorial analyses (see Tables 2, 3 and 4). The type of buying in criteria and boar fattening were the variables with the most influence, as they are the only categorical characteristics remaining in a final model. The type of buying in criteria is a variable that contains information on where farms source their piglets from and the management type behind it. The reference group, type A, stands for the typical VzF farm with pigs originating from a German farm and having BHZP (German Federal Hybrid Breeding Programme) as the dominant sow and boar breed. Types B and C also obtain German piglets with BHZP as the main boar breed, but DAN (DanBred P/S) and PIC (Pig Improvement Company) are the primary sow breeds. Types D and E have mixed sow breeds and BHZP or mixed breeds for boar, respectively. Type F obtains imported piglets from Denmark, whereby the sow and boar breed are not taken into account. Type G was created by merging subtypes that were too few for a separate group. Farms with type E buying in criteria seemed to have fewer exterior lesions, whereas type B farms showed better EXT scores than the reference group. In farms with boar fattening, animals showed higher values in EXT than in farms that purchase castrated males.

With regard to the metric variables, it is noticeable that above all, monetary factors such as production costs or costs for veterinary service seem to play an important role. Thus, in the unifactorial approach, increasing production costs per kg live weight gain (LWG) had the effect of poorer scores for MOR and ADG. However, when the influence of other significant factors is taken into account, production costs no longer seem to have such a strong effect on ADG and are even associated with a considerably lower mortality. Poorer performance in MOR is also reflected by higher veterinary costs. An increase in feed intake per pig and day is reflected in significantly better daily gain and better scores on the EXT.

### On-farm health indicators

Contrasting the characteristics under study by age group showed that significant differences in prevalence level occurred for ten of the remaining on-farm indicators (Table 5). Tail biting lesions, ear haematoma and coughing index did not vary substantially between the age groups. Bursa auxiliaris, bursitis, faecal skin dirtying,

ophthalmic discharge, conjunctivitis and lameness were more often found in age group 2, whereas skin lesions, purulent nasal discharge, flank biting lesions, ear biting lesions and pigs showing signs of diarrhoea at the pen level showed a higher prevalence in age group 1.

Taking the adjustment of Table 5 into account, it was found that Bursa auxiliaris was by far the most frequently observed finding (76.0% ± 18.3%). This was followed by faecal skin dirtying (12.4% ± 15.7%) and ophthalmic discharge (8.9% ± 9.9%), whereas flank biting lesions were found the least (0.7% ± 1.3%). In general, it could be stated that most prevalence levels were very low (< 5%) and therefore showed a right-skewed distribution, which justifies logittransformation and zstandardization. Detailed results are shown in Table 6, and the observed prevalence of all on-farm indicators can be found in Additional file 1.

The majority of the Spearman correlation coefficients for on-farm and slaughter indicators lay beneath 0.2 and the highest coefficient was calculated for tail biting lesions recorded on-farm and at slaughter with 0.33. Nevertheless, the factor map of the principal component analysis of selected on-farm indicators and slaughter scores (Fig. 1) underlines the technical association between scores based on findings in slaughtered pigs and corresponding indicators in living animals. However, Bursa auxiliaris and bursitis were spotted at some distance from the other exterior lesions, and RESP was found to be closer to lameness than to the respiratory indicators. The first component explains 13.2%, and the second component explains 11.8% of the total variance.

The results of the unifactorial ANOVA revealed that none of the selected on-farm indicators could predict

any slaughter score to a higher degree than 8.26% ( $R^2$  of model "EXT = ear haematoma"), and the majority of  $R^2$  values were lower than 1% (Table 7). Ear haematoma and faecal skin dirtying showed significant p-values in two models, whereas skin lesions, flank biting lesions, tail biting lesions, ear biting lesions and lameness each predicted one score significantly, and the remaining factors showed no significant p-value. Therefore, these factors may be considered potential supplemental indicators for health monitoring, as scores based on findings in slaughtered pigs cannot be depicted by any of them.

EXT was the score with the highest number of significant associations with on-farm indicators. There was also a great degree of professional agreement, since all observed indicators concerning exterior lesions, despite bursa and bursitis, predicted EXT significantly. In contrast, RESP indicated a significant association with ear haematoma only, but none of the respiratory indicators and MANG could be significantly predicted by only faecal skin dirtying.

## Discussion

Animal-based indicators may be an important tool for health monitoring, although they cannot replace but supplement good stockmanship or consultancy for early detection of herd health problems [9]. The present study provides insight into the different management and farm characteristics of 154 typical German fattening pig units and their influence on established health scores. Furthermore, an interconnection between selected health scores and welfare indicators recorded on-farm was analysed.

**Table 2** Results of unifactorial analyses (factors with pvalue < 0.1) and final multifactorial model (factors with pvalue < 0.05) for MOR (mortality score).  $R^2$  of final model = 17.54%. Factors remaining in the final model are in bold letters, and the reference level is in italic letters (BFCD benefits free of direct costs, LWG live weight gain)

Factor	factor levels	unifactorial		multifactorial	
		p-value	estimate	p-value	estimate
<b>BFDC/100 kg LWG</b>		< 0.0001	-0.04	< 0.0001	-0.05
<b>feed energy/kg LWG</b>		0.0161	0.08	0.0314	0.08
feed availability		0.0570			
	misc.	0.7548	0.09		
	ad libitum	0.0175	0.42		
	rationed	*	*		
<b>costs for veterinary service/pig</b>		0.0672	0.13	0.0482	0.14
<b>production costs/kg LWG</b>		0.0813	2.23	0.0043	-5.59
feed energy		0.0992	0.08		
	< 13 MJ	0.0992	0.28		
	13–13,4 MJ	*	*		

\* No p-value and estimate for reference level

**Table 3** Results of unifactorial analyses (factors with pvalue < 0.1) and final multifactorial model (factors with pvalue < 0.05) for ADG (average daily gain score). R<sup>2</sup> of final model = 96.10 %. Factors remaining in the final model are in bold letters, and the reference level is in italic letters (*BFGD* benefits free of direct costs, *LWG* live weight gain, *FP* fattening pig place)

Factor	factor levels	unifactorial		multifactorial	
		p-value	estimate	p-value	estimate
<b>needed feed/pig/day</b>		< 0.0001	-3.49	< 0.0001	-4.31
<b>feed energy/kg LWG</b>		< 0.0001	0.18	< 0.0001	0.24
group size		< 0.0001			
	<i>misc.</i>	0.0227	0.46		
	21–50 pigs	0.0144	-0.58		
	13–20 pigs	0.0034	-0.56		
	<i>1–12 pigs</i>	*	*		
space per FP		< 0.0001			
	> 0.9 m <sup>2</sup>	< 0.0001	-1.14		
	0.825 m <sup>2</sup>	0.0277	-0.37		
	<i>0.75 m<sup>2</sup></i>	*	*		
type of buying in criteria**		0.0001			
	G	0.0142	0.75		
	F	0.0010	-0.99		
	E	0.0507	-0.50		
	D	0.2479	0.30		
	C	0.8341	-0.06		
	B	0.5755	-0.12		
	<i>A</i>	*	*		
pigs/FP		0.0011	-0.83		
feeding techniques		0.0022			
	<i>misc.</i>	0.1604	0.38		
	mash	0.0073	-0.46		
	<i>liquid</i>	*	*		
water supply		0.0049			
	public	0.0012	-0.56		
	private well, water treated	0.6106	-0.16		
	<i>private well</i>	*	*		
<b>production costs/kg LWG</b>		0.0083	3.27	< 0.0001	1.29
floor type		0.0145			
	<i>misc.</i>	0.0048	0.70		
	partially slatted	0.3237	0.32		
	<i>fully slatted</i>	*	*		
feed availability		0.0192			
	<i>misc.</i>	0.3581	0.26		
	ad libitum	0.0177	-0.40		
	<i>rationed</i>	*	*		
live weight of losses		0.0227	-0.02		
boar fattening/single sex		0.0245			
	yes/yes	0.8974	0.03		
	no/yes	0.0103	-0.44		
	<i>no/no</i>	*	*		

**Table 3** Results of unifactorial analyses (factors with pvalue < 0.1) and final multifactorial model (factors with pvalue < 0.05) for ADG (average daily gain score). R<sup>2</sup> of final model = 96.10 %. Factors remaining in the final model are in bold letters, and the reference level is in italic letters (*BFCD* benefits free of direct costs, *LWG* live weight gain, *FP* fattening pig place) (Continued)

Factor	factor levels	unifactorial		multifactorial	
		p-value	estimate	p-value	estimate
slaughter weight		0.0272	-0.08		
BFDC/100 kg LWG		0.0311	-0.02		
post fattening		0.0478			
	yes	0.0478	0.37		
	no	*	*		
weight gain/pig		0.0890	-0.04		

\* No p-value and estimate for reference level

\*\* For explanation of factor levels see "Discussion/Impact of farm and management characteristics"

**Herd selection and availability of data**

Based on the results of Grosse-Kleimann et al. [2], only those farms were taken into account that started the fattening period, with pigs having an initial body weight of 24 to 33.5 kg. This bodyweight range is typical in Germany and therefore implemented by the majority of fattening pig farmers. As the health scores in the categories above and below this weight range deviate strongly, the results of the study can only be related to

this collective and would have to be validated for pigs of other weight or age by means of additional investigations. The production data were collected in a harmonized way by the VzF advisors as part of the farm sector analysis and were processed for the project especially as described in a previous publication. For the level classification of categorical factors, it must be kept in mind that a level fulfils the situation on a farm to at least 90 %. This means that a certain inaccuracy in the data

**Table 4** Results of unifactorial analyses (factors with pvalue < 0.1) and final multifactorial model (factors with pvalue < 0.05) for EXT (exterior lesions score). R<sup>2</sup> of final model = 26.12 %. Factors remaining in the final model are in bold letters, and the reference level is in italic letters

factor	factor levels	unifactorial		multifactorial	
		p-value	estimate	p-value	estimate
<b>type of buying in criteria**</b>		0.0014		0.0024	
	G	0.6629	0.08	0.8411	0.04
	F	0.0011	-0.57	0.0107	-0.47
	E	0.9358	0.01	0.6050	0.08
	D	0.4364	0.12	0.8524	0.03
	C	0.0053	-0.50	0.0010	-0.60
	B	0.5943	-0.07	0.2323	-0.15
	A	*	*	*	*
<b>feed intake/pig/day</b>		0.0040	-0.64	0.0400	-0.48
phosphor reduction		0.0135			
	yes	0.0135	-0.26		
	no	*	*		
<b>costs for disinfection***/pig</b>		0.0137	-0.49	0.0069	-0.53
<b>boar fattening/single sex</b>		0.0479		0.0112	
	yes/yes	0.0168	0.36	0.0078	0.38
	no/yes	0.8943	0.01	0.5161	-0.06
	no/no	*	*	*	*
animals sold		0.0531	0.00		

\* No p-value and estimate for reference level

\*\* For explanation of factor levels see "Discussion/Impact of farm and management characteristics"

\*\*\* Within the scope of the final cleaning of the stables

**Table 5** Mean prevalence [%] at the pen level of on-farm indicators and p-value of the t-test for age group 1 (n = 669 pens) and age group 2 (n = 532 pens)

Indicator	age group 1		age group 2		p-value	correction factor
	mean	SD	Mean	SD		
bursa auxiliaris	63.0	23.4	76.4	17.8	< 0.0001	1.21
bursitis	1.7	3.7	2.4	4.4	0.0012	1.41
faecal skin dirtying	9.0	16.6	12.2	18.6	0.0020	1.36
skin lesions	5.5	9.9	3.0	6.6	< 0.0001	0.54
purulent nasal discharge	2.2	4.8	1.3	3.3	0.0002	0.59
ophthalmic discharge	4.8	7.7	8.9	12.7	< 0.0001	1.85
conjunctivitis	2.9	7.4	5.0	10.3	< 0.0001	1.72
flank biting lesions	2.2	5.9	0.7	2.9	< 0.0001	0.32
tail biting lesions	2.7	8.2	3.2	8.6	0.2694	-*
ear haematoma	1.3	3.5	1.6	3.3	0.1415	-*
ear biting lesions	3.1	7.5	0.9	4.9	< 0.0001	0.29
lameness	1.7	3.7	2.3	4.6	0.0091	1.35
coughing index	1.3	0.1	1.6	0.1	0.0843	-*

\* No correction factor calculated because of nonsignificant t-test

cannot be ruled out. Since many pig units have grown over a long period, a 100 % accurate classification is often not possible.

To record the status quo of the study herds, on-farm health information was acquired. Since the aim of the study was to develop factors for monitoring that are easy to implement under practical conditions, each pig unit was visited only once, and the fattening pigs from a maximum of eight randomly selected pens were examined. In contrast, the production data and health scores refer to the aggregated data of a complete half-year and

take into account all pigs that were kept or sold for slaughter during this period. However, with regard to the whole study collective, the large number of visited farms and pigs statistically compensates for this. Hence, to obtain a rough impression of the health status of individual pig herds, the survey is indicated to be suitable.

#### Impact of farm and management characteristics

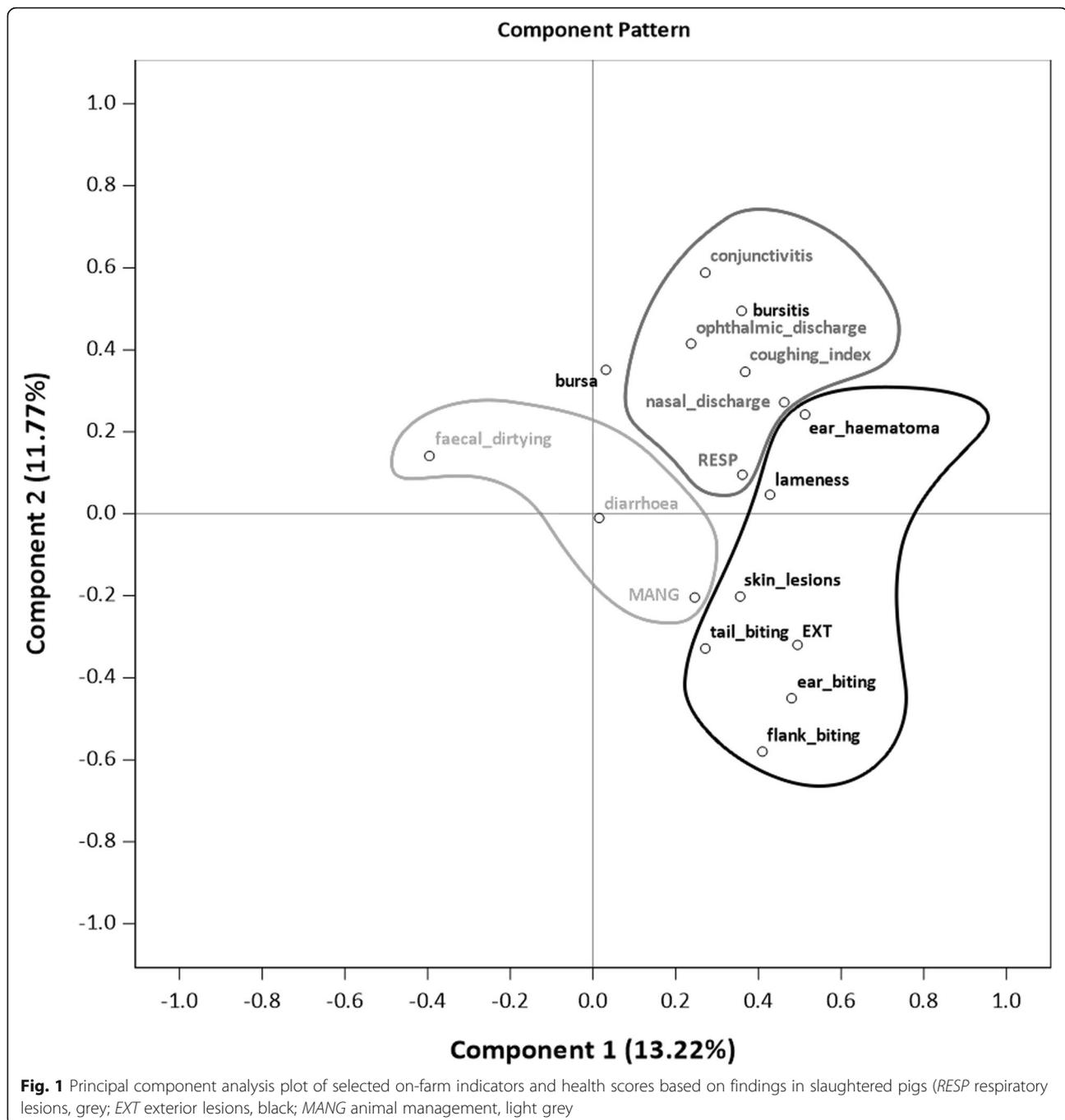
A positive impact on pig performance through record keeping by advisory services in general was demonstrated by van Staaveren et al. [10]. Pandolfi et al. [11]

**Table 6** Descriptive measures of selected health indicators from on-farm health information in the study collective of 154 pig units (farm-specific prevalence, corrected by age group)

Indicator	unit	mean	median	SD	CV	min	max
bursa	%	76.0	80.8	18.3	24.1	10.5	100.0
bursitis	%	2.4	1.9	2.2	90.6	0.0	13.2
faecal skin dirtying	%	12.4	5.6	15.7	126.7	0.0	89.0
skin lesions	%	2.8	1.6	3.6	126.5	0.0	21.5
purulent nasal discharge	%	1.3	0.9	1.6	124.8	0.0	10.0
ophthalmic discharge	%	8.9	5.9	9.9	110.6	0.0	59.3
conjunctivitis	%	4.9	1.1	8.6	176.7	0.0	45.0
flank biting lesions	%	0.7	0.0	1.3	176.2	0.0	8.1
tail biting lesions*	%	2.9	1.6	4.2	147.1	0.0	25.7
ear haematoma*	%	1.5	1.1	1.5	105.6	0.0	8.4
ear biting lesions	%	0.9	0.3	1.5	181.2	0.0	12.1
lameness	%	2.3	1.6	2.2	98.3	0.0	10.6
coughing index***	bouts/min	0.05	0.03	0.06	118.0	0.0	0.4
diarrhoea	%	8.3	0.0	16.3	195.7	0.0	87.5

\* Observed prevalence (no correction factor because of nonsignificant t-test)

\*\* Number of coughing bouts in 2 × 3 min, divided by number of animals per pen



found that farm and management characteristics, such as the ventilation system or group size, can act as potential starting points for improving animal welfare. For the present study, comparison of uni- and multifactorial analyses indicates an association of several factors with the given health scores. However, in an overall context, regarding the other characteristics under study, confounding changes the risk perspective. Since one aim of the study was to find the set of farm characteristics for

each health score that showed the most significant influence and could be usable as regulation screws to improve pig health, the focus was on multifactorial analyses.

Mortality is generally accepted as an indicator in most epidemiological studies describing animal health [8, 12–14]. Here, MOR is primarily associated with direct cost-free benefits as well as production costs: the higher the number of animals that died or were culled pre-

**Table 7** Results of unifactorial analyses for on-farm indicators and selected health scores (age group corrected, logit- and z-transformed, farm specific prevalence). Significant p-values and corresponding indicators are bold

indicator	RESP		EXT		MANG	
	p-value	R <sup>2</sup>	p-value	R <sup>2</sup>	p-value	R <sup>2</sup>
bursa auxiliaris	0.4871	0.32 %	0.7251	0.08 %	0.5778	0.20 %
bursitis	0.9683	0.00 %	0.2307	0.94 %	0.5673	0.22 %
<b>faecal skin dirtying</b>	0.4393	0.39 %	<b>0.0155</b>	3.80 %	<b>0.0168</b>	3.70 %
<b>skin lesions</b>	0.4036	0.46 %	<b>0.0066</b>	4.75 %	0.0587	2.33 %
purulent nasal discharge	0.1004	1.77 %	0.4304	0.41 %	0.3835	0.50 %
ophthalmic discharge	0.7804	0.05 %	0.7521	0.07 %	0.3270	0.63 %
conjunctivitis	0.2180	1.00 %	0.2741	0.79 %	0.5917	0.19 %
<b>flank biting lesions</b>	0.1635	1.27 %	<b>0.0071</b>	4.67 %	0.7537	0.06 %
<b>tail biting lesions</b>	0.8921	0.01 %	<b>0.0121</b>	4.07 %	0.2554	0.85 %
<b>ear haematoma</b>	<b>0.0306</b>	3.04 %	<b>0.0003</b>	8.26 %	0.2598	0.83 %
<b>ear biting lesions</b>	0.0987	1.78 %	<b>0.0011</b>	6.82 %	0.3282	0.63 %
<b>lameness</b>	0.4515	0.37 %	<b>0.0337</b>	2.93 %	0.1237	1.55 %
coughing index	0.0712	2.13 %	0.4843	0.32 %	0.5382	0.25 %
diarrhoea	0.4208	0.43 %	0.3284	0.63 %	0.5431	0.24 %

slaughter, the lower the profits and the more production costs per kilogram of carcass sold arise. Other farm characteristics, however, did not show a significant influence in the multifactorial analysis in the study collective. In contrast, Agostini et al. [15] found that multiple origins of pigs led to higher mortality. They also observed lower mortality in pigs with higher IBW, which is in line with the present study. An investigation of grow-fattening pigs in Belgium [16] also identified the origin of piglets and furthermore the duration of fattening period and season for the beginning of the fattening period as risk factors for higher mortality. Similar results were shown by Oliveira et al. [17]. The different results reveal that the mortality rate is a multicausal event that is difficult to attribute to one specific reason or farm characteristic.

Although average daily gain is a non-animal-based measure, it could be a sign of animal health problems and therefore is an important indicator for an overall assessment of pig health [18]. The results of the present study showed that the ADG is the score that is by far the most influential or predictable by the production data. Goodness of fit well over 90 % can be achieved through modelling with the feed intake per day, the feed energy per kg of LWG and the production costs. These variables are closely related in terms of content, as production costs largely consist of feed costs and thus the daily feed intake and feed energy, respectively. However, these are exclusively variable and measurable factors and not farm-specific conditions that can be adjusted to influence the target value. Nevertheless, during a consultation, it is possible to determine which of the three

factors is not within the normal range and use this information as a starting point for improvement in daily weight gain. Other studies that examined the influence of farm characteristics on average daily gain did not find a significant effect of group size [19] or castration (physical and via vaccination) [20] but a positive impact of straw bedding on average daily gain [21].

The use of slaughter findings as a monitoring and surveillance tool for pig health is frequently implemented in Europe [22–24]. However, the influence of farm characteristics on findings in slaughtered pigs has not been consistently confirmed. Kongsted et al. [25] showed that the production system has an impact on several health-related lesions, whereas Cagienard et al. [26] could not find significant differences in meat inspection data related to the type of housing.

In the current study, regarding the animal-based slaughter score EXT, the herd attributes type of buying in criteria and fattening boars stayed in the final models as significant effects. Type B (German piglets with BHZP as main boar breed and DAN as main sow breed) and E (German piglets with mixed boar and sow breeds) farms seem to have fewer exterior lesions than the reference group. However, it must be taken into account that the comparison groups in this study were relatively small, with 12 and 13 farms, respectively; therefore, the significance of the results needs to be interpreted with caution. The fattening of boars also plays a role in exterior lesions. Farms that housed entire male pigs had more problems with external injuries, such as skin lesions or tail and ear biting lesions. This is in line with common literature, which found increased sexual and aggressive

behaviour against pen mates in single-sex male groups [27, 28]. Concerning metric FC, the intake of feed per pig and day seem to have a protective influence on EXT. The suggestion is that satisfied animals are lazier and therefore show less aggression against pen mates.

#### Reflection of on-farm information in health scores

One aim of the study was to develop health indicators for monitoring that are easy to implement under practical conditions. Therefore, no repeated survey was carried out on the farms, and only a random sample of individual pigs inspected was investigated. In contrast, the production data and health scores refer to the collected data of a complete half-year and take into account all animals that were kept or sold during this period. This aspect is underlined by the relatively low  $R^2$  of the unifactorial on-farm health information models (Table 8). Nevertheless, veterinary examination can give an impression of the health status of individual pig herds and provide clues as to which health problems are not represented by routine processes along the pig production chain [29].

The results of official meat inspection are, on the one hand, animal-based, but on the other hand, the lifetime health status is displayed in a way such that special health events are hidden by aggregation.

The idea of comparing the actively recorded health status of living animals with assessments based on slaughter findings has already been pursued several times. Maisano et al. [30] detected similar prevalence levels of animal-based measures on-farm and slaughter plants, whereas Carroll et al. [31] were able to establish little correlation between health problems in live animals and skin and tail lesions that developed at least ten weeks before slaughter. Van Staaveren et al. [32] also identified skin and tail lesions as “potential iceberg indicators” at the slaughterhouse to replace on-farm assessments. This fits with the results from the present study, as skin lesions and bite lesions on ears, tail and flanks were also significantly mapped by EXT, and a small correlation was found for tail biting lesions on farm and tail lesions detected at slaughter.

The respiratory tract findings, on the other hand, showed only a small statistical correlation between indicators at the slaughter plant and the corresponding indicators at the farm. Other studies that investigated the association of lung or pleural lesions and on-farm indicators for respiratory diseases also found no or little significance [33, 34]. Even though the principal component analysis supports a substantial relationship, RESP could be significantly modelled by blood ears only and none of the technically assigned ones, e.g., coughing index or purulent nasal discharge. One possible explanation for this may be that respiratory lesions in slaughter pigs are

often caused by enzootic pneumonia (*Mycoplasma hyopneumoniae*) [35], which is typically clinically apparent from the sixth to the eighth fattening week [36]. As the study included pigs from the first to the twelfth fattening week, the typical time period for coughing caused by enzootic pneumonia was underrepresented.

MANG could also scarcely be modelled by the on-farm indicators. However, this is to some extent reasonable, as the score includes, among others, the indicator “dermal damage by handling”, which occurs during transport from the farm to the slaughterhouse and therefore is not expected to be found during on-farm examination. Another score component, liver milk spots, usually remains clinically inconspicuous in fattening pigs [36, 37] and therefore cannot show connection to the on-farm indicators.

In addition to respiratory on-farm indicators (ophthalmic and nasal discharge, conjunctivitis and coughing index), Bursa auxiliaris, bursitis and diarrhoea showed no significant association with any of the slaughter scores. From this, the recommendation can be derived to consider these indicators in addition to the slaughter findings for health monitoring.

#### Conclusions

In general, the present study shows that the investigated health scores are influenced to varying degrees by different farm and management characteristics. However, many of the influencing variables examined have a stronger influence when considered individually, but they do not longer play a decisive role in the multifactorial context. In practice, this confounding effect suggests that the whole farm with all its characteristics should always be taken into account for advising and changing of just individual factors has to be avoided.

From a specific perspective, ADG can be explained in a multifactorial model to almost 100 % by production costs and feed intake, whereas models for animal-based scores such as MOR or EXT had a much lesser coefficient of determination. Hence, MOR and EXT seem to be more important for describing animal welfare.

With regard to the on-farm indicators, a general contextual connection could be established via principal component analysis. In connection with the associated slaughter scores, however, it became apparent that only EXT could be mapped significantly by the result of the survey on the farms. Following this, the on-farm collection of respiratory indicators, e.g., coughing index or purulent nasal discharge as well as bursa or bursitis and diarrhoea, could be recommended as supplemental items for health monitoring.

Overall, it can be stated that production data are suitable to map health issues of fattening pigs in the context of monitoring and to obtain an impression of possible

adjusting screws for consultancy in the next step. In this study, the production data are based on aggregated data of a sixth month period, and the on-farm examination of pigs may provide additional information in the form of a snapshot.

#### Abbreviations

ADG: average daily gain score; BHZP: German Federal Hybrid Breeding Programme; EXT: exterior lesions score; FCR: feed conversion ratio score; LWG: live weight gain; MANG: animal management score; MOR: mortality score; PIC: Pig Improvement Company; RESP: respiratory lesions score; TF: treatment frequency score; VzF: Association for Promoting Farming Economics

#### Supplementary information

The online version contains supplementary material available at <https://doi.org/10.1186/s40813-021-00225-y>.

**Additional file 1** Descriptive results of all farm characteristics-factors and on-farm indicators.

**Additional file 2** Results of uni- and multifactorial analyses for all health scores and farm characteristics.

#### Acknowledgements

This work was conducted under the MultIVIS project. The MultIVIS consortium would like to thank all participating farmers for their cooperation.

#### Authors' contributions

LK, HG and HP designed the study. HP and HM were responsible for raw data acquisition of production data. Together with HG, they edited the data and prepared it for transfer to JGK and LK. JGK, BW, IS, EGB, NK, HN, HP, HM, HG and LK were involved in the development of the survey questionnaire for on-farm health information, and BW and IS acquired the information via farm visits. JGK performed further processing steps and analyses of the entire data and wrote the initial draft. LK and HG critically reviewed the manuscript. All authors have read and approved the manuscript.

#### Funding

The project is supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support programme with the number 2817905315. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. This publication was supported by "Deutsche Forschungsgemeinschaft" and University of Veterinary Medicine Hannover, Foundation within the funding programme Open Access Publishing. Open Access funding enabled and organized by Projekt DEAL.

#### Availability of data and materials

The data were collected on an individual basis from farmers and slaughterhouses. Each participant gave written consent with the understanding that data would not be transferred to a third party. Therefore, any data transfer to interested persons is not allowed without an additional formal contract. Data are available to qualified researchers who sign a contract with the University of Veterinary Medicine Hannover and VzF. This contract will include guarantees to the obligation to maintain data confidentiality in accordance with the provisions of the German data protection law. Currently, there exists no data access committee or other body who could be contacted for the data. However, for this purpose, a committee will be founded. This future committee will consist of the authors as well as members nominated by the University of Veterinary Medicine Hannover and the VzF. Interested cooperative partners who are able to sign a contract as described above may contact:

Prof. Dr. Lothar Kreienbrock  
Department of Biometry, Epidemiology and Information Processing  
University of Veterinary Medicine, Hannover  
Bünteweg 2, 30559 Hannover

Email: [lothar.kreienbrock@tiho-hannover.de](mailto:lothar.kreienbrock@tiho-hannover.de)

#### Declarations

##### Ethics approval and consent to participate

The performance data and farm characteristics used within this study were accumulated and maintained by VzF for consulting advice for farmers. Information on antibiotic usage is based on mandatory application and delivery forms, and findings in slaughtered pigs are based on data from official slaughterhouse inspection. On-farm health information was acquired solely for the study. All information was provided voluntarily by farmers, signing individual written consent data to be used by the study team only. The research does not involve any regulated animals, and there were no scientific procedures performed on animals of any kind. For this reason, formal approval by an ethical committee was not necessary under the provisions of the German regulations.

##### Consent for publication

Not applicable.

##### Competing interest

The authors declare that they have no competing interests.

##### Author details

<sup>1</sup>Department of Biometry, Epidemiology and Information Processing, WHO Collaborating Centre for Research and Training for Health in the Human-Animal-Environment Interface, University of Veterinary Medicine Hannover, Foundation, Hanover, Germany. <sup>2</sup>Institute for Animal Hygiene, Animal Welfare and Farm Animal Behaviour, University of Veterinary Medicine Hannover, Foundation, Hanover, Germany. <sup>3</sup>Swine Health Service, Chamber of Agriculture in Lower Saxony, Oldenburg, Germany. <sup>4</sup>Field Station for Epidemiology, University of Veterinary Medicine Hannover, Foundation, Bakum, Germany. <sup>5</sup>VzF e.V, Association for Promoting Farming Economics, Uelzen, Germany. <sup>6</sup>MSG, Marketing Service Gerhardy, Garbsen, Germany.

Received: 12 April 2021 Accepted: 16 July 2021

Published online: 03 August 2021

#### References

- Scientific Opinion on the use of animal-based measures to assess welfare in pigs. 2012.
- Grosse-Kleimann J, Plate H, Meyer H, Gerhardy H, Heucke CE, Kreienbrock L. Health Monitoring of Finishing Pigs by Secondary Data Use – A Longitudinal Analysis. *Porcine Health Management*. 2021;7(20).
- Kish L. A procedure for objective respondent selection within the household. *Journal of the American statistical Association*. 1949;44(247):380–7.
- Dohoo I, Martin SW, Stryhn H. *Veterinary Epidemiologic Research*. 2nd Edition ed. Ver Inc. Charlottetown, Prince Edward Island, Canada. 2009; p. 865.
- Meese GB, Ewbank R. The establishment and nature of the dominance hierarchy in the domesticated pig. *Anim Behav*. 1973;21(2):326–34.
- D'Eath RB. Individual aggressiveness measured in a resident-intruder test predicts the persistence of aggressive behaviour and weight gain of young pigs after mixing. *Appl Anim Behav Sci*. 2002;77(4):267–83.
- Cochran WG. *Sampling techniques* (3rd Edition). John Wiley & Sons; 1977.
- Nienhaus F, Meemken D, Schoneberg C, Hartmann M, Kornhoff T, May T, et al. Health scores for farmed animals: Screening pig health with register data from public and private databases. *PLOS ONE*. 2020;15(2).
- Harley S, More SJ, Boyle LA, O'Connell N, Hanlon A. Good animal welfare makes economic sense: potential of pig abattoir meat inspection as a welfare surveillance tool. *Irish Veterinary Journal*. 2012;65(1):1–12.
- Van Staaveren N, Teixeira DL, Hanlon A, Boyle LA. Pig carcass tail lesions: the influence of record keeping through an advisory service and the relationship with farm performance parameters. *Animal*. 2017;11(1):140–6.
- Pandolfi F, Kyriazakis I, Stoddart K, Wainwright N, Edwards SA. The 'Real Welfare' scheme: Identification of risk and protective factors for welfare outcomes in commercial pig farms in the UK. *Prev Vet Med*. 2017;146:34–43.
- Tiergerechtheitsindex für Mastschweine. In: Gumpenstein BfaLB, editor; 1995.

13. Dickhaus CP, Meemken D, Blaha T. Attempts to quantify the health status of pig herds: developing and validating a Herd Health Score (HHS). Sustainable animal production: the Challenges and Potential Developments for Professional Farming, Wageningen Academic Publishers. 2009:191–201.
14. Pandolfi F, Edwards SA, Maes D, Kyriazakis I. Connecting Different Data Sources to Assess the Interconnections between Biosecurity, Health, Welfare, and Performance in Commercial Pig Farms in Great Britain. *Frontiers in Veterinary Science*. 2018;5:41(41).
15. Agostini PS, Fahey AG, Manzanilla EG, O'Doherty JV, De Blas C, Gasa J. Management factors affecting mortality, feed intake and feed conversion ratio of grow-finishing pigs. *Animal*. 2014;8(8):1312.
16. Maes DGD, Duchateau L, Larriestra A, Deen J, Morrison RB, de Kruijff A. Risk factors for mortality in grow-finishing pigs in Belgium. *Journal of Veterinary Medicine Series B*. 2004;51(7):321–6.
17. Oliveira J, Yus E, Guitián FJ. Effects of management, environmental and temporal factors on mortality and feed consumption in integrated swine fattening farms. *Livestock Science*. 2009;123(2–3):221–9.
18. Kuratorium für Technik und Bauwesen in der Landwirtschaft e.V (KTBL). Tierschutzindikatoren: Leitfaden für die Praxis - Schwein. KTBL-Praktikerleitfaden „Tierschutzindikatoren – Schwein“; 2016.
19. Turner SP, Allcroft DJ, Edwards SA. Housing pigs in large social groups: a review of implications for performance and other economic traits. *Livestock Production Science*. 2003;82(1):39–51.
20. Pauly C, Spring P, O'doherty JV, Kragten SA, Bee G. Growth performance, carcass characteristics and meat quality of group-penned surgically castrated, immunocastrated (Improvac®) and entire male pigs and individually penned entire male pigs. *Animal*. 2009;3(7):1057–66.
21. Douglas SL, Szyzka O, Stoddart K, Edwards SA, Kyriazakis I. Animal and management factors influencing grower and finisher pig performance and efficiency in European systems: a meta-analysis. *Animal*. 2015;9(7):1210–20.
22. Pill K. Untersuchungen zur Verwendung von klinischen und pathologisch/anatomischen Befunden am Schlachthof für die Einschätzung der Tiergesundheit und des Tierschutzes in Schweine- und Rinderbeständen. Bibliothek der Tierärztlichen Hochschule Hannover; 2014.
23. Sanchez-Vazquez MJ, Strachan WD, Armstrong D, Nielsen M, Gunn GJ. The British pig health schemes: integrated systems for large-scale pig abattoir lesion monitoring. *Vet Rec*. 2011;169(16):413–.
24. Willeberg P, Gerbola MA, Petersen BK, Andersen JB. The Danish pig health scheme: Nation-wide computer-based abattoir surveillance and follow-up at the herd level. *Prev Vet Med*. 1984;3(1):79–91.
25. Kongsted H, Sørensen JT. Lesions found at routine meat inspection on finishing pigs are associated with production system. *Vet J*. 2017;223:21–6.
26. Cagienard A, Regula G, Danuser J. The impact of different housing systems on health and welfare of grower and finisher pigs in Switzerland. *Prev Vet Med*. 2005;68(1):49–61.
27. Boyle LA, Björklund L. Effects of fattening boars in mixed or single sex groups and split marketing on pig welfare. *Anim Welf*. 2007;16(2):259–62.
28. Rydhmer L, Zamaratskaia G, Andersson HK, Algers B, Lundström K, editors. Problems with aggressive and sexual behaviour when rearing entire male pigs. Proceedings of the 55th Annual Meeting of the European Association for Animal Production, Bled, Slovenia; 2004.
29. Federation of Veterinarians of Europe. Position Paper - Monitoring of farm animal welfare using animal indicators. 2018.
30. Maisano AM, Luini M, Vitale N, Rota Nodari S, Scali F, Alborali GL, et al. Animal-based measures on fattening heavy pigs at the slaughterhouse and the association with animal welfare at the farm level: a preliminary study. *animal*. 2020;14(1):108–18.
31. Carroll GA, Boyle LA, Hanlon A, Collins L, Griffin K, Friel M, et al. What can carcass-based assessments tell us about the lifetime welfare status of pigs? *Livestock Science*. 2018;214:98–105.
32. Van Staaveren N, Doyle B, Manzanilla E, Calderón Díaz JA, Hanlon A, Boyle LA. Validation of carcass lesions as indicators for on-farm health and welfare of pigs. *J Anim Sci*. 2017;95(4):1528–36.
33. Wilms-Schulze Kump F. Untersuchung der Auswirkung infektiöser und nicht infektiöser Faktoren auf die Atemwegsgesundheit der Mastschweine und deren Schlachtkörper. *Imu*; 2010.
34. van Staaveren N, Vale AP, Manzanilla EG, Teixeira DL, Leonard FC, Hanlon A, et al. Relationship between tail lesions and lung health in slaughter pigs. *Prev Vet Med*. 2016;127:21–6.
35. Fablet C, Marois C, Dorenlor V, Eono F, Eveno E, Jolly JP, et al. Bacterial pathogens associated with lung lesions in slaughter pigs from 125 herds. *Res Vet Sci*. 2012;93(2):627–30.
36. grosse Beilage E, Wendt M. Diagnostik und Gesundheitsmanagement im Schweinebestand. Stuttgart: UTB; 2013.
37. Eriksen L, Nansen P, Roepstorff A, Lind P, Nilsson O. Response to repeated inoculations with *Ascaris suum* eggs in pigs during the fattening period. *Parasitol Res*. 1992;78(3):241–6.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)



## 3 Übergreifende Diskussion

Die vorliegende Arbeit sowie die darin enthaltenen Publikationen sind im Rahmen des MuTiViS-Projektes entstanden. Die folgende übergreifende Diskussion beschäftigt sich damit, welche Aspekte und Erkenntnisse neben den veröffentlichten Ergebnissen eine wichtige Rolle bei der Planung und Durchführung einer umfassenden Tierwohlstudie spielen und welche finalen Schritte für ein nationales Monitoring notwendig sind.

### 3.1 Planung einer umfassenden Tierwohlstudie

#### 3.1.1 Selektion der Projektbetriebe

Die Auswahl des Studienkollektivs steht am Beginn jeder Planung einer Studie und ist folglich ein essentieller Schritt, da hierdurch auch entschieden wird, inwiefern ein repräsentatives Untersuchungsergebnis erreicht werden kann. Daher müssen zunächst Kriterien definiert werden, um die Art der zu untersuchenden Projektbetriebe festzulegen. Dies können Faktoren wie die Betriebsgröße, bestimmte bauliche Gegebenheiten oder grundlegende Managemententscheidungen sein. Wichtig ist dabei, potentielle Stratifizierungsmerkmale im Vorhinein zu beachten, damit die Bestände in den einzelnen Schichten mit einer hinreichenden statistischen Genauigkeit untersucht werden können. Zudem sollten die Unterkollektive ähnliche Größenordnungen aufweisen, um vergleichende Auswertungen mit ausreichender und balancierter Präzision durchführen zu können. Im Falle der vorliegenden Studie betrifft dies vor allem die Einstallgewichte der Ferkel zur Mast. Wie in **Publikation 1** thematisiert wird, hat dies maßgeblichen Einfluss auf einige der untersuchten Zielvariablen (tägliche Zunahmen, Futterverwertung, Mortalität und Therapiehäufigkeit), sodass drei verschiedene Klassen definiert wurden. Da jedoch in den Klassen 1 und 3 wesentlich weniger Betriebe vertreten waren, konnten aussagekräftige Analysen der betroffenen Indikatorvariablen nur für die Klasse 2 durchgeführt werden, welche die typischen Einstallgewichte in deutschen Produktionssystemen abbildet.

Des Weiteren ist die Verfügbarkeit und Vollständigkeit der zu untersuchenden Daten ein entscheidendes Merkmal bei der Wahl potentieller Projektbetriebe. Vor allem bei longitudinalen Auswertungen über einen längeren Untersuchungszeitraum ist darauf zu achten, dass die Daten für die einzelnen Betriebe komplett vorhanden sind oder gegebenenfalls nacherhoben werden können. Im MulTiViS-Projekt führte dies zu einem reduzierten Kollektiv für die Längsschnitt-Analysen im Rahmen der **Publikation 1** sowie dem Ausschluss von Betrieben für die Gegenüberstellung von Produktions- und Erhebungsdaten in **Publikation 2**. Hier lagen teilweise keine oder unvollständige Produktionsdaten für den Zeitraum, in dem die Erhebung stattgefunden hat, vor.

Insgesamt bedeutet dies, dass trotz der erheblichen Größe des Studienkollektivs nicht alle in Deutschland auftretenden Produktionsbedingungen mit hinreichender Genauigkeit abgebildet werden konnten, sodass die Untersuchung nur teilweise als repräsentativ betrachtet werden kann. Jedoch weisen die Projektbetriebe eine große Heterogenität in Bezug auf bestimmte Merkmale wie die Betriebsgröße, das Einkaufsmanagement oder die Fütterung auf, weshalb gewisse Aussagen trotzdem pauschal für die deutsche Mastschweinehaltung getroffen werden können. Zudem ist das gewählte Kollektiv als typisch für den VzF bzw. Nordost-Niedersachsen anzusehen, und deckt dadurch einen großen Teil der deutschen Produktionsbedingungen ab. Dadurch weisen die ermittelten Ergebnisse ein hohes Maß an Generalisierbarkeit für Deutschland auf.

### 3.1.2 Aufbereitung von Sekundärdaten

Für das MulTiViS-Projekt lagen für ausgewählte Schweinemastbetriebe Sekundärdaten in Form von Produktionsdaten vor, die sich auf die Halbjahreszeiträume vom 1. Juli 2017 bis zum 30. Juni 2019 beziehen. Diese beinhalten folgende Teile:

- (1) Biologische und ökonomische Leistungsdaten aus der **Betriebszweiganalyse** des VzF, inklusive allgemeiner Betriebsdaten ("Schlüsselzahlen")

(2) Von QS angeforderte Anwendungs- und Abgabebelege zum **Antibiotikaeinsatz**

(3) **Schlachtbefunde** aus der routinemäßig durchgeführten Fleischuntersuchung

Die Eignung dieser Datenstrukturen zur Beschreibung von Tierwohl im Rahmen eines Monitorings wurde in den **Publikationen 1 und 2** geprüft, wobei die erste auf technisch-statistische Aspekte eingeht und die zweite den fachlich-inhaltlichen Bezug untersucht. Im Folgenden werden die Ergebnisse der Plausibilitätsprüfungen sowie die durchgeführten Aufarbeitungsschritte separat für die einzelnen Datenstrukturen erläutert.

Die Nutzung von Sekundärdaten ist zunächst grundsätzlich mit einem geringeren Aufwand an Ressourcen verbunden als die spezielle Erhebung von Primärdaten. Eine direkte Auswertung solcher Informationen ohne vorherige Aufarbeitung ist allerdings nicht angemessen, da die Nutzung von Daten, die zu einem anderen Zweck erfasst wurden, in der Regel intensiver Vorbereitung bedarf. Dazu gehört die fachliche sowie statistische Plausibilitätsprüfung, um die generelle Eignung der Daten festzustellen. Mit Hilfe deskriptiver Statistiken können einzelne Indikatorvariablen anhand ihrer Verteilungen beurteilt und anschließend gegebenenfalls transformiert, zusammengefasst oder anderweitig prozessiert werden.

So führen beispielsweise fehlende Daten (Non-Response) dazu, dass einige Betriebe für die Auswertungen ausgeschlossen werden. Dabei besteht die Gefahr, dass es sich um Betriebe handelt, die typisch für eine bestimmte Produktionsart sind und dadurch einen Selection Bias verursachen können, wie in Kapitel 3.1.1 beschrieben. Neben dieser Problematik spielt bei der Nutzung von Sekundärdaten aber vorwiegend die Verzerrung durch die genutzte Information eine entscheidende Rolle (Information Bias), denn der ursächliche Erhebungszweck und der Nutzungszweck sind zunächst verschieden und es stellt sich somit die Frage, ob die Daten für diese Art der Nachnutzung überhaupt geeignet sind.

#### *Betriebszweiganalyse des VzF*

Vorab ist darauf hinzuweisen, dass biologische und ökonomische Leistungsdaten bzw. Schlüsselzahlen, wie sie im Rahmen der Betriebszweiganalyse des VzF erhoben werden, nicht

pauschal für alle Betriebe in Deutschland vorliegen. Da die spezielle landwirtschaftliche Beratung, beispielsweise durch Erzeugerringe, freiwillig ist, werden diese Daten somit nicht standardisiert erfasst. Vor allem regionale Unterschiede spielen hier eine wichtige Rolle, da sich die Größe, Dichte und Struktur der schweinehaltenden Betriebe in den einzelnen Regionen Deutschlands zum Teil erheblich unterscheiden (Merle *et al.* 2012). Zudem fehlen für einige Indikatoren einheitliche Definitionen bei der Erfassung. So gibt es allein für die Ermittlung der Mortalität vielfältige Definitionen und Methoden der Erfassung und Dokumentation. Diese unterscheiden sich meist im Bezugszeitraum bzw. dem Referenzkollektiv und in der Hinzunahme von notgeschlachteten oder euthanasierten Tieren sowie anderen praktischen Aspekten. Nienhaus *et al.* 2020 zeigten anhand von aktuellen Untersuchungen, dass die Diskrepanz der Mortalitätsinformation dabei erheblich sein kann. Im Hinblick auf die Nutzung der Produktionsdaten für ein nationales Monitoring sind diese Aspekte daher besonders zu beachten.

Für das MULTiViS-Projekt lagen aus der Betriebszweiganalyse quantitative Variablen zu ökonomischer und biologischer Leistung (Mastplaner) und qualitative Variablen bezüglich baulicher Gegebenheiten und grundsätzlicher Managemententscheidungen vor (Schlüsselzahlen). Dabei fiel zunächst auf, dass die Bezugsgrößen im Mastplaner stark variieren. Manche der Werte beziehen sich auf Kilogramm Zuwachs oder verkaufte Tiere, während andere als absolute Zahlen in Euro oder Prozent angegeben werden. Daher wurden für eine generelle Vergleichbarkeit und gemeinsame Auswertung der Daten manche Variablen, zum Beispiel die Kosten für tierärztliche Behandlung, auf die Anzahl der verkauften Tiere bezogen, um eine betriebsübergreifende und größenunabhängige Bewertung zu ermöglichen.

Zudem ergab die fachliche Plausibilitätsprüfung, dass wichtige Leistungsgrößen wie zum Beispiel die täglichen Zunahmen oder die Mortalität entscheidend vom Einstallgewicht der Ferkel zur Mast abhängen (s.o. und **Publikation 1**). In deutschen Produktionssystemen liegt das durchschnittliche Einstallgewicht bei circa 24-32 kg. Einige Betriebe jedoch stallen sogenannte "Baby-Ferkel" mit circa 8-12 kg ein und andere, meist solche mit angegliederter

Ferkelproduktion, praktizieren eine längere Vormast und stallen dann Läufer mit > 30 kg ein. Hinter diesen Entscheidungen stehen unterschiedliche Management-Philosophien, die sich auf den Verlauf der weiteren Mast auswirken. So haben Baby-Ferkel eine bessere Futtermittelverwertung als ältere Schweine, verzeichnen jedoch absolut gesehen zunächst weniger tägliche Zunahmen (Wellock *et al.* 2004). Zudem sind sie aufgrund des noch schwächer ausgeprägten Immunsystems anfälliger für Krankheiten, was in höherer Mortalität resultieren kann. Die schweren Läufer hingegen sind den meisten Krankheiten im jungen Ferkelalter entwachsen und benötigen daher durchschnittlich weniger antibiotische Behandlungen. Die Gültigkeit dieser Thesen für das Studienkollektiv wurden durch statistische Tests bestätigt und daraufhin drei Einstallgewichts-Klassen definiert:

- (1) leicht: < 24 kg
- (2) mittel: 24–33,5 kg
- (3) schwer: > 33,5 kg

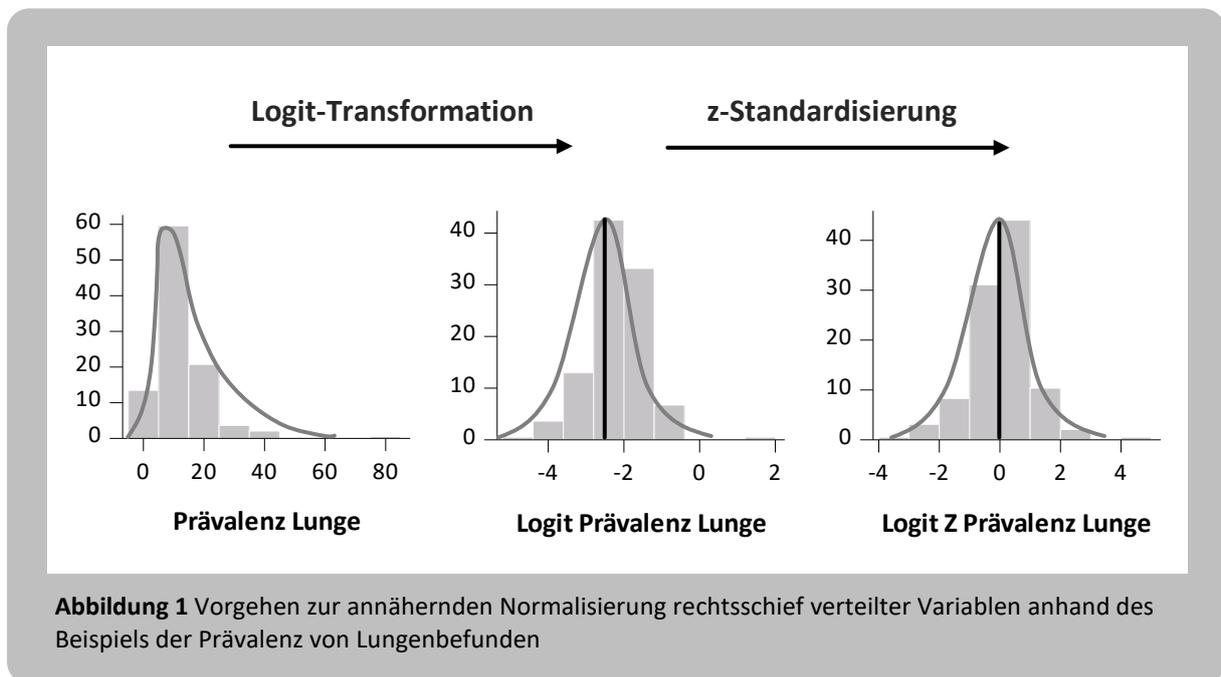
Die hier aufgeführten Grenzen entsprechen nicht der oben erwähnten typischen Verteilung und vor allem die Obergrenze der leichten Einstallgewichts-Klasse 1 (24 kg) liegt deutlich über dem oberen Durchschnittsgewicht der leichten Baby-Ferkel (12 kg). Dies ist der Tatsache zuzuschreiben, dass einige Betriebe des Untersuchungskollektivs innerhalb eines Halbjahres sowohl leichte als auch mittelschwere Schweine eingestallt haben und dadurch ein durchschnittliches Einstallgewicht ermittelt wird, das weder der einen, noch der anderen Klasse eindeutig zuzuordnen ist. Dieses Phänomen kann potentiell bei allen quantitativen Variablen auftreten und zu einer Verzerrung führen. Es entsteht dadurch, dass sich die Produktionsdaten jeweils auf Halbjahresräume beziehen und nicht auf einzelne Mastdurchgänge. Etabliert hat sich dieser Verallgemeinerungsschritt beispielsweise beim nationalen Antibiotikamonitoring (Deutscher Bundestag 2013). Für ein nationales Tierwohlmonitoring bestünde allerdings die Option, die Bezugszeiträume nicht standardisiert auf Halbjahresbasis festzulegen, sondern an die individuellen Mastgruppen anzupassen. Dies würde jedoch vermutlich mit verminderter Vergleichbarkeit einhergehen, da sich Dauer und Beginn bzw. Ende der Mastperioden von Betrieb zu Betrieb unterscheiden.

Im Rahmen der GAK-Maßnahmen (Gemeinschaftsaufgabe zur Verbesserung der Agrarstruktur und des Küstenschutzes) wurden von den Vereinigten Informationssystemen Tierhaltung (vit) in Zusammenarbeit mit den Erzeugerringen und der Landwirtschaftskammer Niedersachsen sogenannte Schlüsselzahlen erarbeitet. In einem Pflichtenheft (BRS 2020), das regelmäßig von einer Expertengruppe des Bundesverband Rind und Schwein e.V. (BRS) überprüft und aktualisiert wird, sind bestimmte Standards für die Schweinehaltung festgehalten. Da jedoch eine eindeutige Erfassung von Informationen und die Zuordnung in Kategorien teilweise nicht ohne weiteres möglich ist, wurde vom BRS der Wert "99" (indifferent) definiert. Dieser wird vergeben, wenn weniger als 90% der Eigenschaften eines Betriebes auf das entsprechende Kriterium zutreffen. Unter Praxisbedingungen zeigte sich, dass viele Betriebe eine gewachsene Struktur aufweisen und somit bei mehreren Schlüsselzahlen in die "99er"-Kategorie fallen. Daher muss generell überlegt werden, wie mit dieser Kategorie umgegangen wird. Handelt es sich um verhältnismäßig wenige Betriebe, besteht die Möglichkeit, diese für weitere Auswertungen auszuschließen. Aufgrund der Vielzahl an Schlüsselzahlen kann dies jedoch schnell in einer (zu) starken Reduktion des Gesamtkollektivs resultieren. Eine Alternative ist das Zusammenfassen von Kategorien, was allerdings durch das Einbüßen an Informationsschärfe zu verminderter fachlich-inhaltlicher Aussagekraft der Modellierung führt. Auch diese Problematik spricht daher dafür, dass für ein nationales Monitoring die spezifische Erfassung von Mastgruppen eher geeignet ist.

#### *Informationen zum Antibiotikaeinsatz von QS*

Im Gegensatz zu den Ergebnissen der Betriebszweiganalyse des VzF liegen die Informationen zum Antibiotikaeinsatz bundesweit für alle Schweinemastbetriebe vor, da die Dokumentation der Anwendung von antimikrobiellen Substanzen in der Nutztierhaltung gesetzlich vorgeschrieben ist (Deutscher Bundestag 2013; Europäische Kommission 2019). Im Rahmen der Qualitätssicherung durch QS, welche circa 95% aller Betriebe in Deutschland umfasst (QS 2019), ist zudem eine annähernd standardisierte und harmonisierte Datenerhebung garantiert. Jedoch sind auch für diese Daten Plausibilitätsprüfungen zwingend erforderlich, da Eingabe- oder Übertragungsfehler vorkommen können und sie gewisse Problematiken

bereithalten können. So zeigten die entsprechenden Histogramme und Boxplots, dass einzelne Variablen nicht normal, sondern rechtsschief verteilt sind. Ursache hierfür ist das gehäufte Vorkommen von niedrigen Werten bzw. Nullwerten. Zur besseren statistischen Auswertung wurden die Daten deshalb durch Logarithmierung bzw. Logit-Transformation und anschließende z-Standardisierung in eine annähernde Normalverteilung überführt (s. Abbildung 1).

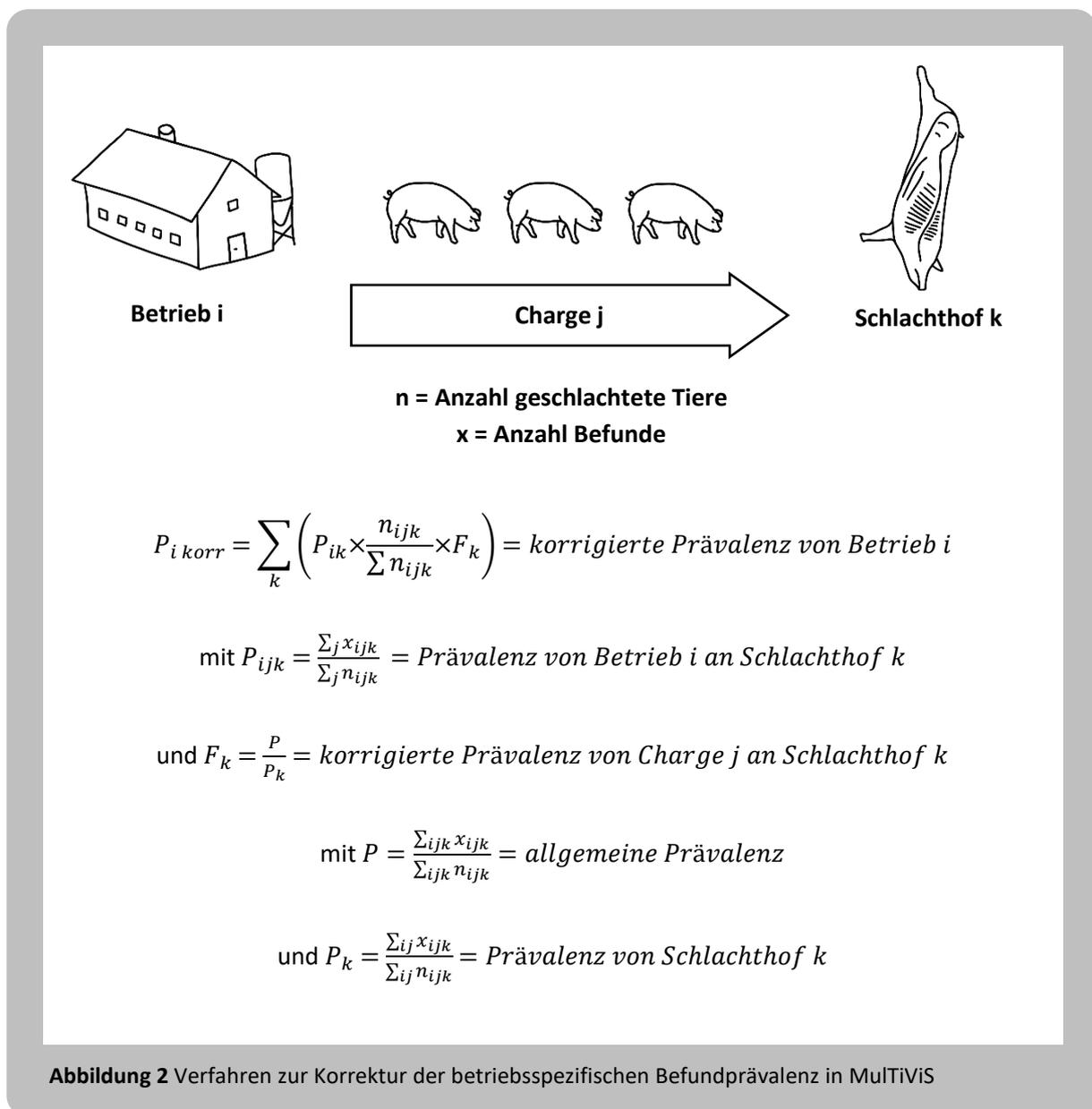


Zur Beschreibung des betriebspezifischen Antibiotikaverbrauchs hat QS einen Therapieindex definiert, der weitestgehend konform ist mit der bundesweit einheitlich festgelegten Methode zur Berechnung der Therapiehäufigkeit im Rahmen der 16. AMG-Novelle (BMELV 2013; Deutscher Bundestag 2013; Hemme *et al.* 2017; QS 2018). Diese Formel bezieht sich auf die durchschnittlich gehaltenen Tiere pro Halbjahr, wobei diese Angabe allerdings im Rahmen des Projektes nicht zur Verfügung stand und daher durch die Anzahl der Mastplätze ersetzt wurde.

### *Schlachtbefunde von QS*

Auch die Erfassung von tierwohlrelevanten Befunden am Schlachthof ist gesetzlich vorgeschrieben (Deutscher Bundestag 2013; Europäische Kommission 2019). Für QS-Betriebe erfolgt zudem seit 2018 die Erhebung im Rahmen der routinemäßigen Fleischuntersuchung nach einem einheitlichen Schema, doch garantiert dies keine schlachthofübergreifende, harmonisierte Datenstruktur. Die Umstellung der Erhebungsmethoden bzw. die Vorgabe neuer Befunde im Untersuchungsplan erfordern umfassende Schulungen des Schlachthofpersonals. Dieser Prozess wurde während der Projektlaufzeit eingeleitet und ist noch nicht überall vollständig umgesetzt, wodurch eine gewisse Unschärfe und mangelnde Validität der Daten zu erklären ist.

Zudem wird die generelle Nutzbarkeit der Befunddaten sowohl in der Praxis als auch in der Wissenschaft vielfach diskutiert und infrage gestellt (Deutscher Bundestag 2017; Hoischen-Tauber *et al.* 2011). Aufgrund mangelnder Standardisierung und Harmonisierung der Erfassungsmethoden bestehen zum Teil erhebliche Unterschiede in den Prävalenzniveaus zwischen den einzelnen Schlachtstätten, was als sogenannter "Schlachthof-Effekt" bezeichnet wird. Um diesem entgegenzuwirken, wurde für das MultiViS-Projekt das Verfahren von Nienhaus *et al.* 2020 angepasst, indem eine nach Schlachthoflieferungen adjustierte und gewichtete Prävalenz pro Betrieb und Halbjahr ermittelt wurde (s. Abbildung 2).



Kleinere Chargen mit wenigen Tieren fallen dadurch weniger ins Gewicht und die unterschiedlichen Prävalenzlevel an den verschiedenen Schlachthöfen können ausgeglichen werden. Allerdings handelt es sich hierbei um eine theoretische, künstlich ermittelte Größe, da die "wahre Prävalenz" aufgrund der Tatsache, dass jedes Schwein nur einmal am Schlachthof untersucht wird, nicht exakt bestimmt werden kann. Daher können die so ermittelten Werte nur zu Vergleichen im Kollektiv aber nicht als absolute Messgrößen verwendet werden. Dies ist ebenfalls ein übliches Verfahren in der allgemeinen

demoskopischen Standardisierung (Kreienbrock *et al.* 2012).

### 3.1.3 Erhebungskonzept

#### *Genereller Aufbau*

Zur Klärung der Frage, ob bereits vorhandene Produktionsdaten allein für eine Tierwohlbeurteilung ausreichend sind oder eine spezielle Untersuchung im Betrieb notwendig ist, wurde eine übergreifende Bewertung durchgeführt. Für eine gute Vergleichbarkeit sind dazu ähnliche Datenstrukturen bezüglich Umfang, Bezugsgröße bzw. -zeitraum und fachlicher Definition notwendig. Die Produktionsdaten beziehen sich jeweils auf alle in einem Betrieb und Halbjahr verkauften Schweine. Für eine kleine, einstellige Anzahl von Projektbetrieben wäre es möglich und vom Aufwand vertretbar, die gleichen Schweine nochmals speziell zu untersuchen. Allerdings erfordert dies so viel Zeit und Aufwand, dass eine größer angelegte Studie oder ein Monitoringprogramm unter Praxisbedingungen dies nicht leisten könnten. Um jedoch aussagekräftige Ergebnisse zu erhalten, ist ein gewisser Studenumfang unerlässlich. Bei der Planung einer Erhebung tritt also der Konflikt zwischen Praxistauglichkeit bzw. finanziellen Möglichkeiten und der statistischen sowie fachlichen Aussagekraft auf. Eine mögliche Lösung dieser Problematik ist die Untersuchung einer beschränkten Zufallsauswahl von Tieren an einem oder mehreren Stichtagen, welche auch in der vorliegenden Studie angewendet wurde. Bei über 200 Projektbetrieben dauerte die Erhebungsphase mit nur einem Besuch pro Betrieb bereits ein ganzes Jahr. Limitierende Faktoren waren hierbei die zum Teil sehr weiten Anfahrtswege sowie die Abstimmung von Terminen mit den Tierhalter:innen im normalen Betriebsalltag bzw. während spezieller Situationen wie zum Beispiel der Ernte.

Um eine aussagekräftige aber überschaubare Anzahl von Schweinen für die Untersuchung auszuwählen, ist ein standardisiertes Verfahren notwendig. Dazu bietet sich der sogenannte Schwedenschlüssel an (Kish 1949). Dieser ermöglicht eine zufällige aber dennoch standardisierte Auswahl von Buchten anhand einer vorgefertigten Tabelle mit Zufallszahlen.

Allerdings ist eine praktische Umsetzung teilweise technisch erschwert, da Betriebe durch ihre gewachsenen Strukturen eine sehr heterogene Bauweise bzw. Aufteilung von Abteilen und Buchten in den Ställen aufweisen können. Deswegen kann ein solches Verfahren keine vollständig standardisierte Zufallswahl garantieren und muss im Einzelfall durch das Untersuchungsteam den örtlichen Gegebenheiten angepasst werden.

Außerdem ist zu beachten, dass die Zufallsauswahl von bei der Erhebung untersuchten Tieren nicht zwangsläufig eine Teilmenge der Tiere sein muss, auf die sich die Produktionsdaten beziehen. Die Grundlage für den Berechnungszeitraum der Produktionsdaten bildet das Verkaufsdatum (Schlachtdatum). Wenn also zum Beispiel die Erhebung am 15. Juni stattgefunden hat, wurden die Produktionsdaten aus dem ersten Halbjahr (1. Januar bis 30. Juni) zugeordnet. Handelte es sich bei den untersuchten Tieren allerdings um junge Schweine die erst im zweiten Halbjahr verkauft wurden, sind diese nicht durch die Produktionsdaten aus dem ersten Halbjahr abgebildet. Es kann jedoch davon ausgegangen werden, dass in der Regel zwischen den Durchgängen keine erheblichen Unterschiede im Gesundheitsstatus und vor allem bei den umgebenden Einflussfaktoren auftraten, sodass die Stichprobe trotzdem als angemessen repräsentativ (für den einzelnen Betrieb) angesehen werden kann. Umgehen könnte man dieses Phänomen durch die genaue Erfassung des Einstalldatums bzw. des voraussichtlichen Verkaufsdatums und eine entsprechende Korrektur der Zuordnung des Produktionshalbjahres. Allerdings ist eine solch exakte Zuordnung in der betrieblichen Praxis schwer in eine Erfassungsroutine umzusetzen.

Ein weiterer Aspekt bei einmaligen und an einem zufälligen Datum stattfindenden Betriebsbesuchen ist der Effekt der Jahreszeiten. Trotz vorrangiger Stallhaltung können extreme Temperaturen einen erheblichen Einfluss auf das Wohlbefinden der Schweine haben. Dazu zählen starke Hitze im Sommer, bei der gängige Kühlungsmaßnahmen nicht mehr ausreichen, niedrige Temperaturen im Winter, bei denen nicht genügend geheizt wird oder Übergangszeiten mit extremen Temperaturschwankungen zwischen Tag und Nacht. Doch auch hier ist es in der Praxis nahezu unmöglich, eine große Anzahl von Betrieben unter ähnlichen klimatischen Bedingungen zu besuchen. Da sich im Falle von MuTiViS die Erhebung

jedoch über den Zeitraum eines ganzen Jahres erstreckt hat, verliert dieser Effekt statistisch gesehen dadurch an Bedeutung, dass die Betriebsbesuche sich auf die verschiedenen Wetterlagen verteilen und Extreme ausgeglichen werden. Für zukünftige Monitoringprogramme sollte jedoch – analog zur Standardisierung zwischen den Schlachthöfen – eine Korrektur von Variablen vorgenommen werden, wenn bei diesen eine saisonale Abhängigkeit unterstellt werden kann.

### *Erhebungsbogen*

In den letzten Jahren haben sich bereits viele verschiedene Methoden zur Erfassung und Beschreibung von Tierwohl in Schweinemastbetrieben etabliert. Am bekanntesten und häufigsten angewendet ist dabei wohl das Animal Welfare Pig Protocol (Welfare Quality 2009). Dieses umfasst einen Katalog von 30-50 Messgrößen, die in drei Schritten zu einer Gesamtbeurteilung zusammengefasst werden. Die vollständige Bearbeitung dieses Protokolls ist jedoch sehr umfangreich und wird mit fünfeinhalb Stunden veranschlagt, sodass dies für ein nationales Monitoring ungeeignet erscheint. Zur besseren Umsetzbarkeit und Praxistauglichkeit für das MulTiViS-Projekt wurde daher eine ausführliche Literaturrecherche in nationalen und internationalen Quellen durchgeführt (Bartussek 1995; Dickhaus *et al.* 2007; Schrader *et al.* 2016; Welfare Quality 2009). Auf Basis dieser Vorauswahl von Tierwohlintikatoren und Einflussfaktoren wurde durch die Expertise der Projektpartner:innen sowie exemplarische Anwendung in einigen Testbetrieben ein reduziertes Untersuchungsprotokoll erstellt, welches in maximal zweieinhalb Stunden bearbeitet werden kann. Im Rahmen eines abschließenden Testmonitorings wurde dieses Programm erneut reduziert, sodass nur noch maximal eine Stunde Bearbeitungszeit erforderlich ist. Die finale Version des Erhebungsbogens umfasst drei Teile:

#### **(1) Interviewbogen**

Für einen generellen Überblick über den Betriebsaufbau und Informationen zu Management, Hygiene, Tierbetreuung und bauliche Gegebenheiten der Schweinemastställe.

**(2) Abteilmogen**

Tierwohlrelevante Faktoren aus der Tierumgebung, die nicht nur einzelne Buchten, sondern das ganze Abteil betreffen.

**(3) Buchtenbogen**

Informationen zur Buchtenausstattung und zur Gesundheit der darin gehaltenen Tiere durch Erfassung buchtenspezifischer Prävalenzen von Schweinen mit abweichenden Befunden.

Dabei war für die Selektion von Erhebungsvariablen zunächst die grundsätzliche Zielsetzung zu klären, ob die Validierung der bereits vorhandenen Produktionsdaten oder eine komplementäre Erfassung zur ergänzenden Untersuchung durchgeführt werden soll. Bei MulTiViS handelt es sich um eine Mischform mit sowohl äquivalenter als auch ergänzender Erfassung von Informationen in Anlehnung an die Produktionsdaten. Dabei wurden hauptsächlich die Einflussfaktoren aus der Tierumgebung komplementär zu den vorhandenen Informationen aus der Betriebszweiganalyse gewählt, da hier die Erfassung durch die Beraterinnen und Berater als valide gilt und so der zusätzliche Erhebungsaufwand möglichst geringgehalten werden kann. Im Hinblick auf die Plausibilitätsprüfung sollten allerdings einige ausgewählte Variablen gezielt doppelt erfasst werden. Dafür bietet sich zum Beispiel an, die Anzahl der Mastplätze oder die Information, ob Eber gemästet werden, zu verwenden.

In Bezug auf die Tierwohlintikatoren war zudem eine äquivalente Erhebung teilweise gar nicht möglich, wenn es um die Bewertung der Lungen oder anderer innerer Organe geht, wie sie am Schlachthof erfolgt. Daher musste hier auf alternative Indikatoren zurückgegriffen werden, wie beispielsweise den Hustenindex als stellvertretenden Indikator für die Lungengesundheit. Äußere Befunde hingegen, wie zum Beispiel Verletzungen an Ohren und Schwänzen, liegen in Form von Schlachtbefunden vor und können unmittelbar am lebenden Tier erfasst werden. Ein Vergleich der Prävalenzlevel, also des durchschnittlichen Anteils von auffälligen Befunden, zeigte jedoch, dass hier teilweise deutliche Unterschiede zwischen der Erhebung im Projekt und in den Schlachthöfen vorliegen. Ein Grund für das geringe Maß an Übereinstimmung ist in den zum Teil abweichenden Definitionen für auffällige Befunde sowie

der Erfassungsgenauigkeit zwischen den Schlachthöfen und der Erhebung im Projekt zu vermuten. Bei der Entwicklung des Erhebungsbogens wurde der Ansatz verfolgt, ein möglichst detailliertes Untersuchungsprotokoll zu erstellen, um eine aussagekräftige und umfassende Tierwohlbeurteilung vornehmen zu können. So wurden zum Beispiel die Verletzungen an den Ohren der Tiere über vier einzelne Indikatorvariablen beschrieben (Blutohren, Ohrrandnekrosen, Bissspuren und Substanzverlust). Im Laufe der Erhebung stellte sich jedoch für einige der Tierwohlindikatoren heraus, dass die genaue tierärztliche Analyse sehr zeitaufwändig und daher diese Art der Unterscheidung von Einzelbefunden unter Praxisbedingungen teilweise nicht möglich ist. Die deskriptive Statistik zeigte außerdem, dass bestimmte Indikatoren oder Einflussfaktoren zwar wichtig für eine Tierwohlbeurteilung sind, aber so selten vorkommen, dass eine statistisch aussagekräftige Auswertung nicht möglich ist.

Die aufgeführten Problematiken verdeutlichen das bestehende Defizit bei der standardisierten Erfassung von Tierwohlindikatoren und die Notwendigkeit der Entwicklung eines praxisnahen Erhebungskonzepts. Insgesamt kann jedoch die geringe Similarität der erhobenen Tierwohlindikatoren mit den Schlachtbefunden als Informationsgewinn angesehen werden, der die Wichtigkeit einer zusätzlichen (tierärztlichen) Erhebung im Betrieb unterstreicht.

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

Im Folgenden wird thematisiert, wie die beschriebenen Projektdaten weiter aufbereitet und in welcher Form sie ausgewertet wurden. Grundlegend dafür ist ein relatives Benchmarking-System, welches auf dem Konzept der z-Standardisierung ausgewählter Indikatoren zu Tiergesundheits-Scores beruht. Anschließend wird beschrieben, inwiefern bestimmte Einflussfaktoren als Präventionsmaßnahmen dienen können und es erfolgt die Gegenüberstellung von Tiergesundheits-Scores am Schlachthof und äquivalente erfassten Tierwohlindikatoren in den Betrieben.

### 3.2.1 Festlegung der Tierwohlindikatoren und Transformation zu Tiergesundheits-Scores

Um ein Tierwohlmonitoring auf Basis bestehender Daten durchzuführen, müssen aus dem umfassenden Set an Variablen zunächst die zu untersuchenden Zielgrößen, d.h. Indikatoren zur Bewertung des Tierwohls in einem Betrieb, ausgewählt werden. Zur Übersicht wurde daher im Rahmen der vorliegenden Studie eine umfangreiche Liste mit allen für das Projekt verwendeten Indikatoren und Einflussfaktoren angelegt. Wie bereits in der Einleitung erläutert und durch die umfassenden Ergebnisse der Literaturrecherche für den Erhebungsbogen bestätigt, gibt es eine Vielzahl an potentiellen Tierwohlindikatoren. Welche von diesen nun die wichtigsten sind, ist jedoch nicht eindeutig festgelegt. Als Basis für die MulTiViS-Studie wurden die vom Kuratorium für Bau und Technik in der Landwirtschaft (KTBL) fortlaufend zu erhebenden und halbjährlich auszuwertenden Indikatoren zur Eigenkontrolle beim Schwein ausgewählt (Schrader *et al.* 2016), da diese als ein entwickelter Standard in Deutschland angesehen werden können. Zu diesen Indikatoren gehören die Tierverluste (Mortalität), tägliche Zunahmen, Therapiehäufigkeit (Einsatz von Antibiotika) und die Schlachtbefunde. Zusätzlich wurde die Futtermittelverwertung mit aufgenommen, da diese ebenfalls als wichtiger Indikator gilt und vom Projektteam als notwendige Information erachtet wurde. Dieses Variablen-set enthält somit eine umfassende Auswahl von direkten und indirekten Indikatoren der Tiergesundheit, welche sich vollständig mit den für das Projekt vorhandenen Daten abbilden lassen.

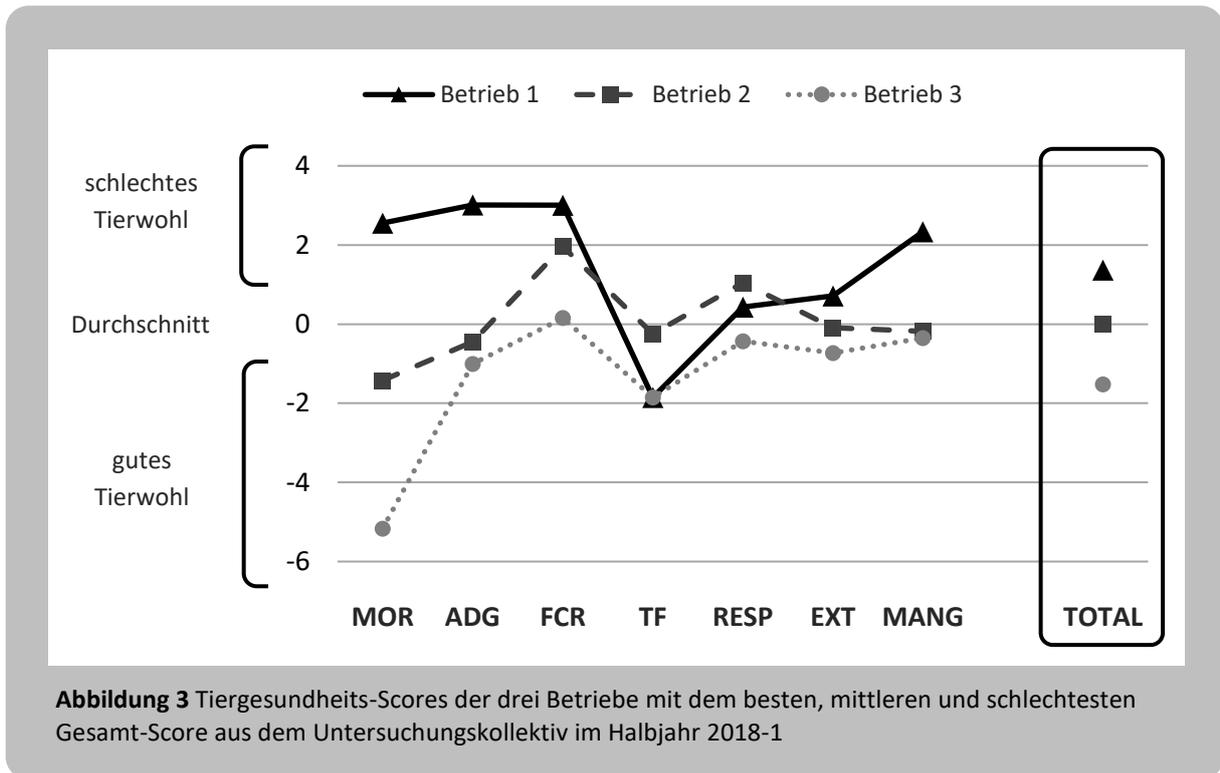
Wie bereits in Kapitel 3.1.2 erwähnt, werden die Zielgrößen in verschiedenen Einheiten gemessen, liegen also auf unterschiedlichen Skalen und sind größtenteils nicht normalverteilt. Zudem gibt es in der Literatur und in der Praxis keine einheitliche Definition von Cut-Off-Werten, die zur individuellen Beurteilung der Betriebe herangezogen werden können. Die Einführung solcher Cut-Offs birgt zudem ein hohes Potential für einen Diskurs zwischen den verschiedenen Stakeholdern. Gründe hierfür sind zum Beispiel die starken regionalen Unterschiede in den Tierhaltungssystemen oder die unterschiedlichen Fachkontexte, in denen die Leistungsgrößen beurteilt werden.

Daher bietet es sich an, die Indikatoren anhand der in Kapitel 3.1.2 und **Publikation 1**

beschriebenen Methodik (Logit-Transformation und z-Standardisierung) für ein relatives Monitoring vorzubereiten. Dabei stehen negative Score-Werte für eine gute und positive Score-Werte für eine schlechte Tiergesundheit. Durch die einheitslose z-Skala erfolgt keine absolute Bewertung der einzelnen Tierwohlindikatoren, sondern das Kollektiv der Betriebe wird in eine Reihenfolge gebracht und somit die individuelle Leistung stets im Gesamtkontext der Gruppe von untersuchten Betrieben betrachtet. Bekannt ist dieses Prinzip bereits aus dem nationalen Antibiotikamonitoring (Deutscher Bundestag 2013) sowie der Publikation von Nienhaus *et al.* 2020. Vorteil dieser Methode ist der ständige Ansporn zur Verbesserung, da sich der individuelle Score-Wert nicht nur bei Verminderung des Tierwohls auf dem eigenen Betrieb, sondern auch durch Verbesserung der anderen Betriebe bei gleichbleibendem Status des eigenen Betriebs verändern kann. Dies führt langfristig zu einer absolut gesehenen Weiterentwicklung der Gesamtqualität des Kollektivs, die allerdings nicht mit einer besseren Leistung des einzelnen Betriebes einhergehen muss. Ein großer Nachteil dieser Methode – welcher jedoch gleichzeitig auch als Vorteil gesehen werden kann – ist die Tatsache, dass unterschiedliche Einheiten und Größenordnungen der einzelnen Indikatoren verschleiert werden. Dies ist vor allem für die Schlachtbefunde von Bedeutung, bei denen sich die einzelnen Befunde auf zum Teil sehr unterschiedlichen Prävalenzniveaus bewegen. So kommen zum Beispiel die respiratorischen Indikatoren wesentlich häufiger vor als Läsionen an den Ohren. Durch die z-Standardisierung werden diese auf eine einheitliche Skala gesetzt und die Unterschiede sind nicht mehr erkennbar. Andererseits ermöglicht dies jedoch erst die Zusammenfassung der Einzelbefunde zu Teil-Scores, wie im folgenden Abschnitt beschrieben, sowie den anschließenden Vergleich mit Tiergesundheits-Scores aus anderen Bereichen.

Um die Komplexität des Indikatoren-Sets zu reduzieren und damit praxistauglicher bzw. leichter interpretierbar zu machen, wurden die 13 Schlachtbefunde des QS-Erhebungsprotokolls zusammengefasst. Mittels Expertengewichten von Nienhaus *et al.* 2020 erfolgte die Reduktion zu den drei Schlacht-Scores RESP (respiratorische Gesundheit), EXT (äußere Läsionen) und MANG (Tiermanagement). Außerdem wurde ein gewichteter Gesamt-Score (TOTAL) berechnet, welcher neben den drei Schlacht-Scores vier weitere Tiergesundheits-Scores enthält, nämlich MOR (Mortalität), ADG (tägliche Zunahmen), FCR

(Futterverwertung) und TF (Therapiehäufigkeit). Dieser Gesamt-Score kann für eine grobe Einschätzung genutzt werden. Allerdings zeigte sich, dass durch eine Mittelung die Gefahr besteht, dass gute oder schlechte Einzelwerte ausgeglichen werden und somit die Gesamtbeurteilung einzelne Problembereiche des Tierwohls verdeckt (s. Abbildung 3).



Am Beispiel der drei ausgewählten Betriebe in Abbildung 3 kann man erkennen, dass Betrieb 1 trotz einer hohen Mortalität einen Gesamt-Score hat, der nur leicht schlechter als der Durchschnitt ist. Grund hierfür ist die sehr geringe Therapiehäufigkeit, die auf dem gleichen niedrigen Niveau liegt wie bei dem Betrieb mit dem besten Gesamt-Score. Dieses Phänomen unterstreicht die oben beschriebene Problematik der "Ausmittlung" von Einzelleistungen. Aus **Publikation 1** folgt genau diese Notwendigkeit, neben einem Gesamt-Score stets auch die einzelnen Tiergesundheits-Scores für eine nachhaltige Tierwohlbeurteilung zu berücksichtigen. Außerdem beleuchtet das Beispiel von Betrieb 1 in Abbildung 3 eine fragliche inhaltlich-fachliche Aussagekraft der Therapiehäufigkeit. So kann der geringe Einsatz von

Antibiotika zwar einerseits für ein umsichtiges Betriebsmanagement und eine gute Tiergesundheit stehen, andererseits aber auch für die unzureichende Behandlung kranker Tiere, was meist in einer hohen Mortalität resultiert. Hinzu kommt die teils schwierige Interpretierbarkeit von biologischen Leistungsgrößen wie zum Beispiel den täglichen Zunahmen. Es steht außer Frage, dass ein schwer krankes Schwein keine hohen täglichen Zunahmen erbringen kann. Jedoch lässt sich im Umkehrschluss nicht allein durch gute tägliche Zunahmen auf ein gutes Tierwohl schließen.

### 3.2.2 Untersuchung potentieller Einflussfaktoren aus den Produktionsdaten

Die Daten aus der Betriebszweiganalyse des VzF enthalten ausführliche Informationen über die Betriebe und deren Management, wodurch die Frage entsteht, wie sich diese als Einflussfaktoren auf das Tierwohl auswirken und in welcher Form sie als potentielle Präventionsmaßnahmen dienen können. Dazu wurden in **Publikation 2** die Tiergesundheits-Scores durch die potentiellen Einflussfaktoren aus den Produktionsdaten modelliert. In vorbereitenden Assoziations- bzw. Korrelationsprüfungen zur Vermeidung von Confounding konnten bereits zwei Variablen aussortiert werden. Die anschließend durchgeführten ein- und mehrfaktoriellen Varianzanalysen zeigten, dass tatsächlich keine starke Verzerrung auftritt, da sich zwar die Anzahl der signifikanten Einflussfaktoren verringert, sich aber die Schätzer nicht wesentlich geändert haben. Am auffälligsten war die Tatsache, dass sich lediglich die beiden Tiergesundheits-Scores ADG und FCR zu täglichen Zunahmen und der Futterverwertung mit einem Bestimmtheitsmaß von über 0,64 modellieren ließen. Dieser Wert kann als Untergrenze für ein gut erklärtes Modell, also einen hohen Erklärungswert der Einflussvariablen auf die Variabilität der Zielvariable, interpretiert werden (Fahrmeir *et al.* 2016). Bei den täglichen Zunahmen und der Futterverwertung handelt es sich um indirekte Tierwohlindikatoren und die im Endmodell verbleibenden Einflussvariablen "Futter/Tier/Tag", "Futterenergie/kg Zuwachs" und "Produktionskosten/kg Zuwachs" sind quantitative Messgrößen mit eher wirtschaftlichem als biologischem Hintergrund. Dies deutet darauf hin, dass tiernahe Indikatoren wie zum Beispiel die Schlacht-Scores oder die Mortalität betriebsindividuellen

Schwankungen unterliegen, die sich nicht gut durch die untersuchten Einflussfaktoren beschreiben lassen. Dabei spielt sicherlich auch die strukturelle Heterogenität der Betriebe eine wichtige Rolle, da die Faktoren, welche ein Krankheitsgeschehen oder Verletzungen verursachen, stark von den einzelnen Managemententscheidungen oder baulichen Gegebenheiten abhängen.

Es bleibt also festzuhalten, dass sich die biologische Leistung der Tiere gut durch andere wirtschaftliche Messgrößen beschreiben lässt. Allerdings ist es anhand der vorliegenden Daten und Ergebnisse nicht möglich, eindeutige und universell gültige Präventionsmaßnahmen zur Verbesserung des Tierwohls zu nennen, da die Produktionsvielfalt hier erhebliche Unterschiede in den Rahmenbedingungen vorgibt, die in jedem Falle bei der Bewertung zu berücksichtigen sind. Im Zuge landwirtschaftlicher oder tierärztlicher Beratung müssen also auch die betriebsspezifischen Merkmale genauer betrachtet und spezifisch bewertet werden.

### 3.2.3 Gegenüberstellung der Indikatoren aus Produktions- und Erhebungsdaten

Zur Beantwortung der Frage, ob allein die Produktionsdaten ausreichend sind, um eine aussagekräftige Tierwohlbeurteilung der Betriebe vornehmen zu können, ist eine übergreifende Auswertung mit den Informationen aus der Erhebung erforderlich. Wie bereits in Kapitel 3.1.3 erläutert, liegt ein Hauptproblem dieser Gegenüberstellung in der unterschiedlichen Bezugsgröße. Auf der einen Seite stehen die Daten aller verkauften Tiere aus einem Halbjahr und auf der anderen Seite die Informationen von Schweinen aus lediglich acht zufällig ausgewählten Buchten an einem Stichtag in dem entsprechenden Halbjahr. Aufgrund der Tatsache, dass hier also nicht identische Grundgesamtheiten von gehaltenen Tieren verglichen werden, ist per se von einer eingeschränkten Übereinstimmung auszugehen. Dies hat sich in der vorliegenden Studie für die meisten der untersuchten Gesundheitsvariablen bestätigt. So lagen in **Publikation 2** die Korrelationskoeffizienten der Schlachtbefunde mit den entsprechenden Erhebungsindikatoren bei maximal 0,33 und die

Bestimmtheitsmaße der einfaktoriellen Modelle alle auf einem niedrigen einstelligen Niveau.

Hinzu kommt die Tatsache, dass bei der Erhebung nicht alle aus den Produktionsdaten definierten Tierwohlintikatoren bzw. Tiergesundheits-Scores äquivalent erhoben wurden. Diejenigen, die sich auf die biologische Leistung beziehen (tägliche Zunahmen und Futtermittelverwertung) werden nicht direkt am Tier gemessen, sondern geben nur indirekten Aufschluss zum Tierwohl. Die Mortalität gilt zwar als direkter Indikator, ist jedoch nicht einem konkreten Stichtag zuzuordnen, sondern nur über einen Zeitraum berechenbar. Die Therapiehäufigkeit ist ebenfalls ein indirekter Indikator, der an der Tiergruppe gemessen wird. Da eine eindeutige Interpretation, wie oben erwähnt, nur eingeschränkt möglich ist, ist sie zudem nur in Zusammenhang mit anderen Größen, wie zum Beispiel der Mortalität zu beurteilen.

Somit bleiben als direkt vergleichbare Indikatoren diejenigen, welche auf den Schlachtbefunden beruhen. In **Publikation 2** werden zunächst die einzelnen Indikatoren aus der Erhebung auf Zusammenhänge mit den inhaltlich passenden Schlachtbefunden geprüft. Dabei wurde die These bestätigt, dass vor allem die äußeren Verletzungen eine verhältnismäßig hohe Übereinstimmung zwischen Erhebung im Betrieb und am Schlachthof aufweisen. Jedoch zeigten sich hier auch einige Zusammenhänge, die fachlich nicht plausibel erscheinen und erwartete Übereinstimmungen konnten teilweise nicht bestätigt werden. So war zum Beispiel kein eindeutig signifikanter Zusammenhang des Hustenindex zum respiratorischen Tiergesundheits-Score RESP darstellbar.

Laut den Ergebnissen aus **Publikation 2** verbleiben sieben Indikatoren, die keinen der drei Schlacht-Scores signifikant beschreiben können und somit den Schluss nahelegen, diese als zusätzliche Indikatoren in zukünftige Standarderhebungsprotokolle aufzunehmen. Allerdings ist die Belastbarkeit dieser Folgerung eingeschränkt, da der Erklärungswert dieser Variablen insgesamt auf einem sehr niedrigen Niveau liegt.

### 3.3 Erforderliche Schritte für ein nationales Tierwohlmonitoring

Die grundsätzlichen Analysen und diskutierten Ergebnisse können als wesentlicher Schritt für die Entwicklung eines nationalen Tierwohlmonitoring angesehen werden. Bevor man ein solches Programm aber umsetzt, bleiben drei näher zu untersuchende Aspekte, die trotz des umfangreichen Studiendesigns von MuTiViS und der ausführlichen Analysen noch nicht abschließend geklärt werden konnten.

#### *Verfügbarkeit und Aufbereitung der Daten*

Hier ist zunächst die Verfügbarkeit der Produktionsdaten aus der Betriebszweiganalyse zu nennen. Wie in der Einleitung und Kapitel 3.1.2 erläutert, ist zwar eine bundesweit einheitliche und verpflichtende Erfassung von wichtigen Tierwohlindikatoren wie der Mortalität oder den täglichen Zunahmen gefordert, jedoch steht hierzu noch kein verbindliches Konzept zur Verfügung. Bis ein solches umgesetzt und fest etabliert ist, müsste man zunächst auf andere Quellen zurückgreifen und lückenhafte Datengrundlagen in Kauf nehmen.

Generell ist unbedingt darauf hinzuweisen, dass für ein nationales Monitoring oder eine nationale Tierwohlbank eine automatisierte Erfassung der benötigten Daten nicht ausreichend ist, sondern diese umfassend aufbereitet werden müssen. Die vorliegende Arbeit hat gezeigt, dass umfangreiche Plausibilitätsprüfungen und Transformationsschritte notwendig sind, um die Sekundärdaten für diese Nutzung vorzubereiten.

#### *Qualität der Schlachthofdaten*

Ein weiterer Punkt ist die Qualität der Schlachthofbefunde. Aufgrund der mangelhaften Harmonisierung bei der Erfassung sind diese für belastbare Auswertungen bislang nur eingeschränkt nutzbar. Es ist daher zwingend erforderlich, durch qualitätsverbessernde Maßnahmen eine bessere Nutzbarkeit zu erzielen. Neben ausführlichen Schulungen des untersuchenden Personals gehört hierzu die Überarbeitung des Merkmalskatalogs der

Schlachttier- und Fleischuntersuchungsstatistik (Koch 2016; Wanda *et al.* 2012). Dabei sind unter anderem die Unterschiede bei den praktischen und technischen Gegebenheiten vor Ort zu berücksichtigen. Da jeder Schlachthof individuell aufgebaut ist, ist zum Beispiel die Festlegung einer fixen Position der untersuchenden Personen nicht unbedingt möglich. Zudem ist für eine einheitliche und konsistente Befundermittlung die detaillierte Beschreibung der zu erhebenden Merkmale in Form einer leicht verständlich aber klar definierten Anleitung erforderlich. Außerdem ist für ein betriebsspezifisches Monitoring die Abgrenzung von Befunden, die auf den Transport zur Schlachtstätte zurückzuführen sind, notwendig. So sind zum Beispiel die Treibespuren als einer der 13 standardmäßig von QS erfassten Befunde nicht eindeutig dem Betriebsmanagement zuzuordnen und daher im Vergleich zu den anderen Tiergesundheits-Scores nur bedingt als gleichwertiges Mittel zur Tierwohl-Kategorisierung geeignet.

Nur durch Umsetzung dieser Punkte kann eine harmonisierte und vor allem vergleichbare Erfassung von tierwohlrelevanten Befunden über alle Schlachthöfe in Deutschland ermöglicht werden.

#### *Praxistauglichkeit und Durchführung der Datenerfassung*

Außerdem sollte durch weitere Untersuchungen geklärt werden, welche der potentiell infrage kommenden Tierwohlindikatoren aus der Erhebung tatsächlich in ein allgemeines Untersuchungsprotokoll aufgenommen werden. Dabei ist die Praxistauglichkeit der Erfassung und die Validierung durch ein erweitertes Studienkollektiv zu berücksichtigen. Sind die finalen Indikatoren ausgewählt, bleibt zu entscheiden, in welcher Form diese zukünftig erfasst werden sollen. Neben der gesetzlich vorgeschriebenen Eigenkontrolle durch die Tierhalter:innen könnte die Erhebung durch (landwirtschaftliche) Berater:innen bei Routinebesuchen auf den Betrieben durchgeführt werden und die ausgewählten Indikatoren als eine Art zusätzlicher Schlüsselzahlen erfasst werden, wie sie bereits vom BRS vorgegeben sind. Eine weitere Möglichkeit ist das Hinzuziehen von Tiermediziner:innen zur gezielten Untersuchung der Tiere bezüglich möglicher Tierwohlprobleme. Bei den in der vorliegenden

Arbeit ausgewerteten Informationen handelt es sich hauptsächlich um Daten, die in der landwirtschaftlichen Praxis erfasst werden. Lediglich die Anwendungs- und Abgabebelege zum Antibiotikaeinsatz werden von Tierärzt:innen ausgefüllt. Daher scheint es durchaus sinnvoll, im Rahmen einer nationalen Tiergesundheitsdatenbank neben den vorhandenen Agrar-Daten auch solche veterinärmedizinischen Ursprungs mit einzubeziehen, um die fachliche Qualität zu sichern.

### **3.4 Abschließende Bewertung**

Grundsätzlich sind die hier erarbeiteten Ergebnisse aus der vorliegenden Studie und den beiden Publikationen als Basis für ein nationales Monitoring geeignet und es wurden Erkenntnisse gewonnen, die nun im größeren Kontext angewandt werden können. Dabei ist vor allem darauf zu achten, dass der Gesamt-Score nur als Screening-Instrument für einen ersten Status Quo des Tierwohlstatus der Betriebe genutzt werden kann. Um den komplexen Zusammenhängen der Tierwohlindikatoren gerecht zu werden, darf jedoch der Fokus auf die einzelnen Tiergesundheits-Scores nicht verloren gehen. Daher ist für nötigen Handlungsbedarf und weitergehende Beratung die multivariate Betrachtung der einzelnen Tiergesundheits-Scores zwingend erforderlich. Neben der multivariaten Auswertung fehlt allerdings auch eine "multivariate Interpretation". Ein klassisches Benchmarking, wie etwa bei den relativen z-Scores, kann z. B. auch mit Hilfe einer multivariaten Diskriminanzfunktion erfolgen. Hierzu müsste jedoch eine geschichtete, spezifische Analyse für jede Klasse von Produktionsbedingungen erfolgen, so dass hierzu in jedem Falle eine Erweiterung der Datenbasis erforderlich wäre.

Zudem ist aufgrund der heterogenen Produktionsbedingungen die Empfehlung von Einzelmaßnahmen zur Verbesserung des Tierwohls kaum zwingend kausal, weshalb nicht nur eine multivariate, sondern auch eine multifaktorielle Betrachtungsweise notwendig ist. Wegen der fehlenden Kausalität sollte daher besser ein gesamtheitlicher Ansatz verfolgt und Abhängigkeiten der betrieblichen Faktoren untereinander berücksichtigt werden. Ansonsten

besteht das Risiko, den an einer Stelle erwirkten Erfolg durch eine Verschlechterung an anderer Stelle einzubüßen. Zudem sind einige der untersuchten Einflussvariablen nicht nur als potentielle Präventionsmaßnahmen zu sehen, sondern auch als wichtige Stratifizierungsmerkmale. So haben die Einstallgewichtsklasse und die Altersgruppe bei der Untersuchung so maßgeblichen Einfluss auf einige der Tierwohlindikatoren, dass nur eine gesonderte Betrachtung sinnvoll ist.

In Bezug auf die Erfassung zusätzlicher Indikatoren ist die weiterführende Forschung zur Validierung der vorgeschlagenen Indikatoren daher unbedingt empfehlenswert. Um eine kontinuierliche und vor allem nachhaltige Erhebung zu ermöglichen, sind zudem intensive Observer-Schulungen anzuraten. Außerdem sollten die Komplexität und der Umfang des Untersuchungsprotokolls angepasst werden und klare Definitionen der zu erhebenden Gesundheitsmerkmale festgelegt werden, um eine möglichst harmonisierte und qualitativ korrekte Erfassung zu garantieren.

Abschließend ist festzuhalten, dass die Untersuchungen und Auswertungen der vorliegenden Studie gezeigt haben, dass die Nutzung von Produktionsdaten nach intensiver Vorbereitung für ein Tierwohlmonitoring genutzt werden können. Außerdem wurden potentielle zusätzliche Indikatoren identifiziert. Zu deren Aufnahme in Standarduntersuchungsprotokolle sind jedoch weitere Validierungsschritte notwendig, da aufgrund des Studiendesigns von MulTiViS die Aussagekraft des Vergleichs von Schlachtbefunden und den Ergebnissen der speziellen tierärztlichen Erhebung nur eingeschränkt belastbar ist. Ursache hierfür ist der Konflikt zwischen Praxistauglichkeit und statistischer Aussagekraft.

## 4 Zusammenfassung

Julia Große-Kleimann

### **Tierwohlbewertung von Mastschweinen durch übergreifende Analyse vorhandener Produktionsdaten und aktiv erfasster Variablen im Betrieb**

Das Tierwohl von Nutztieren hat in den vergangenen Jahren zunehmend an Aufmerksamkeit gewonnen. Trotz gesetzlicher Vorgaben gibt es bisher jedoch noch keine einheitlich festgelegten Indikatoren zur Messung des Tierwohls. Daher wird zur Bestimmung der Tiergerechtheit von Haltungssystemen für Nutztiere auf verschiedene direkte und indirekte Methoden zurückgegriffen. Neben der aktiven Untersuchung der Tiere spielt auch die Nutzung bereits vorhandener Daten aus der Produktionskette eine wichtige Rolle. Um eine fortlaufend aktuelle Einschätzung des Tierwohlstatus in großem (nationalen) Umfang vornehmen zu können, bietet sich daher die Sekundärdatennutzung an. Das Verbundprojekt MULTiViS befasst sich dazu mit der Fragestellung, ob im Rahmen eines Tierwohlmonitorings von Mastschweinen routinemäßig erfasste und anschließend aufgearbeitete Produktionsdaten die tierärztliche Untersuchung im Bestand ergänzen oder sogar ersetzen können.

Im Rahmen des Projektes wurden ca. 200 Mitgliedsbetriebe des VzF einmalig an einem Stichtag besucht und eine Zufallsauswahl von Schweinen tierärztlich untersucht. Dabei spielte die Wahl der Projektbetriebe eine entscheidende Rolle, um eine gute Repräsentativität zu erreichen und Verzerrungen zu vermeiden. Die erfassten Indikatoren und Einflussfaktoren wurden größtenteils komplementär zu den vorliegenden Produktionsdaten erfasst. Letztere beinhalten die Ergebnisse aus der Betriebszweiganalyse des VzF, Informationen zum Antibiotikaeinsatz sowie die Schlachtbefunde. Für die statistische Auswertung mussten diese Daten zunächst aufbereitet werden. Dazu wurden u.a. Variablen zusammengefasst oder transformiert und eine Adjustierung der Befunddaten aus der routinemäßigen Fleischuntersuchung vorgenommen, um den sogenannten Schlachthof-Effekt auszugleichen.

Zudem wurde das Einstallgewicht der Ferkel zur Mast als essentielles Stratifizierungsmerkmal herausgestellt, da dieses für grundlegende Management-Philosophien steht, die erheblichen Einfluss auf wichtige Tierwohlintikatoren haben.

Als Vorbereitung für ein relatives Benchmarking wurden die Indikatoren durch z-Standardisierung auf eine einheitslose Skala gebracht. Dadurch kann jedem Betrieb ein spezifischer Score-Wert pro Indikator zugeordnet werden kann, welcher die Leistung im Vergleich zu den anderen Betrieben im Untersuchungskollektiv beschreibt. Dies birgt den Vorteil, dass keine absoluten Cut-Off-Werte definiert werden müssen und zudem ein stetiger Ansporn zur Verbesserung für den einzelnen Betrieb gegeben ist. Auf Basis ausgewählter Indikatoren wurden sieben Tiergesundheits-Scores und ein Gesamt-Score gebildet, welcher als Screening-Instrument zur Einschätzung des Tierwohlstatus dienen kann. Als Grundlage für eine landwirtschaftliche oder tierärztliche Beratung ist jedoch die multivariate Berücksichtigung der Einzel-Scores unerlässlich.

Die Untersuchung von möglichen Einflussgrößen auf die Tiergesundheits-Scores ergab, dass sich die Scores, welche sich auf die biologische Leistung der Schweine beziehen, durch wirtschaftliche Größen modellieren lassen. Jedoch spielt die große Produktionsvielfalt der Betriebe eine entscheidende Rolle, da durch die Heterogenität des Untersuchungskollektivs keine universell gültigen Aussagen zu spezifischen Präventionsmaßnahmen getroffen werden können. Aufgrund der Abhängigkeit einzelner Faktoren untereinander sollte zudem ein multifaktorieller Ansatz verfolgt werden. Insgesamt zeigten verschiedene varianzanalytische Modelle, dass die Übereinstimmung von Produktionsdaten und Erhebung eher gering ist, was für die Wichtigkeit der speziellen tierärztlichen Untersuchung im Bestand spricht. So konnten die Tiergesundheits-Scores, welche auf den Schlachtbefunden beruhen, durch die entsprechenden Indikatoren aus der Erhebung nur zu einem geringen Maß dargestellt werden.

Zur Umsetzung der Studienergebnisse in ein nationales Monitoring müssen zunächst die Verfügbarkeit der Daten bzw. die Praxistauglichkeit geklärt werden. Außerdem sollte die Qualität und Belastbarkeit der Befunddaten aus der Fleischuntersuchung verbessert werden.

## 5 Summary

Julia Grosse-Kleimann

### **Assessment of animal welfare in fattening pigs by integrative data analysis of existing production data and results of on-farm examination**

The welfare of farm animals has gained increasing attention in recent years. However, despite legal requirements, there are still no uniformly defined indicators for measuring animal welfare. Therefore, various direct and indirect methods are used to determine the animal welfare of farm animal husbandry systems. In addition to the active examination of the animals, the use of existing data from the production chain also plays an important role. In order to be able to make a continuously updated assessment of the animal welfare status on a large (national) scale, the use of secondary data is therefore an option. The joint project MultiViS is addressing the question of whether production data routinely collected and subsequently processed within the framework of animal welfare monitoring of fattening pigs can supplement or even replace veterinary on-farm examinations.

In the project, about 200 member farms of the VzF were visited once on a specific date and a random selection of pigs was examined by a team of veterinarians. The choice of project farms plays a decisive role in achieving good representativeness and avoiding biases. The indicators and influencing factors recorded were largely complementary to the available production data. The latter include the results from the VzF's farm branch analysis, information on antibiotic consumption and the slaughter findings. For the statistical evaluation, these data first had to be processed. For example, variables were combined or transformed and an adjustment of the slaughter findings was made to compensate for the so-called slaughterhouse effect. In addition, the initial body weight of piglets stalled in for fattening was identified as an essential stratification characteristic, as this represents fundamental

management philosophies that have a considerable influence on important animal welfare indicators.

In preparation for a relative benchmarking, the indicators were brought to a unitless scale through the so-called z-standardisation. This makes it possible to assign a specific score value per indicator to each farm, which describes the performance in comparison to the other farms in the study collective. This has the advantage that no absolute cut-off values have to be defined and, in addition, a constant incentive for improvement is given to the individual farm. On the basis of selected indicators, seven animal health scores and one overall score were calculated, which can be used as a screening tool to assess the animal welfare status. However, multivariate consideration of the individual scores is essential as a basis for agricultural or veterinary advice.

The investigation of possible variables influencing the animal welfare scores showed that the scores relating to the biological performance of the pigs can be modelled by economic variables. However, the large production diversity of the farms plays a decisive role, as no universally valid statements on specific prevention measures can be made because of the heterogeneity of the study collective. Moreover, due to the dependency of individual factors on each other, a multifactorial approach should be pursued. Overall, different variance-analytical models showed that the agreement between production data and results of the survey is rather low, which speaks for the importance of the special veterinary on-farm examination. Thus, the animal health scores, which are based on the slaughter findings, could only be represented to a small extent by the corresponding indicators from the survey.

In order to implement the results of the study in national monitoring, the availability of the data and the practicality and responsibility of the execution must first be clarified. In addition, the quality and reliability of the findings data from the meat inspection should be improved.

## 6 Literaturverzeichnis

Bartussek H. Tiergerechtheitsindex für Mastschweine. Gumpenstein BfaLB. 1995.

BMEL. Nutztierhaltungsstrategie. 2016.

BMELV Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz.

Bekanntmachung des Berechnungsverfahrens zur Ermittlung der Therapiehäufigkeit eines Tierhaltungsbetriebes durch die zuständige Behörde. Bonn 2013.

Brambell F W R. Report of the technical committee to enquire into the welfare of animals kept under intensive livestock husbandry systems. Her Majesty's Stationery Office 1965.

BRS Bundesverband Rind und Schwein e. V. Pflichtenheft der Erzeugerringe. 2020.

Dawkins M S. Using behaviour to assess animal welfare. *Animal welfare* 2004; 13: 3-7.

DESTATIS. Viehbestand - Fachserie 3 Reihe 4.1 - 3. November 2018.

Deutscher Bundestag. 16. Gesetz zur Änderung zur Änderung des Arzneimittelgesetzes. Bonn 2013.

Deutscher Bundestag. Antwort der Bundesregierung auf die Anfrage Bündnis 90/DIE GRÜNEN: Monitoring-Daten zum Tierwohl aus dem Nutztierbereich. 2017.

Dickhaus C P, Meemken D, Blaha T. Epidemiological Analysis and Classification of the Health Status of Pig Herds – The Heard Health Index (HHI). In: 13th International Congress in Animal Hygiene. Tartu, Estonia: Proceedings of 13th International Congress in Animal Hygiene, 2007: 199-204.

Duncan I J H. Science-based assessment of animal welfare: farm animals. *Revue scientifique et technique-Office international des epizooties* 2005; 24: 483.

Duncan I J H, Dawkins M S. The problem of assessing “well-being” and “suffering” in farm animals. In: *Indicators relevant to farm animal welfare*: Springer, 1983, pp. 13-24.

EFSA European Food Safety Authority. Scientific Opinion on the use of animal-based measures to assess welfare in pigs. *EFSA Journal* 2012.

Europäische Kommission. DURCHFÜHRUNGSVERORDNUNG (EU) 2019/627 DER KOMMISSION vom 15. März 2019 zur Festlegung einheitlicher praktischer Modalitäten für die Durchführung der amtlichen Kontrollen in Bezug auf für den menschlichen Verzehr bestimmte Erzeugnisse tierischen Ursprungs gemäß der Verordnung (EU) 2017/625 des Europäischen Parlaments und des Rates und zur Änderung der Verordnung (EG) Nr. 2074/2005 der Kommission in Bezug auf amtliche Kontrollen. *Amtsblatt der Europäischen Union* L 131/51 2019.

Fahrmeir L, Heumann C, Künstler R, Pigeot I, Tutz G. *Statistik: Der Weg zur Datenanalyse* Springer-Verlag, 2016, pp.

FAWC Farm Animal Welfare Council. *Fife Freedoms*. 1979.

Grunert K G, Sonntag W I, Glanz-Chanos V, Forum S. Consumer interest in environmental impact, safety, health and animal welfare aspects of modern pig production: Results of a cross-national choice experiment. *Meat Science* 2018; 137: 123-129.

Hemme M, Käsbohrer A, von Münchhausen C, Hartmann M, Merle R, Kreienbrock L. Unterschiede in der Berechnung des betriebsbezogenen Antibiotika-Einsatzes in Monitoringsystemen in Deutschland – eine Übersicht. *Berliner und Münchener Tierärztliche Wochenschrift* 2017; 130: 93-101.

Hoischen-Tauber S, Blaha T, Werner C, Sundrum A. Zur Reproduzierbarkeit der Befunderfassung am Schlachthof für Merkmale der Tiergesundheit. *Archiv für Lebensmittelhygiene* 2011; 62(3):82-7.

- IGN Internationale Gesellschaft für Nutztierhaltung. Tierschutzindikatoren am Schlachthof. 2016.
- Initiative Tierwohl. Gesellschaft zur Förderung des Tierwohls in der Nutztierhaltung mbH. <https://initiative-tierwohl.de/>. Zuletzt besucht am 12. August 2021.
- Kish L. A procedure for objective respondent selection within the household. *Journal of the American statistical Association* 1949; 44: 380-387.
- Koch M. Neukonzeption der Schlachttier- und Fleischuntersuchungsstatistik. 2016.
- Kreienbrock L, Pigeot I, Ahrens W. *Epidemiologische Methoden*. Berlin Heidelberg, Springer Spektrum, 2012, pp.
- Merle R, Busse M, Rechter G, Meer U. Regionalisation of Germany by data of agricultural structures. *Berliner und Münchener Tierärztliche Wochenschrift* 2012; 8: 52-59.
- Nienhaus F, Meemken D, Schoneberg C, Hartmann M, Kornhoff T, May T, et al. Health scores for farmed animals: Screening pig health with register data from public and private databases. *PLOS ONE* 2020; 15.
- QS Qualität und Sicherheit GmbH. Leitfaden Antibiotikamonitring Schwein. Bonn 2018.
- QS Qualität und Sicherheit GmbH. Zum Hofe. 2019.
- Rohleder R. Bio-Siegel in der deutschen Lebensmittelbranche (Bachelor Thesis). University of Applied Sciences, Mittweida 2017.
- Schrader L. Indikatoren für Tiergerechtigkeit. Bayerische Arbeitsgemeinschaft Tierernährung, Tagungsband 51. Jahrestagung: Tierernährung und Tierwohl 2013.
- Schrader L, Czycholl I, Krieter J, Leeb C, Zapf R, Ziron M. Tierschutzindikatoren: Leitfaden für die Praxis - Schwein. *KTBL-Praktikerleitfaden*. 2016: 37-51.

- Sonoda L T, Fels M, Oczak M, Vranken E, Ismayilova G, Guarino M, et al. Tail biting in pigs – causes and management intervention strategies to reduce the behavioural disorder. A review. *Berliner und Münchener tierärztliche Wochenschrift* 2013; 126: 104-112.
- Tiedemann U. Bericht zur Lage anlässlich des 28. Deutschen Tierärztetags in Dresden. *Deutsches Tierärzteblatt* 2018; 10: 1380 - 1388.
- Tonsor G T, Olynk N J, Wolf C A. Consumer preferences for animal welfare attributes: The case of gestation crates. *Journal of Agricultural and Applied Economics* 2009; 41: 714-730.
- USDA National Agricultural Library, U.S. Department of Agriculture.  
<https://www.nal.usda.gov/awic/animal-welfare-act>. Zuletzt besucht am 12. August 2021.
- Valros A, Välimäki E, Nordgren H, Vugt J, Fàbrega E, Heinonen M. Intact Tails as a Welfare Indicator in Finishing Pigs? Scoring of Tail Lesions and Defining Intact Tails in Undocked Pigs at the Abattoir. *Frontiers in Veterinary Science* 2020; 7.
- Valide Datenerfassung in der Schlachttier- und Fleischuntersuchung durch Schlungsmaßnahmen. Garmisch-Partenkirchen 2012.
- Wegner B, Spiekermeier I, Nienhoff H, Grosse-Kleimann J, Rohn K, Meyer H, et al. Application of the voluntary human approach test on commercial pig fattening farms: a meaningful tool? *Porcine Health Management* 2020; 6.
- Welfare Quality. Assessment Protocol for pigs. 2009.
- Welfare Quality®. System zur Gesamtbeurteilung des Wohlergehens landwirtschaftlicher Nutztiere. 2009.
- Wellock I J, Emmans G C, Kyriazakis I. Describing and predicting potential growth in the pig. *Animal Science* 2004; 78: 379-388.

## 7 Danksagung

An erster Stelle möchte ich mich bei meinem Doktorvater und dem Institutsdirektor Herrn Prof. Dr. rer. nat. Lothar Kreienbrock bedanken. Während aller Hochs und Tiefs in der langen Projektlaufzeit stand er stets verlässlich hinter mir. Zudem nahm er sich jedes Mal Zeit für meine Fragen und konnte mit seiner Expertise immer wieder Lösungsansätze für die kniffligsten Probleme aufzeigen. Außerdem fand er stets gute Auswege aus noch so schwierig erscheinenden Situationen.

Ein weiterer Dank gilt außerdem den MulTiViS-Kolleg:innen, die während der gemeinsamen Projektzeit freundlich und konstruktiv mit mir zusammengearbeitet haben. Es war sehr hilfreich, die Herausforderungen und lehrreichen Erfahrungen in diesem Projekt teilen zu können.

Außerdem möchte ich mich bei meinen Kolleg:innen des Instituts für Biometrie, Epidemiologie und Informationsverarbeitung bedanken. Maria danke ich für die geduldige Einarbeitung in SAS und Corinna für ihre fleißige Unterstützung beim Programmieren. Mit seinem Wissen im Themenbereich Statistik hat mich außerdem Martin unterstützt. Zudem möchte ich mich noch mal ganz herzlich bei Fenja für das Gegenlesen bedanken und bei ihr und Katharina für die freundschaftliche Unterstützung. Sie standen mir jederzeit mit Rat und Tat zur Seite, wenn es mal nicht wie geplant verlief.

Zuletzt geht ein riesiges Dankschön an meine Familie, vor allem an David, der mir die Möglichkeit und die nötigen Freiräume für diese Arbeit geschaffen hat und zudem stets ein offenes Ohr für mich hatte.