Deep learning from videography as a tool for measuring infection in poultry

Neil Scheidwasser[1†], Louise Ladefoged Poulsen[2†], Mark Poulsen Khurana[2], Maider Iglesias-Carrasco[3], Prince Ravi Leow[1], Daniel Joseph Laydon[4, 5], Christl Ann Donnelly[6], Anders Miki Bojesen ✉[2*], Samir Bhatt[*] ✉[1, 4*] & David Alejandro Duchêne ✉[1*]

[1]Section of Epidemiology, Department of Public Health, University of Copenhagen, 1353 Copenhagen, Denmark
[2]Section of Veterinary Clinical Microbiology, Department of Veterinary and Animal Sciences, University of Copenhagen, 1870 Frederiksberg, Denmark
[3]Center for Evolutionary Hologenomics, Globe Institute, University of Copenhagen, 1353 Copenhagen, Denmark
[4]MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, W12 0BZ, United Kingdom
[5]Centre for Health Economics & Policy Innovation, Department of Economics & Public Policy, Imperial College Business School, Imperial College London, London, United Kingdom
[6]Department of Statistics, University of Oxford, Oxford, OX1 3LB, United Kingdom

[†] Equal contribution

[*] Equal contribution

Correspondence: *{miki;david.duchene;samir.bhatt}@sund.ku.dk*

Poultry farming is threatened by regular outbreaks of *Escherichia coli* (*E. coli*) that lead to significant economic losses and public health risks. However, traditional surveillance methods often lack sensitivity and scalability. Early detection of infected poultry using minimally invasive methods is thus essential for preventing epidemics. To that end, we leverage recent advancements in computer vision, employing deep learning-based tracking to detect behavioral changes associated with *E. coli* infection in a case-control trial comprising two groups of broiler chickens: (1) an uninfected, unvaccinated negative control group and (2) *E. coli*-infected chickens without prior vaccination using a field strain from the poultry industry. More specifically, we can use tracking data to extract simple features pertaining to kinematic behavior, such as distance travelled, rate of change in body area, and time spent near the food source. These revealed markedly reduced activity in the challenged group when compared to the negative control. Post-mortem physiological markers of infection and inflammation, including serum amyloid A and organ lesion scores, confirmed infection severity differences among groups. Overall, this study demonstrates the potential of deep learning-based behavioral analysis as a scalable and objective tool for early infection detection in poultry farming, paving the way for improved animal welfare and disease management.

Keywords. deep learning, poultry farming, broiler chickens, animal behavior, immunology

1. Introduction

Poultry farming is a cornerstone of the agriculture industry and is essential to global nutritional security, accounting for nearly 35% of the world's meat production [1]. In many poultry-producing regions, *Escherichia coli* (*E. coli*) infection is a leading cause of morbidity and mortality, including China [2, 3], the United States [4, 5], Brazil [6], and several European countries such as the United Kingdom [7] and Denmark [8, 9]. The diversity of pathogens is also of concern for public health, particularly those of economic importance and frequent zoonosis, which include several viruses, bacteria, fungi, and roundworms [10–12]. Consequently, detecting early signs of infection in domesticated birds, such as broiler chickens, is crucial to animal welfare, economics, and public health.

Over recent decades, a plethora of techniques have been proposed to leverage data from physiological markers (e.g., body temperature [13–15]) or behavioural outputs (e.g., locomotion [16]; sound [17–20]) to assess welfare impairments and detect pathogen outbreaks. Measuring temperature requires precise thermal imaging cameras, which are expensive and sensitive to environmental disturbances. In contrast, acquiring sound and video recordings is logistically simpler and less expensive, but usually requires computer-intensive postprocessing to extract meaningful behavioural features. In addition, obtaining precise estimates of individual patterns remains challenging. For instance, commonly used video processing methods such as optical flow [16, 21] provide insights at the level of a flock and not of the individual. To obtain individual-level measurements, gait-scoring systems have been developed and implemented in broiler chickens [22, 23], but these are impractical for real-time applications. Other frameworks based on markers or wearable sensors (e.g., accelerometers [24] or RFID chips [25–27]) are promising for their high resolution, but are intrusive and not easily scalable to industrial settings.

Alternatively, recent progress in deep learning [28] for computer vision has catalysed the development of versatile frameworks for markerless pose estimation and kinematic covariate extraction for individual [29–32] or multiple animals simultaneously [33–36], which have especially garnered interest in neuroscience [37, 38] and ecology [39]. As to animal behaviour and welfare monitoring, the potential advantages of deep learning-based videography analysis are multi-factorial. Indeed, state-of-the-art solutions are non-invasive, adaptable to various environments, user-friendly, and may allow for individual-level analysis with near-human-level precision. Although these systems initially require data-intensive training of deep neural networks on videos of interest, transfer learning allows for fast deployment of pre-trained models on new video data.

To this end, we develop a feasibility study that demonstrates the capacity of deep learning-based, markerless tracking (i.e., tracking without pre-placed physical markers) from video recordings to detect phenotypical and physiological differences in broiler chickens following *E. coli* infection. We designed a trial in which 40 chickens were randomly assigned to two, equal-sized, separately-housed groups: (1) an uninfected group (control), and (2) a group infected with *E. coli* (Fig. 1). Infection with *E. coli* in many poultry-producing regions is a leading cause of morbidity and mortality [5, 40]. In addition, *E. coli* outbreaks are linked with intensive use of antibiotics, promoting the selection of resistant strains that have the potential to spillover into human populations [41].

Here, we use the multi-animal version of DeepLabCut [34] to identify behavioural changes associated with *E. coli* infection in a trial with two groups of broiler chickens: (1) uninfected group (negative control), and (2) an *E. coli*-infected group using a standard industry dose. Using automated measurements from deep learning, we found chickens infected with *E. coli* to move and travel less in their pens and spend less time near the food source (used as a proxy to measure feeding time). This was in agreement with necropsy data confirming distinct patterns of *E. coli* infection and inflammation in infected compared with uninfected groups. Thus, our data provide evidence that deep learning on video is a useful and objective tool for detecting the behavioural and physiological changes associated with infection. The simplicity and objectivity of the statistics gleaned from videos, including distance travelled and changes in body area, indicate that scalability to industrial settings is feasible. Addressing sensitivity to other pathogens, high animal densities, and trade-offs between labelling individual animals and image resolution will be crucial for implementing reliable deep learning-based outbreak detection methods in broiler farms.

2. Methods

2.1. Animals and housing conditions

40 day-old, *E. coli*-unvaccinated, Ross 308 broiler chickens (mixed gender) were obtained from Blenta AB (Blentarp, Sweden) and followed for 38 days (Fig. 1a). The animals were accommodated in two groups of 20 (see Experimental design) in an 8.64 m^2 coop (3.6 m x 2.4 m; Fig. 1c) with a dust bath to a 16h light (06h-22h):1h dim (22.00-23.00):7h dark (23.00-06.00) schedule. Feed was provided *ad libitum* twice daily to accommodate the increased growth (from 15 to 200 grams/day) according to the feeding schedule of Aviagen (Bække, Denmark) and consisted of commercial wholefood for broiler chickens of age 0-8 weeks (Brogaarden, Lynge, Denmark). Water was available *ad libitum* at two different sources. Standard procedures were conducted to prevent cross-contamination from other pathogens: experimenters wore personal protective equipment and used sanitized lab equipment before conducting daily checks of each pen.

2.2. Experimental design

On the day of arrival, each chicken was randomly assigned to one of two groups: uninfected control group, or infected with *E. coli*. For the infected group, *E. coli* (strain ST117 E44; accession number LXWV00000000.1) was administered intratracheally (9.9×10^5 cfu/chicken in 0.5 ml) (week 5 of the experiment; Fig. 1a), while chickens from the control group were sham-challenged intratracheally (phosphate-buffered saline; (PBS)) following the same timeline. The strain used for inoculation was provided by the Section of Veterinary Clinical Microbiology at the University of Copenhagen and prepared as described in Kromann *et al.* [40] (storage at -80°C, streaking on blood agar base supplemented with 5% bovine blood (BA) and incubation at 37°C overnight). A single colony was incubated in brain heart infusion (BHI) broth overnight at 37°C. Subsequently, 1 ml of the culture were transferred to 100 ml BHI and incubated until the optical density at 600 nm (OD_{600}) reached 1.56 (approximately two hours). At OD_{600} = 1.56 (exponential phase), 308 µL of culture was transferred to 199.69 ml PBS (Thermo Fisher Scientific; Roskilde, Denmark) and a volume of 0.5 ml per animal was used as inoculum and kept on ice until infection. The colony-forming unit (CFU) count in the inoculum was confirmed by 10-fold dilutions prepared in triplicates and plated (100 µL) on BA, incubated overnight before counting.

2.3. Biological data sampling

Blood samples were collected twice: firstly a day before the challenge and secondly before euthanisation for analysis of serum amyloid A using the LZ-SAA assay (Eiken Chemical Co., Tokyo, Japan). Euthanisation was performed by cervical dislocation directly following the induction of unconsciousness by blunt head trauma. Gross pathology was performed two days post-infection during the necropsy in a randomised and blinded manner to mitigate experimenter biases. An overall lesion score was calculated by visually assessing and grading pathological lesions in the peritoneum, airsacs, lungs and spleen (see Table S1). To measure pulmonary *E. coli* CFU counts, each lung was first weighed and then blended with a 1/1 (weight/vol) 0.9% PBS. The resulting blend was serially diluted tenfold (10^{-1}-10^{-7}) and 3×30 µL per dilution was spotted separately for overnight incubation. CFU counts were normalised by lung weight.

2.4. Behavioural data sampling

For each group, five videos were recorded over three days: two in the first two days (morning and afternoon, after feeding) and one on the third morning (Fig. 1a). Each recording lasted between 20 and 30 minutes and was acquired at 30 frames per second (FPS) at a resolution of 1080×1920 pixels using GoPro HERO10 Black cameras (https://gopro.com). The cameras were placed 130 cm away from the floor. For each video, the first two minutes were removed to minimise any potential experimenter intervention.

2.5. DeepLabCut training and inference

Animal tracking was performed in markerless fashion using the multi-animal version of DeepLabCut [34]. For each video, twenty frames were selected using k-means clustering. For each frame, annotations for five body parts (head, centre left, centre right, saddle, and tail) were added for each animal. We used a pre-trained *DLCRNet_ms5* [34], which consists of

a multi-scale convolutional neural network (CNN) built on top of a ResNet-50 [42] feature extractor and stacked with deconvolutional layers to predict where the body parts of interest are and which animals they belong to. The multi-scale architecture fuses high and low resolution maps, thus allowing to accurately detect body parts with both fine-grained detail and broader contextual information, enabling robust performance across varying poses, scales, and inter-animal occlusions. The model was trained with default parameters (Adam [43] optimisation with multi-stage learning rate (η) scheduling (stage 1: $\eta = 1 \times 10^{-4}$ for 7,500 iterations; stage 2: $\eta = 5 \times 10^{-5}$ for 12,000 iterations; stage 3: $\eta = 1 \times 10^{-5}$ until the end of the experiment), batch size of 8) using 95% of the labelled frames for 200,000 training iterations. The root mean squared errors (RMSEs) were: 3.03 pixels (train set) and 8.68 pixels (test set) for an image size of 1080×1920 pixels. The trained network was then used to extract the (x, y) positions of each body part for each animal in all videos. An example frame is shown in Figure 1e.

2.6. Post-processing

For each video, DeepLabCut outputs an array of (x-coordinate, y-coordinate, and likelihood) triplets of size $F \times (N \times B \times 3)$, where F is the number of frames, N the number of animals, and B the number of body parts. Predictions with a likelihood below 0.9 were dropped. To recover some of the dropped values, forward linear interpolation was applied at a body part level, where at most 15 consecutive missing values could be filled (0.5 s). (x, y) positions were finally smoothed using Savitzky-Golay filtering with a polynomial order of 7 and a window length of 15 frames. Subsequently, three behavioural features could be extracted (Fig. 1b). First, distance travelled (in px) was estimated on a frame-to-frame basis by taking the average Euclidean distance of each body part position from one frame to the next. Second, to quantify other forms of behavioural activity not necessarily related to locomotion (e.g., stretching for improved ventilation), the rate of change in body area was estimated in a similar vein to [38], using the shoelace formula based on a polygon defined by the head, the left centre, the saddle, and the right centre (Fig. 1d). Similar to the travel distance, absolute frame-wise differences in body area were collected. To account for potential radial and tangential distortions due to the camera setup, for each frame, we normalised the measured velocity and rate of change in body area of each animal by its head-to-saddle distance. Third, the time spent near the food source (assumed as a proxy for time spent eating) was estimated for each animal by computing the number of frames in which the animal's head was in the vicinity of the food source (Fig. 1c). The vicinity was defined as a rectangle whose scale of 1.05 times the scale of the food source in the video.

2.7. Statistical analysis

For all analyses of physiological covariates, nonparametric Mann-Whitney U tests were conducted to examine between-subject differences. The t-test assumption of normality was assessed using quantile-quantile plots (Q-Q plots) while Levene's test was used to test for homoscedasticity. For all markers, both normality and homoscedasticity could not be reasonably assumed for lesion scores and SAA levels, leading us to opt for Mann-Whitney U tests to assess between-group disparities. For all pairwise comparisons, Cohen's d was reported to reflect the effect size.

For each behavioural feature, a Bayesian mixed-effects model was employed to investigate between-group variations (with the control group as a reference), while accounting for random variations across the different recordings. Our mixed effects model had the following linear predictor

$$\mu_{ij} = \alpha + \beta X + \gamma_{0i}$$

where μ_{ij} is a behaviour feature of interest (e.g., distance travelled) for a day i and individual j, α is an intercept that corresponds to the control group, β is a scalar for the group X (infected or not), and γ is a random effect for the ith day. As the distance travelled and body area change were normalised by the body length (thus constituting rates distributed in (0, 1)), we assumed their values were generated from a Beta distribution. Whilst the time spent near food source is also a rate, a zero-inflated version of the Beta distribution was used to account for the frequent zero-valued observations. All models were fitted with the default prior distributions suggested by `brms` [44]. To facilitate efficient sampling and minimise divergent transitions, a target acceptance rate of 0.99 was chosen and a maximum tree depth of 15 was selected for building during the trajectory phase of the Hamiltonian Monte Carlo no-U-turn sample (NUTS) [45]. For all variables, we compared the beta family models (with a logit link function) against models with the default family parameter (Gaussian with an identity link function) in `brms` and picked the best family based on a Bayesian estimate of the expected log pointwise predictive density (ELPD) from leave-one-out cross-validation (LOOCV) (Table S4). Finally, convergence

plots and posterior predictive checks were plotted to check model fit (see Figures S1 to S4). For all variables, the models fitted using a beta distribution were preferable.

2.8. Implementation

Video pre-processing was performed using `ffmpeg` (https://ffmpeg.org) and tracking was implemented using the multi-animal version of DeepLabCut 2.3.5 [29, 34]. Fine-tuning of the pre-trained DLCRNet was performed on a single NVIDIA RTX A6000 graphical processing unit (GPU). Feature extraction and statistical tests were performed using custom scripts in Python using `NumPy` (Harris *et al.* [46], version 1.25.0), `pandas` (McKinney [47], version 1.5.3), `SciPy` (Virtanen *et al.* [48], version 1.11.1) and `Pingouin` (Vallat [49]; version 0.5.4). Bayesian mixed-effects modelling was implemented in R 4.4.0 using the `brms` package [44]. The code for reproducing feature extraction and statistical analyses is available at: https://github.com/Neclow/dlc2ecoli.

2.9. Ethics

The study was approved by the Danish Animal Experiments Inspectorate under the Danish Ministry of Environment and Food, and all animal procedures were performed in accordance with this approval (license no. 2019-15-0201-01611) and with ARRIVE guidelines. The licence granted, and guidelines hereof, is in agreement with the Danish law on animal experiments and the EU directive 2010/63. Predefined humane endpoints were determined and animals were observed every 30 min for the initial six hours after inoculation, subsequently at maximally eight-hour intervals for three days after inoculation with increased frequency in the event of clinical signs in any of the groups. If clinical signs were present, e.g., ruffled feathers, depression, anorexia, lethargy or dyspnoea, the bird was either treated with 0.1 mg/kg buprenorphine and observed with increased frequency or euthanised.

3. Results

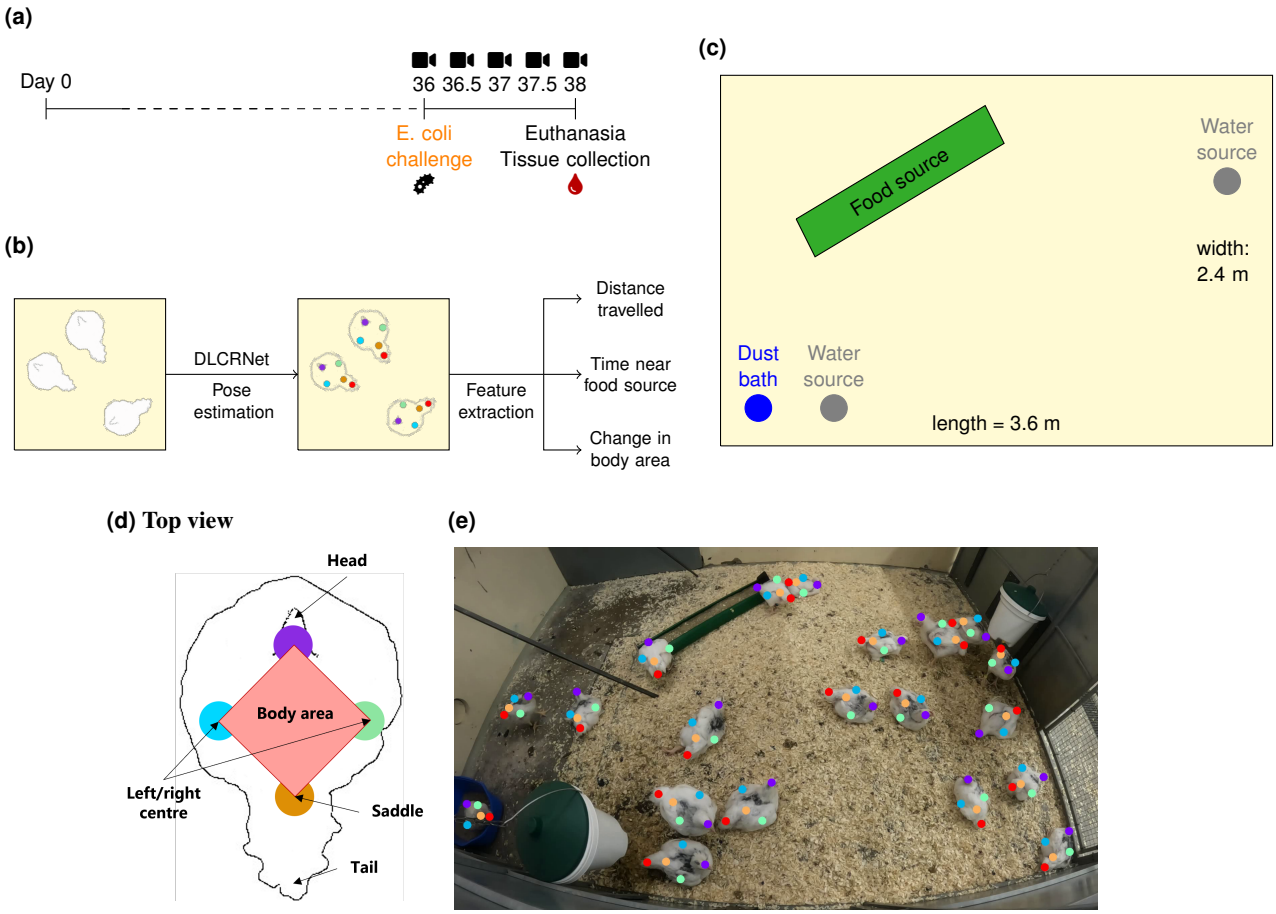

Figure 1. Experimental setup. **(a)** Timeline of the experiment. **(b)** Pipeline for behavioural feature extraction using DLCRNet, the standard model of the multi-animal version of DeepLabCut (maDLC)[34]. **(c)** Diagram of the chicken coops. **(d)** Standardised body parts used for tracking and polygon used to approximate body area. **(e)** Example predictions from maDLC, based on one frame.

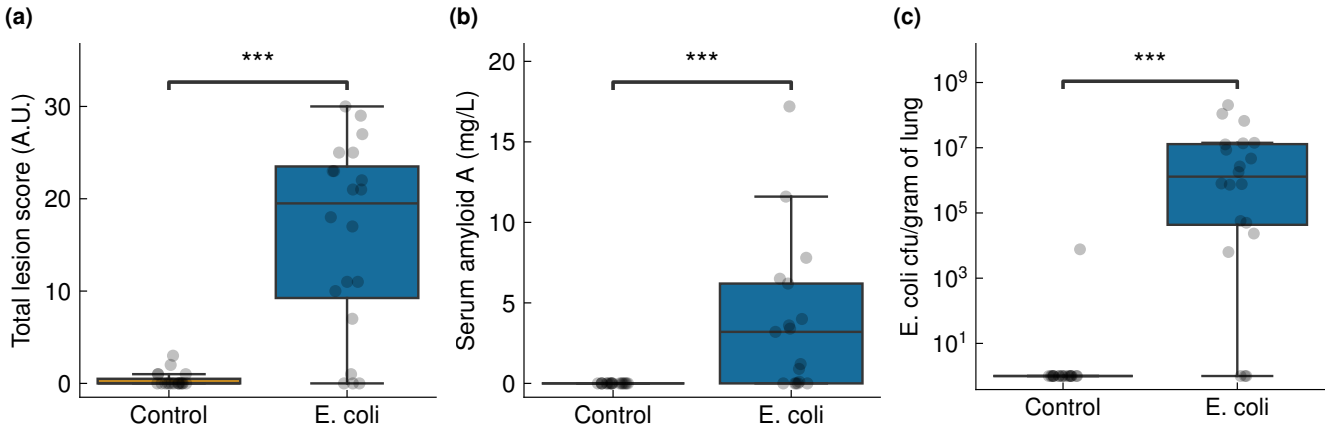

Figure 2. Comparison of physiological biomarkers of inflammation in response to *E. coli* challenge. Significance levels: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

Post-mortem examination of the chickens by necropsy revealed significant disparities between the groups in terms of lesion score, SAA levels, and number of CFUs (Fig. 2; Mann-Whitney U: U_{lesion} = 39.5, U_{SAA} = 45.0, U_{cfu} = 31.0; p-values < 0.001) with medium to very large effect sizes (Cohen's d: d_{lesion} = -2.121, d_{SAA} = -1.153, d_{cfu} = -0.607). Chickens from the control group had near-zero presence of SAA and colony-forming units as well as very few lesions, while *E. coli*-infected groups showed significantly higher values for all markers (Table S2).

Between-group behavioural differences in distance travelled, body area, and time near food source were evaluated using

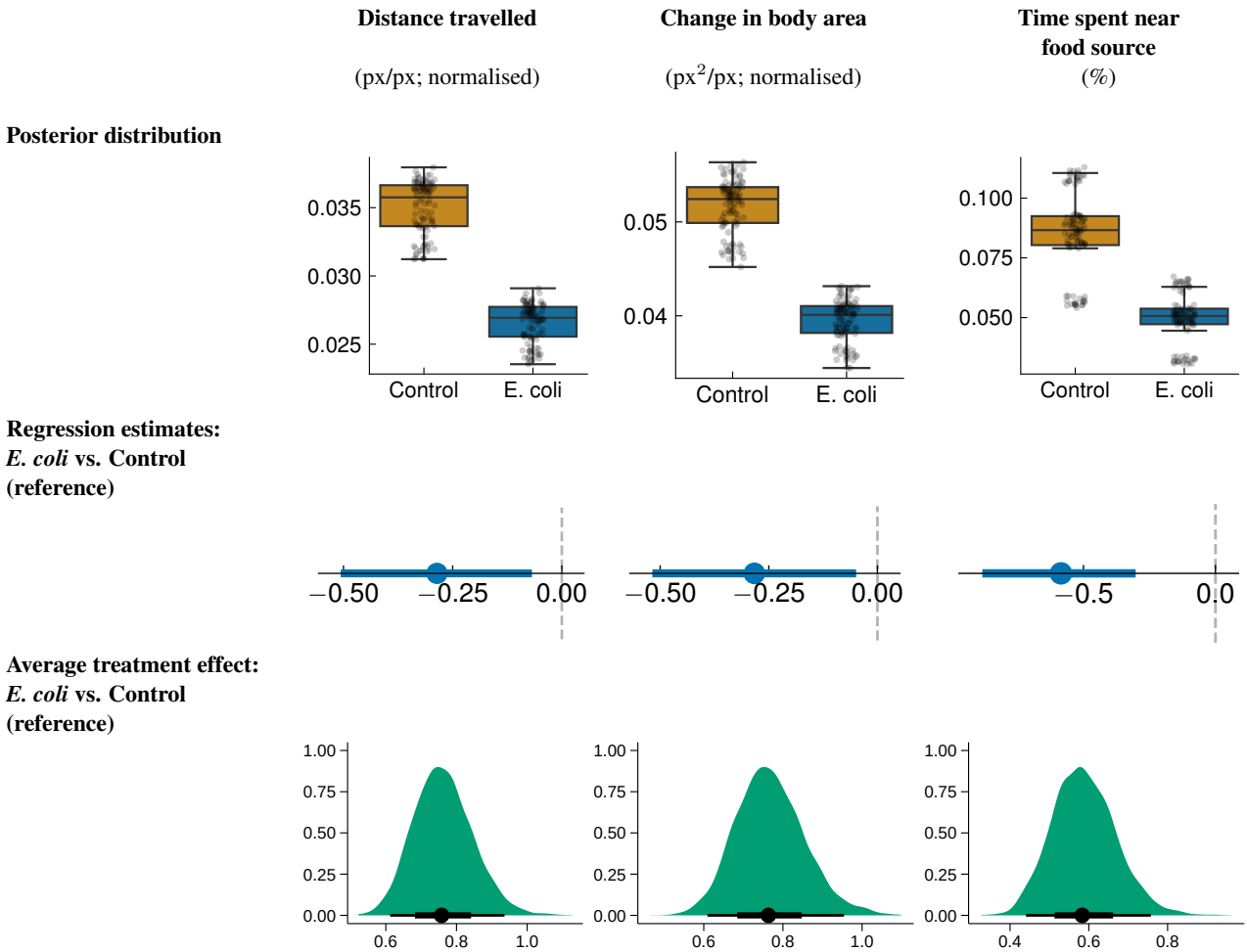

Figure 3. Comparison of behavioural features in response to *E. coli* challenge. Top: posterior distributions from Bayesian mixed-effects modelling to investigate between-group variations while accounting for random variations across the different recordings, predicting the group (infected or not). Middle: mean estimates of the same behaviour features for the *E. coli* group, with the negative control group serving as the reference. Bottom: average treatment effect for each behaviour feature, computed as a ratio between the E. coli and control groups. Error bars indicate 95% confidence intervals.

Bayesian mixed-effects modelling, with random effects applied to the recording time (Fig. S1- S4). We found the infected group to be significantly less active than the uninfected control group as measured by the distance travelled (average treatment effect (ATE) ratio: 24% reduction (95% CI: 6-39%)) and rate of change in body area (ATE ratio: 23% reduction (95% CI: 4-39%)) (Fig. 3; Table S3). In addition, we posit that the infected group fed less as its individuals spent less time near the food source on average (ATE ratio: 41% reduction (95% CI: 24-56%)). Individual chicken identity was not recovered from one recording to another, such that a chicken labelled “chicken1" in a given recording could not be re-identified in subsequent recordings, preventing analyses within-subjects. Overall, video analysis was effective at discriminating behavioural patterns between the control and infected groups, which were confirmed by traditional post-mortem analysis of infection biomarkers.

4. Discussion

In this proof-of-concept study, we demonstrate the capacity of deep learning-based, markerless tracking from video recordings to detect phenotypical differences in broiler chickens following *E. coli* infection. This method holds significant promise for promptly identifying infection or compromised animal well-being at the individual level, without human intervention or intrusive sensors [24–27] - a capability previously limited to flock-level analyses by conventional techniques [16, 21]. Moreover, the simplicity and objectivity of the statistical analyses derived from video data suggest promising avenues for future research in disease detection and management within the poultry industry. Extracting three intuitive features from multi-animal DeepLabCut tracking [34] revealed distinct behavioural patterns between the infected and uninfected chickens, corresponding to observed differences in common physiological markers of *E. coli* infection, including serum amyloid A levels, *E. coli* concentrations and organ lesion scores.

Identifying infections using videography has several challenges, some of which are associated with the inherent conditions from the poultry industry settings, where individual animal identification is laborious. Accordingly, we did not preserve bird identity across recordings, and therefore did not perform within-subject analyses. Although mixed-effects modelling was used to compare trends between groups, analysing individual differences in behavioural and physiological data would facilitate high-resolution analyses. Future research might explore long-term video monitoring at lower resolutions, or the use of individual features for identifications at maximal video resolution. Similarly, we focused on coarse behaviours that might not reflect the behavioural nuances from the immune state of individuals. Upcoming work could explore a greater range of behavioural responses.

Scaling up deep learning-based methods of pose estimation to farm-level settings can be done via a series of technological and practical solutions. Critically, scalability benefits from minimising the number of cameras used, thus maintaining individual chicken identity and data efficiency. Current tracking of individual body parts becomes challenging when the number of individuals is large (e.g., > 50), but solutions are being quickly developed to mitigate the errors arising from crowded environments [50, 51]. Other tracking frameworks such as TRex [33] allow for the tracking of up to 256 individuals with reasonable precision, but do not track individual body parts and are still insufficient for farm-level conditions where several hundreds of animals can coexist. To that end, we posit that using robust deep neural networks trained on semantic segmentation (e.g., SAM [52], SAM-2 [53], or DEVA [54]) or in a self-supervised fashion to build holistic visual representations (e.g., DINOv2 [55]) could be used to compare embeddings frame-by-frame. This framework would operate in a similar manner to optical flow, but we hypothesise that embedding comparison could capture differences more accurately.

5. Conclusion

Our findings demonstrate the potential of deep learning-based video analysis in detecting subtle behavioral and physiological changes linked to infection, offering an objective and scalable tool for outbreak detection in industrial poultry farming. The simplicity and objectivity of the statistical analyses derived from video data, coupled with advancements in deep learning technology, suggest promising avenues for future research in disease detection and management within the poultry industry.

References

1. OECD & FAO. *OECD-FAO Agricultural Outlook 2023-2032* https://www.oecd-ilibrary.org/content/publication/08801ab7-en (OECD, 2023).
2. Liu, C. *et al.* Longitudinal monitoring of multidrug resistance in Escherichia coli on broiler chicken fattening farms in Shandong, China. *Poult. Sci.* **100,** 100887 (2021).
3. Zhou, W. *et al.* Antimicrobial resistance and genomic characterization of Escherichia coli from pigs and chickens in Zhejiang, China. *Front. Microbiol.* **13,** 1018682 (2022).
4. Johnson, T. J. *et al.* Identification of minimal predictors of avian pathogenic Escherichia coli virulence for use as a rapid diagnostic tool. *J. Clin. Microbiol.* **46,** 3987–3996 (2008).
5. Fancher, C. A. *et al.* Avian pathogenic Escherichia coli and Clostridium perfringens: Challenges in no antibiotics ever broiler production and potential solutions. *Microorganisms* **8,** 1533 (2020).
6. Braga, J. F. V. *et al.* Diversity of Escherichia coli strains involved in vertebral osteomyelitis and arthritis in broilers in Brazil. *BMC Vet. Res.* **12,** 1–12 (2016).
7. Kemmett, K. *et al.* The contribution of systemic Escherichia coli infection to the early mortalities of commercial broiler chickens. *Avian Pathol.* **43,** 37–42 (2014).
8. Poulsen, L. L. *et al.* Longitudinal study of transmission of Escherichia coli from broiler breeders to broilers. *Vet. Microbiol.* **207,** 13–18 (2017).
9. Kromann, S. *et al.* Dramatic increase in slaughter condemnations due to Escherichia coli ST23 and ST101 within the Danish broiler production. *Vet. Microbiol.* **280,** 109696 (2023).
10. Miskiewicz, A., Kowalczyk, P., Oraibi, S. M., Cybulska, K. & Misiewicz, A. Bird feathers as potential sources of pathogenic microorganisms: a new look at old diseases. *Antonie van Leeuwenhoek* **111,** 1493–1507 (2018).
11. Mor-Mur, M. & Yuste, J. Emerging bacterial pathogens in meat and poultry: an overview. *Food and Bioprocess Technology* **3,** 24–35 (2010).
12. Agunos, A., Pierson, F. W., Lungu, B., Dunn, P. A. & Tablante, N. Review of nonfoodborne zoonotic and potentially zoonotic poultry diseases. *Avian diseases* **60,** 553–575 (2016).
13. Cook, N., Smykot, A., Holm, D., Fasenko, G. & Church, J. Assessing feather cover of laying hens by infrared thermography. *J. Appl. Poult. Res.* **15,** 274–279 (2006).
14. Okada, H., Itoh, T., Suzuki, K. & Tsukamoto, K. *Wireless sensor system for detection of avian influenza outbreak farms at an early stage* in *IEEE SENSORS* (2009), 1374–1377.
15. Noh, J.-Y. *et al.* Thermal image scanning for the early detection of fever induced by highly pathogenic avian influenza virus infection in chickens and ducks and its application in farms. *Front. Vet. Sci.* **8,** 616755 (2021).
16. Colles, F. M. *et al.* Monitoring chicken flock behaviour provides early warning of infection by human pathogen Campylobacter. *Proc. Biol. Sci.* **283,** 20152323 (2016).
17. Banakar, A., Sadeghi, M. & Shushtari, A. An intelligent device for diagnosing avian diseases: Newcastle, infectious bronchitis, avian influenza. *Comput. Electron. Agric.* **127,** 744–753 (2016).
18. Du, X., Lao, F. & Teng, G. A sound source localisation analytical method for monitoring the abnormal night vocalisations of poultry. *Sensors* **18,** 2906 (2018).
19. Rowe, E., Dawkins, M. S. & Gebhardt-Henrich, S. G. A systematic review of precision livestock farming in the poultry sector: Is technology focussed on improving bird welfare? *Animals* **9,** 614 (2019).
20. Li, N., Ren, Z., Li, D. & Zeng, L. Automated techniques for monitoring the behaviour and welfare of broilers and laying hens: towards the goal of precision livestock farming. *Animal* **14,** 617–625 (2020).
21. Stamp Dawkins, M., Donnelly, C. A. & Jones, T. A. Chicken welfare is influenced more by housing conditions than by stocking density. *Nature* **427,** 342–344 (2004).
22. Kestin, S., Knowles, T., Tinch, A. & Gregory, N. Prevalence of leg weakness in broiler chickens and its relationship with genotype. *Vet Rec.* **131,** 190–194 (1992).
23. Webster, A., Fairchild, B., Cummings, T. & Stayer, P. Validation of a three-point gait-scoring system for field assessment of walking ability of commercial broilers. *J. Appl. Poult. Res.* **17,** 529–539 (2008).
24. Buijs, S. *et al.* Behavioural and physiological responses of laying hens to automated monitoring equipment. *Appl. Anim. Behav. Sci.* **199,** 17–23 (2018).
25. Zhang, F. *et al.* Monitoring behavior of poultry based on RFID radio frequency network. *International Journal of Agricultural and Biological Engineering* **9,** 139–147 (2016).
26. Oliveira, J., Xin, H. & Wu, H. Impact of feeder space on laying hen feeding behavior and production performance in enriched colony housing. *Animal* **13,** 374–383 (2019).
27. Astill, J., Dara, R. A., Fraser, E. D., Roberts, B. & Sharif, S. Smart poultry management: Smart sensors, big data, and the internet of things. *Comput. Electron. Agric.* **170,** 105291 (2020).

28. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521,** 436–444 (2015).
29. Mathis, A. *et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21,** 1281–1289 (2018).
30. Graving, J. M. *et al.* DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. *eLife* **8,** e47994 (2019).
31. Günel, S. *et al.* DeepFly3D, a deep learning-based approach for 3D limb and appendage tracking in tethered, adult Drosophila. *eLife* **8,** e48571 (2019).
32. Pereira, T. D. *et al.* Fast animal pose estimation using deep neural networks. *Nat. Methods* **16,** 117–125 (2019).
33. Walter, T. & Couzin, I. D. TRex, a fast multi-animal tracking system with markerless identification, and 2D estimation of posture and visual fields. *eLife* **10,** e64000 (2021).
34. Lauer, J. *et al.* Multi-animal pose estimation, identification and tracking with DeepLabCut. *Nat. Methods* **19,** 496–504 (2022).
35. Pereira, T. D. *et al.* SLEAP: A deep learning system for multi-animal pose tracking. *Nat. Methods* **19,** 486–495 (2022).
36. An, L. *et al.* Three-dimensional surface motion capture of multiple freely moving pigs using MAMMAL. *Nat. Commun.* **14,** 7727 (2023).
37. Mathis, M. W. & Mathis, A. Deep learning tools for the measurement of animal behavior in neuroscience. *Curr. Opin Neurobiol.* **60,** 1–11 (2020).
38. Sturman, O. *et al.* Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacology* **45,** 1942–1952 (2020).
39. Tuia, D. *et al.* Perspectives in machine learning for wildlife conservation. *Nat. Commun.* **13,** 792 (2022).
40. Kromann, S., Olsen, R. H., Bojesen, A. M., Jensen, H. E. & Thøfner, I. Development of an aerogenous Escherichia coli infection model in adult broiler breeders. *Sci. Rep.* **11,** 19556 (2021).
41. Roth, N. *et al.* The application of antibiotics in broiler production and the resulting antibiotic resistance in Escherichia coli: A global overview. *Poult. Sci.* **98,** 1791–1804 (2019).
42. He, K., Zhang, X., Ren, S. & Sun, J. *Deep residual learning for image recognition* in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016), 770–778.
43. Kingma, D. & Ba, J. *Adam: A Method for Stochastic Optimization* in *ICLR* (San Diega, CA, USA, 2015).
44. Bürkner, P.-C. brms: An R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80,** 1–28 (2017).
45. Hoffman, M. D. & Gelman, A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **15,** 1593–1623 (2014).
46. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585,** 357–362 (Sept. 2020).
47. McKinney, W. *Data Structures for Statistical Computing in Python* in *Proceedings of the 9th Python in Science Conference* (eds van der Walt, S. & Millman, J.) (2010), 56–61.
48. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17,** 261–272 (2020).
49. Vallat, R. Pingouin: statistics in Python. *J. Open Source Softw.* **3,** 1026 (2018).
50. Li, J. *et al. Crowdpose: Efficient crowded scenes pose estimation and a new benchmark.* in *Proc. CVPR* (2019), 10863–10872.
51. Zhou, M., Stoffl, L., Mathis, M. W. & Mathis, A. *Rethinking pose estimation in crowds: overcoming the detection information bottleneck and ambiguity.* in *Proc. ICCV* (2023), 14689–14699.
52. Kirillov, A. *et al. Segment anything.* in *Proc. ICCV* (2023), 4015–4026.
53. Ravi, N. *et al.* Sam 2: Segment anything in images and videos. *arXiv preprint arXiv:2408.00714* (2024).
54. Cheng, H. K., Oh, S. W., Price, B., Schwing, A. & Lee, J.-Y. *Tracking anything with decoupled video segmentation.* in *Proc. ICCV* (2023), 1316–1326.
55. Oquab, M. *et al.* DINOv2: Learning Robust Visual Features without Supervision. *Transact. on Mach. Learn. Res.* (2024).
56. Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. *Statist. Sci.* **7,** 457–472 (1992).

6. Author contributions

N.S., L.L.P., A.M.B., S.B., D.A.D.: study conception and design. L.L.P., N.S.: practical execution. L.L.P., A.M.B.: necropsies and microbiology. N.S.: data analysis. M.P.K., M.I-C., P.R.L.: interpretation of results. L.L.P., A.M.B., D.A.D., S.B.: supervision. N.S., M.P.K., S.B., D.A.D: writing of the original draft. N.S., L.L.P., M.P.K, M.I-C., P.R.L., D.J.L., C.A.D., A.M.B., S.B., D.A.D.: edition of the original draft.

7. Competing interests

The authors declare no competing interests.

8. Funding

SB and DJL acknowledge support from the MRC Centre for Global Infectious Disease Analysis (MR/R015600/1), jointly funded by the UK Medical Research Council (MRC) and the UK Foreign, Commonwealth & Development Office (FCDO), under the MRC/FCDO Concordat agreement, and also part of the EDCTP2 programme supported by the European Union. SB is funded by the National Institute for Health Research (NIHR) Health Protection Research Unit in Modelling and Health Economics, a partnership between the UK Health Security Agency, Imperial College London and LSHTM (grant code NIHR200908). Disclaimer: "The views expressed are those of the author(s) and not necessarily those of the NIHR, UK Health Security Agency or the Department of Health and Social Care." SB acknowledges support from the Novo Nordisk Foundation via The Novo Nordisk Young Investigator Award (NNF20OC0059309). SB acknowledges support from the Danish National Research Foundation via a chair grant (DNRF160) which also supports NS and MPK. SB acknowledges support from The Eric and Wendy Schmidt Fund For Strategic Innovation via the Schmidt Polymath Award (G-22-63345) which also supports PRL. SB and NS acknowledge the Pioneer Centre for AI as affiliate researchers. DAD is funded by a Novo Nordisk Fonden Data Science Investigator Award (NNF23OC0084647). DJL acknowledges funding from the Wellcome Trust UK for the Vaccine Impact Modelling Consortium (VIMC) Climate Change Research Programme (grant ID: 226727_Z_22_Z). AMB acknowledges support from the Innovation Fund Denmark (Grant no. 071-00001) and the Danish Poultry Levy Foundation (Grant no. 122656).

8.1. Acknowledgements

The authors thank Alexandros Katsiferis for valuable insight on analysing the `brms` outputs.

8.2. Data and code availability

All code and data relevant to reproduce the experiments is available online: https://github.com/Neclow/dlc2ecoli. Release of video and physiological data is pending approval from the Department of Veterinary and Animal Sciences at the University of Copenhagen.

A. Appendix

A.1. Supplementary Figures

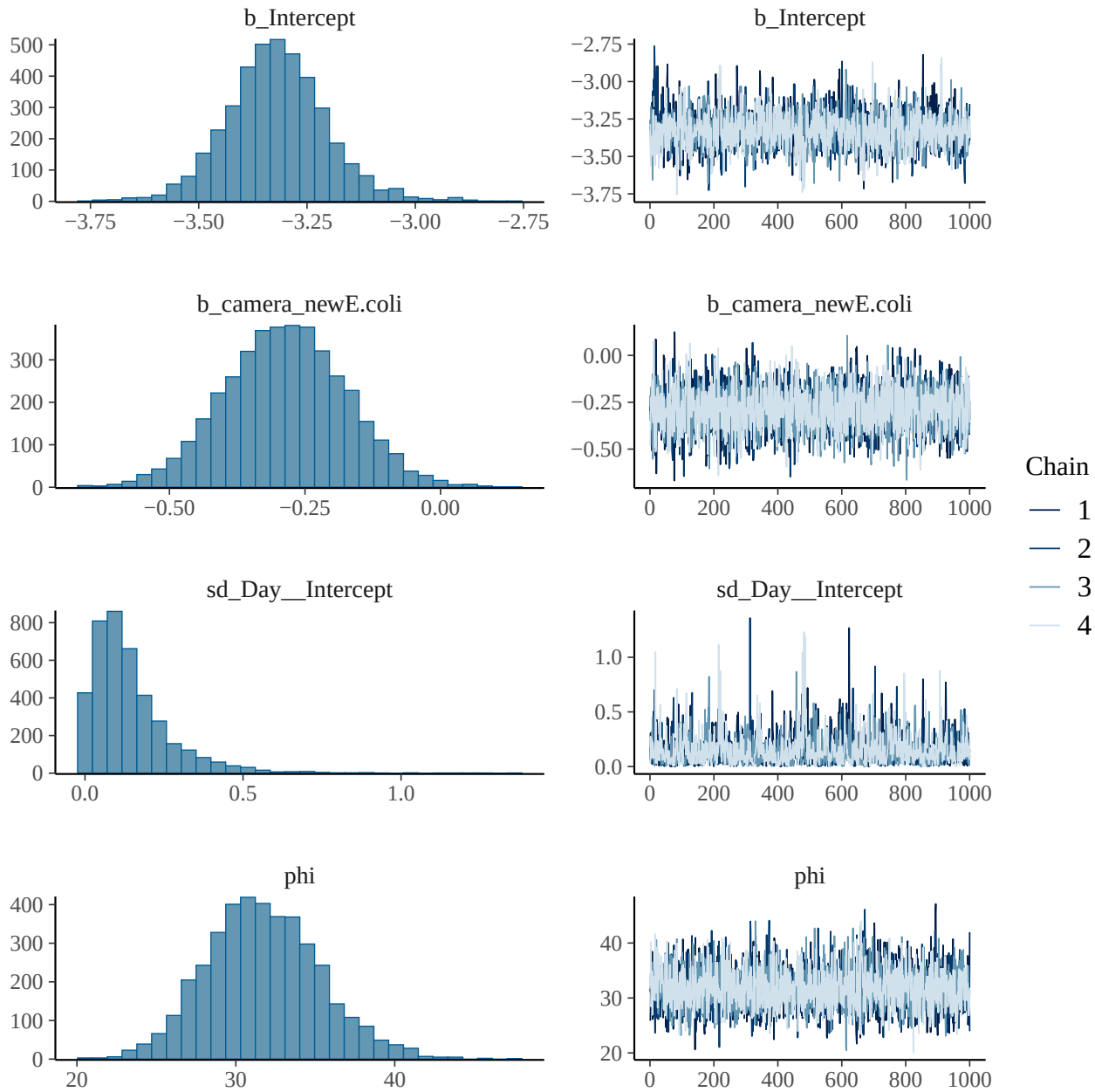

Figure S1. `brms` output for distance travelled, showing density plots for each factor (left) and caterpilar plots (right).

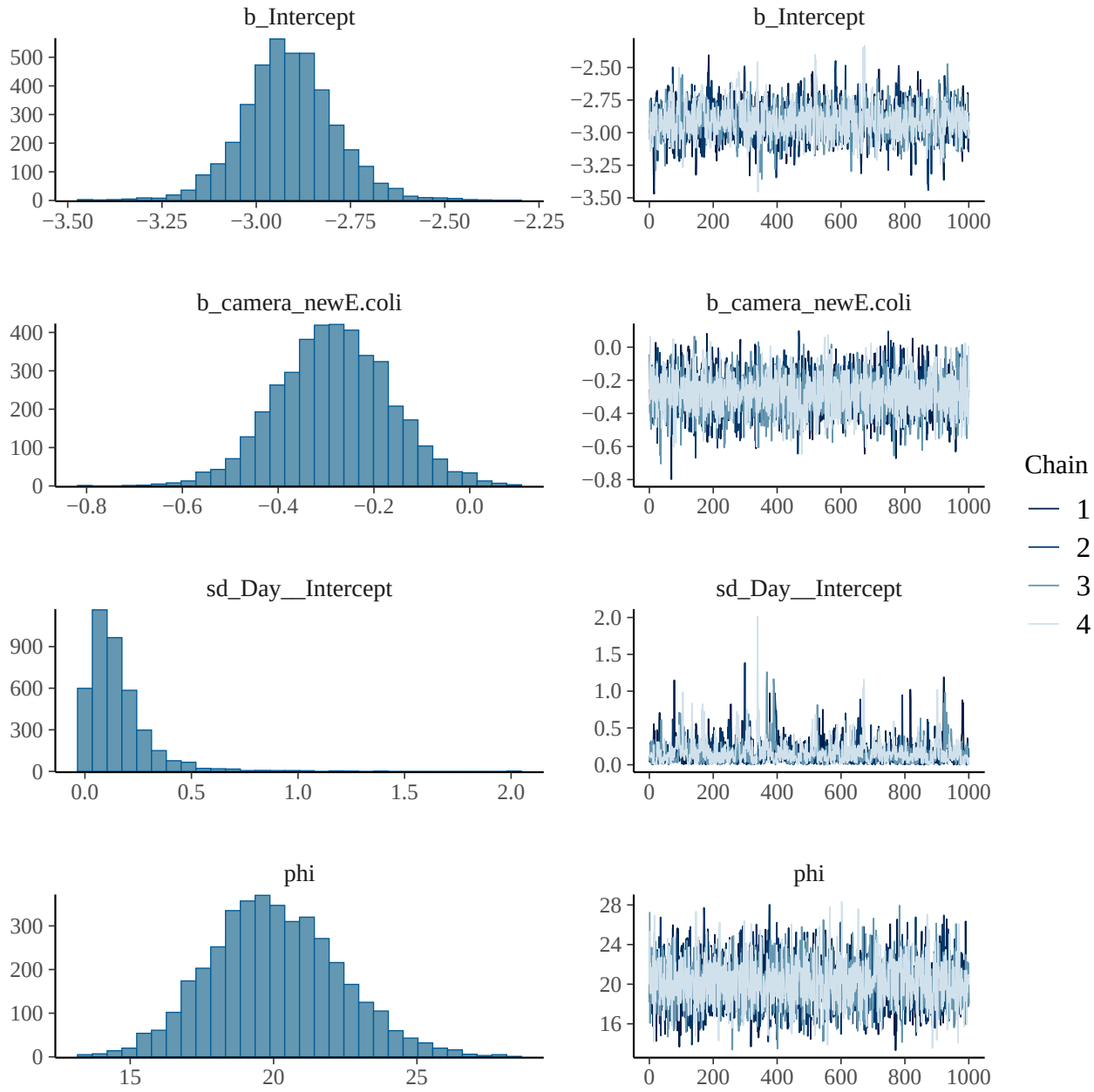

Figure S2. `brms` output for the change in body area, showing density plots for each factor (left) and caterpilar plots (right).

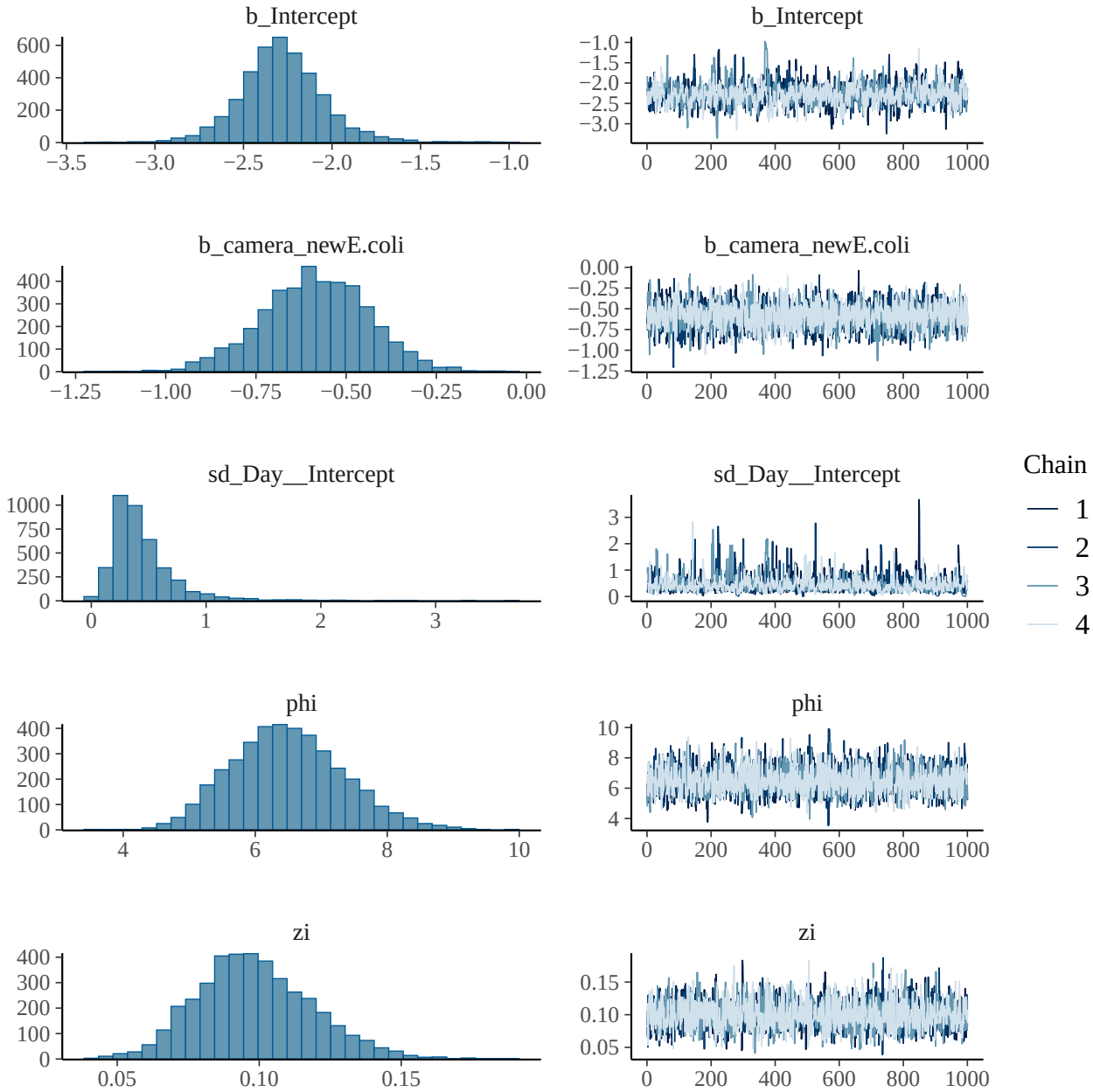

Figure S3. `brms` output for the time spent near the food source., showing density plots for each factor (left) and caterpilar plots (right).

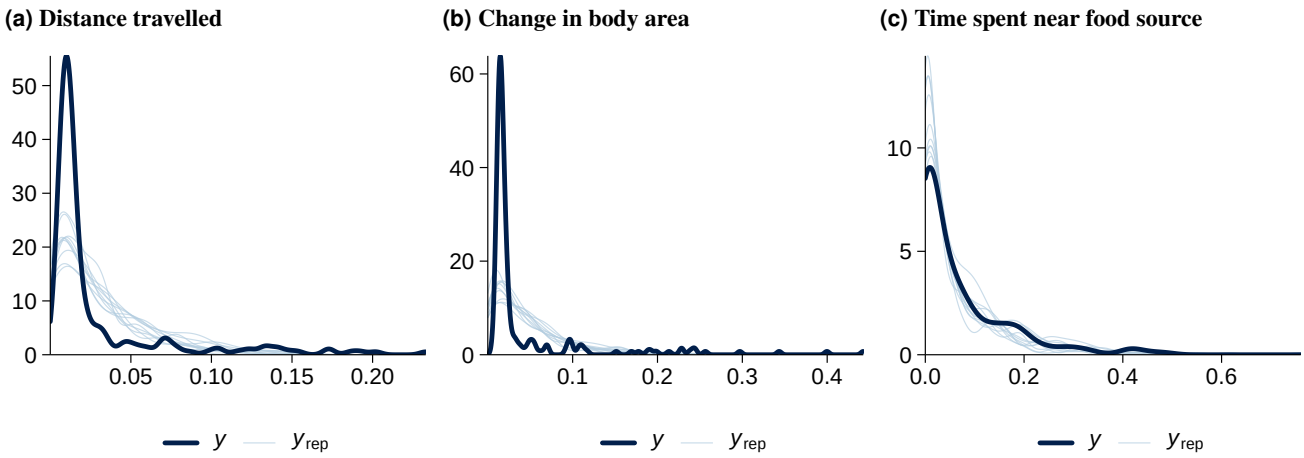

Figure S4. Posterior predictive checks for the time spent near the food source.

A.2. Supplementary Tables

Table S1. Description of lesion score levels

Organ	Characteristic	Levels
Peritoneum	Inflammatory reaction	0: None 1: Local at the entrance 2: Around the ovary 3: In the omentum 4: All peritoneum
	Amount of exudate	0: None 1: Sparse 2: Some 3: Abundant
	Type of exudate	0: None 1: Aqueous 2: Lump of pus 3: Flakes 4: Confluent
	Transparency of peritoneum	0: Clear 1: Nuclear 2: Cloudy 3: Milky 4: Opaque
Liver	Inflammatory reaction	0: None 1: Fibrin covering $< 25\%$ 2: Fibrin covering $> 25\%$ 3: Fibrin covering all the liver
Airsac (left)	Transparency of airsac	0: Clear 1: Nuclear 2: Cloudy 3: Milky 4: Opaque
	Amount of exudate	0: None 1: Sparse 2: Some 3: Abundant
Lung (left)	Amount of exudate	0: None 1: Sparse 2: Some 3: Abundant
	Congestion	0: None 1: Sparse ($< 25\%$ of lung tissue) 2: Pronounced ($> 25\%$ of lung tissue)
	Lung lesions (oedema)	0: None 1: Slight oedema of the alveolar walls 2: Moderate oedematous thickening of alveolar walls with occasional alveoli containing coagulated oedema fluids 3: Extensive occurrence of alveolar and interstitial oedema
	Lung lesions (size)	0: No lesions 1: Small lesions between 1 cm^2 and 2 cm^2 2: Moderate-size lesions between 3 cm^2 and 4 cm^2 3: Extensive lesions with more than 4 cm^2 and covering almost all the area of the lungs
	Granulomas	0: None 1: Few small white caseous/spherical nodules 2: Numerous small white caseous/spherical nodules 3: Many caseous lesions/nodules and green-gray mold due to sporulation
Spleen	Proliferation	0: None 1: Little proliferation 2: Distinct proliferation

Table S2. Mann-Whitney U tests to assess between-group differences in serum amyloid A levels, lesions scores, and number of colony-forming units.

A	B	U-val	p-value	Cohen's d
Serum amyloid A (mg/L)				
Control	E. coli	45.0	**2.77e-05**	-1.153
Lesion score				
Control	E. coli	39.5	**1.064e-05**	-2.121
Colony-forming units (per gram of lung)				
Control	E. coli	31.0	**1.252e-06**	-0.607

Table S3. `brms` summary tables for the behavioural features. Reference group: Control (uninfected). Significant differences (determined if the confidence intervals exclude zero) with respect to the negative control are shown in bold. CI: confidence interval (l: lower, u: upper). $\hat{R}$: Gelman-Rubin statistic [56] for Markov chain Monte Carlo convergence; convergence is deemed suitable when $\hat{R}$ is close to 1. ESS: effective sample size.

	Estimate	Est.Error	l-95% CI	u-95% CI	$\hat{R}$	Bulk ESS	Tail ESS
Distance travelled							
Intercept	-3.32	0.116	-3.543	-3.079	1.003	2353.594	1771.796
E. coli	-0.286	0.112	**-0.506**	**-0.069**	1.002	3717.03	2639.991
Rate of change in body area							
Intercept	-2.912	0.125	-3.148	-2.658	1.001	1750.502	1543.872
E. coli	-0.283	0.117	**-0.518**	**-0.049**	1.001	3034.374	2758.279
% Time near food source							
Intercept	-2.281	0.245	-2.74	-1.761	1.002	1333.479	1306.324
E. coli	-0.585	0.148	**-0.883**	**-0.302**	1.001	3112.788	2167.955

Table S4. Comparison of different families for mixed-effects modelling using `brms`. ELPD: expected log pointwise predictive density. LOOCV: leave-one-out cross-validation. SE: standard error.

Family (link)	ΔELPD		LOOCV ELPD	
	Estimate	SE	Estimate	SE
Distance travelled				
Beta (logit)			498.315	19.847
Gaussian (identity)	-153.134	8.685	345.181	19.389
Rate of change in body area				
Beta (logit)			418.771	22.354
Gaussian (identity)	-170.662	10.437	248.109	24.957
% Time spent near food source				
Zero-inflated Beta (logit)			292.026	31.327
Gaussian (identity)	-105.782	30.388	186.244	19.72