



THE UNIVERSITY OF BRITISH COLUMBIA

Animal Welfare Program

Faculty of Land and Food Systems

Which Cow is Most Lame? Redefining Lameness Assessment Using Hierarchy



Sky (Kehan)
Sheng¹



Borbala
Foris^{1,3}



Marina A.G.
von
Keyserlingk¹



Tiffany-Anne
Timbers²



Varinia
Cabrera^{1,4}



Daniel M.
Weary¹

¹Animal Welfare Program, The University of British Columbia

²Department of Statistics, The University of British Columbia

³Centre for Animal Nutrition and Welfare, University of Veterinary Medicine

⁴Institute of Electrical Engineering, University of the Republic, Uruguay



**How are training videos
labelled to teach the machine
about lame cows?**





**Gait scoring:
5 experts, 3 rounds**



30 videos

Traditional gait scoring from 5 experts 3 rounds



Interobserver reliability

ICC = 0.42 - 0.47 (0.44 ± 0.02)

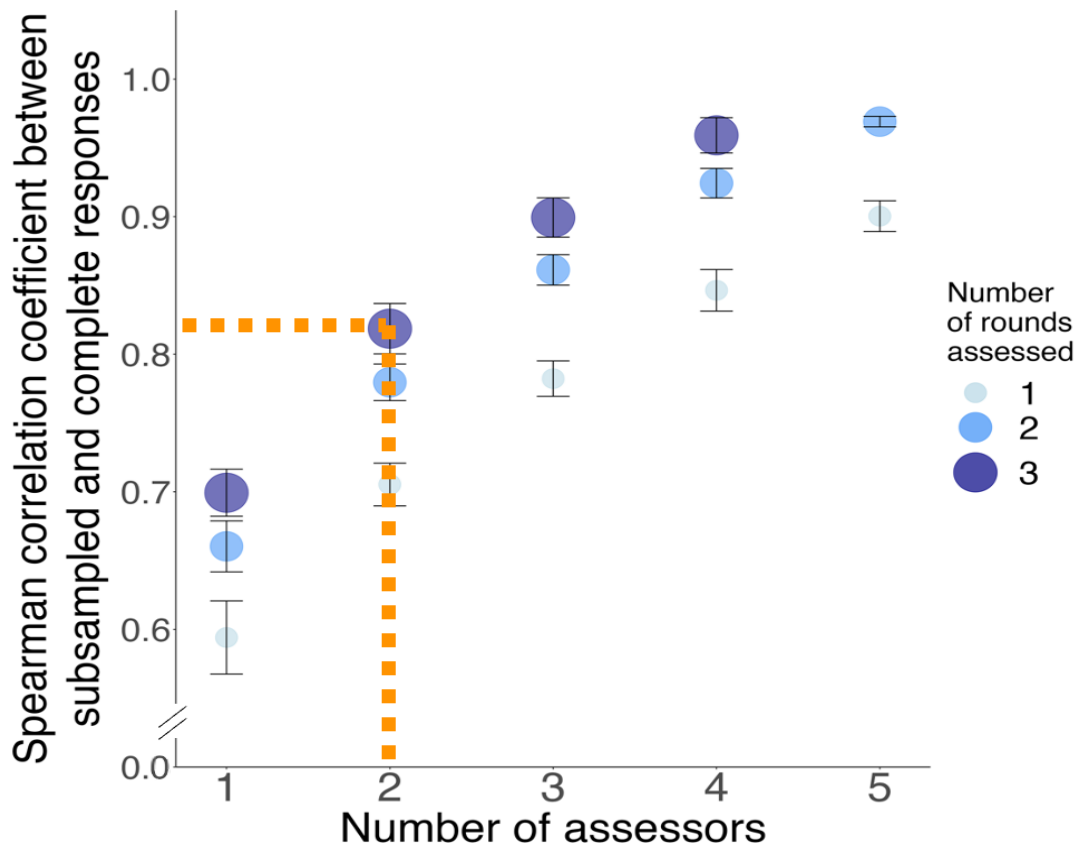


X 3

Intraobserver reliability

ICC = 0.49 - 0.73 (0.62 ± 0.09)

How many experts and how many rounds do we need when using gait scoring?





Crowdsourcing remote comparative lameness assessments for dairy cattle

Kehan Sheng,^{1*} Borbala Foris,^{1*} Marina A. G. von Keyserlingk,¹ John Gardenier,² Cameron Clark,³ and Daniel M. Weary^{1†}

¹Animal Welfare Program, Faculty of Land and Food Systems, The University of British Columbia, Vancouver, BC, V6T 1Z6, Canada

²Australian Centre for Field Robotics, Faculty of Engineering, The University of Sydney, Darlington, NSW 2006, Australia

³Livestock Production and Welfare Group, School of Life and Environmental Sciences, Faculty of Science, The University of Sydney, Camden, NSW 2570, Australia

Which cow is more lame, and by how much? (10 questions)

Question 3 of 10



The left cow is
much more lame



The two cows are
about the same

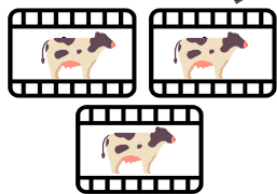


The right cow is
much more lame

☐ The video(s) is not showing.



**Gait scoring:
5 experts, 3 rounds**



30 videos



**435 pairwise
lameness
comparisons**



4 experts

EloSteepness



**Lameness
hierarchy**

Lameness hierarchies from 4 experts



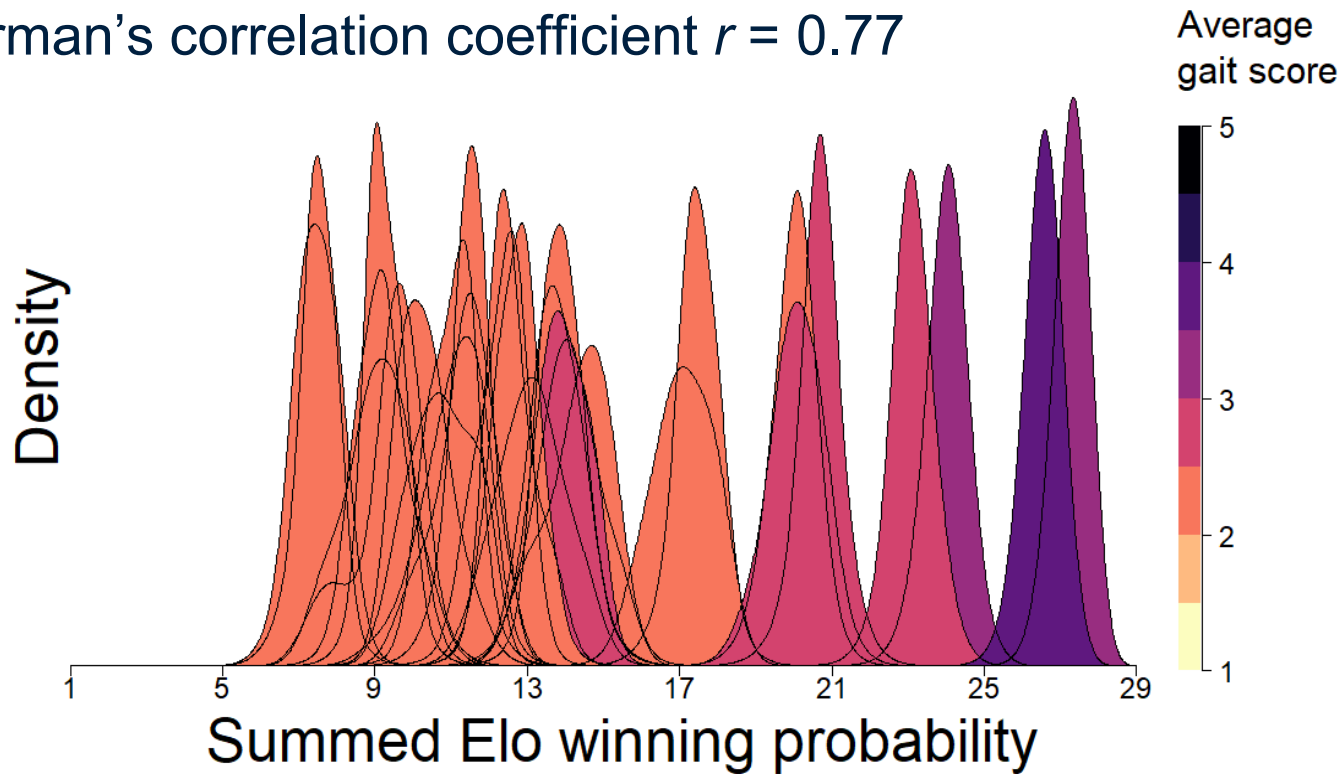
Interobserver reliability

$$\text{ICC} = 0.81$$

Lameness hierarchy based on pairwise assessment VS absolute gait scoring

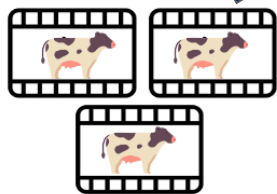


The Spearman's correlation coefficient $r = 0.77$





**Gait scoring:
5 experts, 3 rounds**



30 videos



**435 pairwise
lameness
comparisons**



4 experts

EloSteepness



**Lameness
hierarchy**



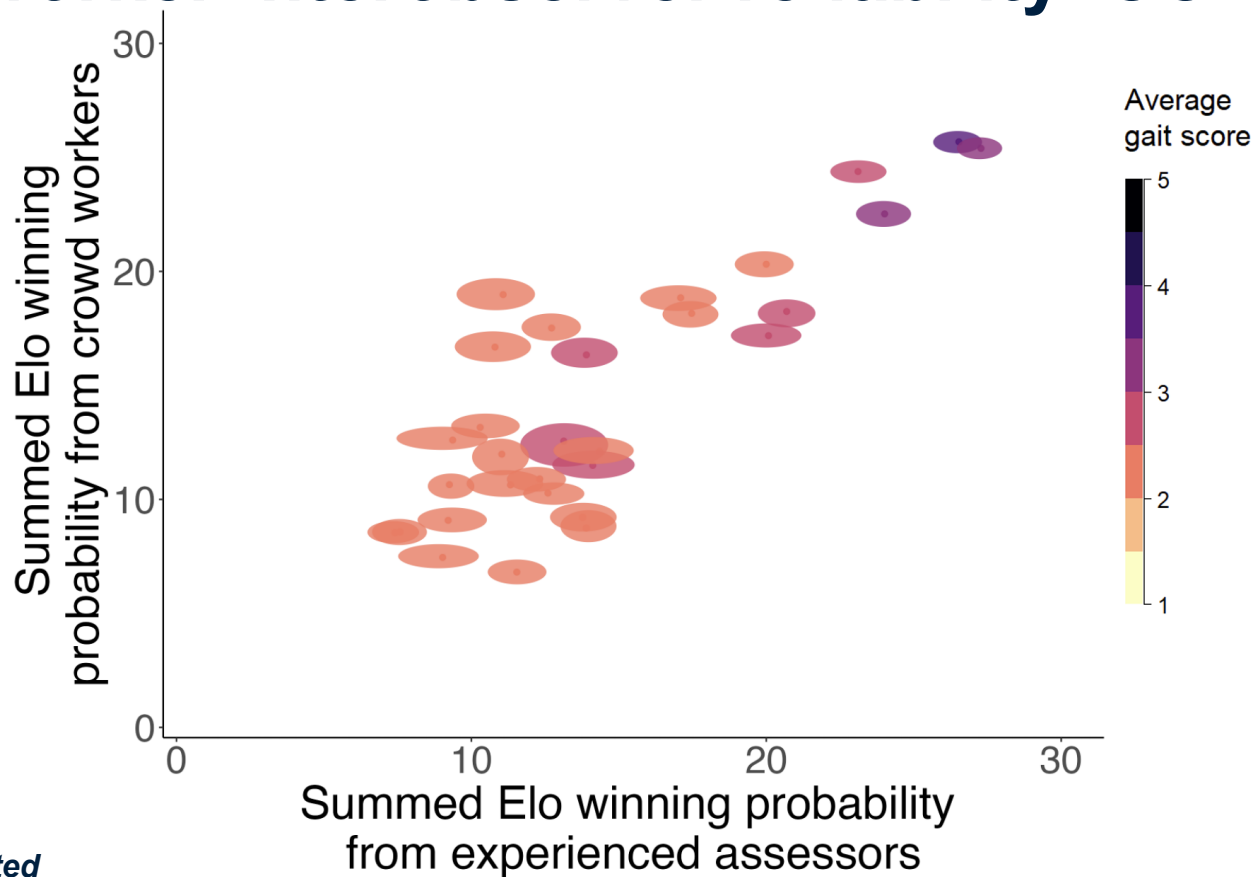
**22 crowd workers
from Amazon MTurk**

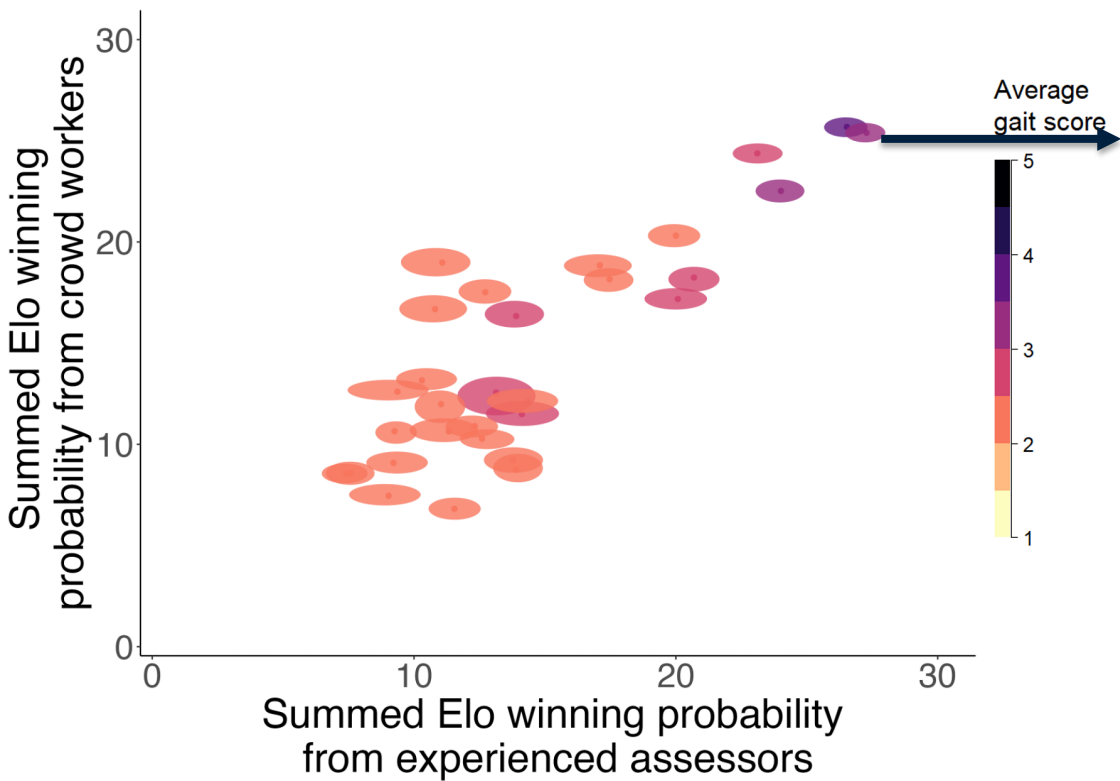
EloSteepness

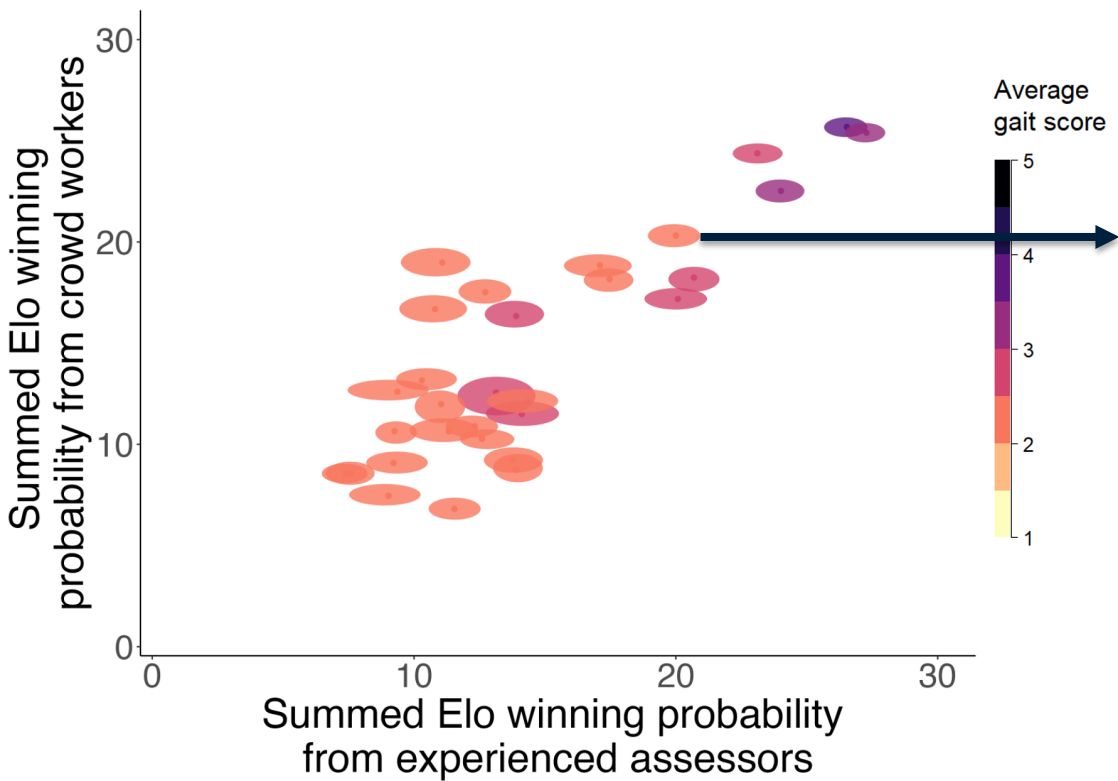


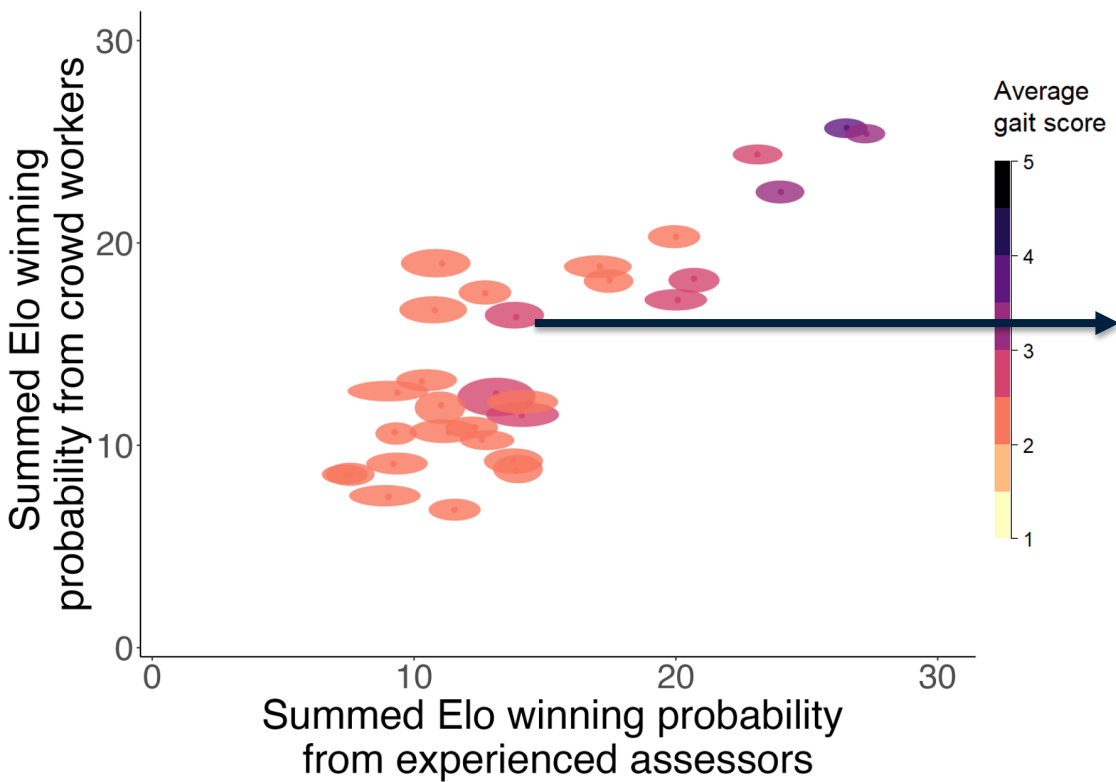
**Lameness
hierarchy¹²**

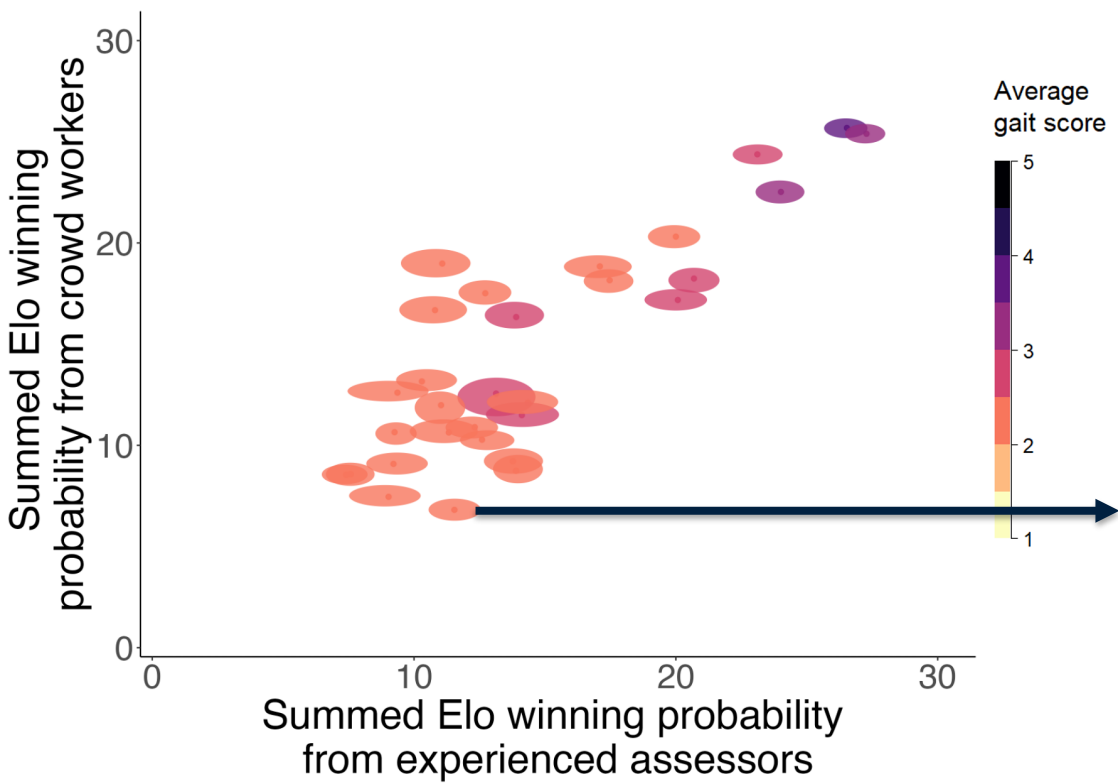
Lameness hierarchy from all experts VS all crowd worker Interobserver reliability ICC = 0.85







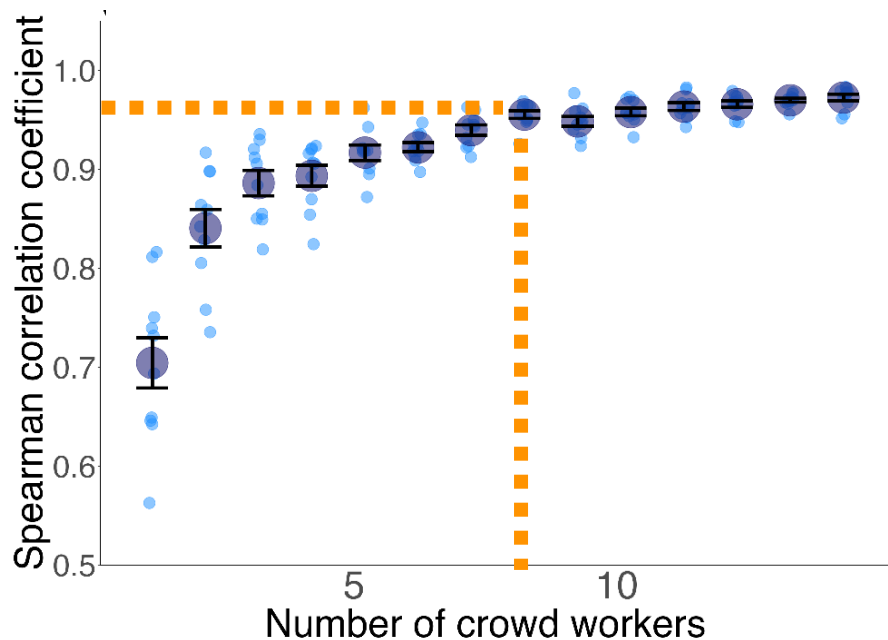




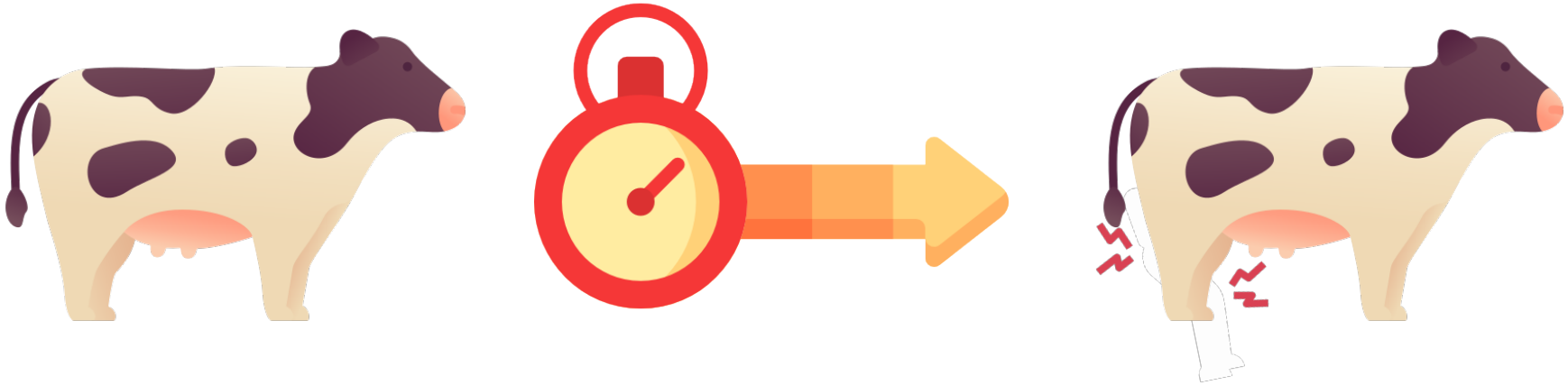
More efficient algorithms



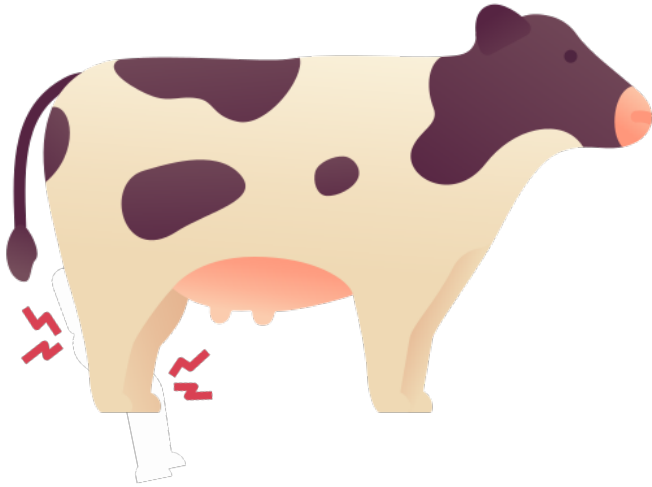
- Reduction of 61% pairwise comparisons
- Minimum of 8 crowd workers



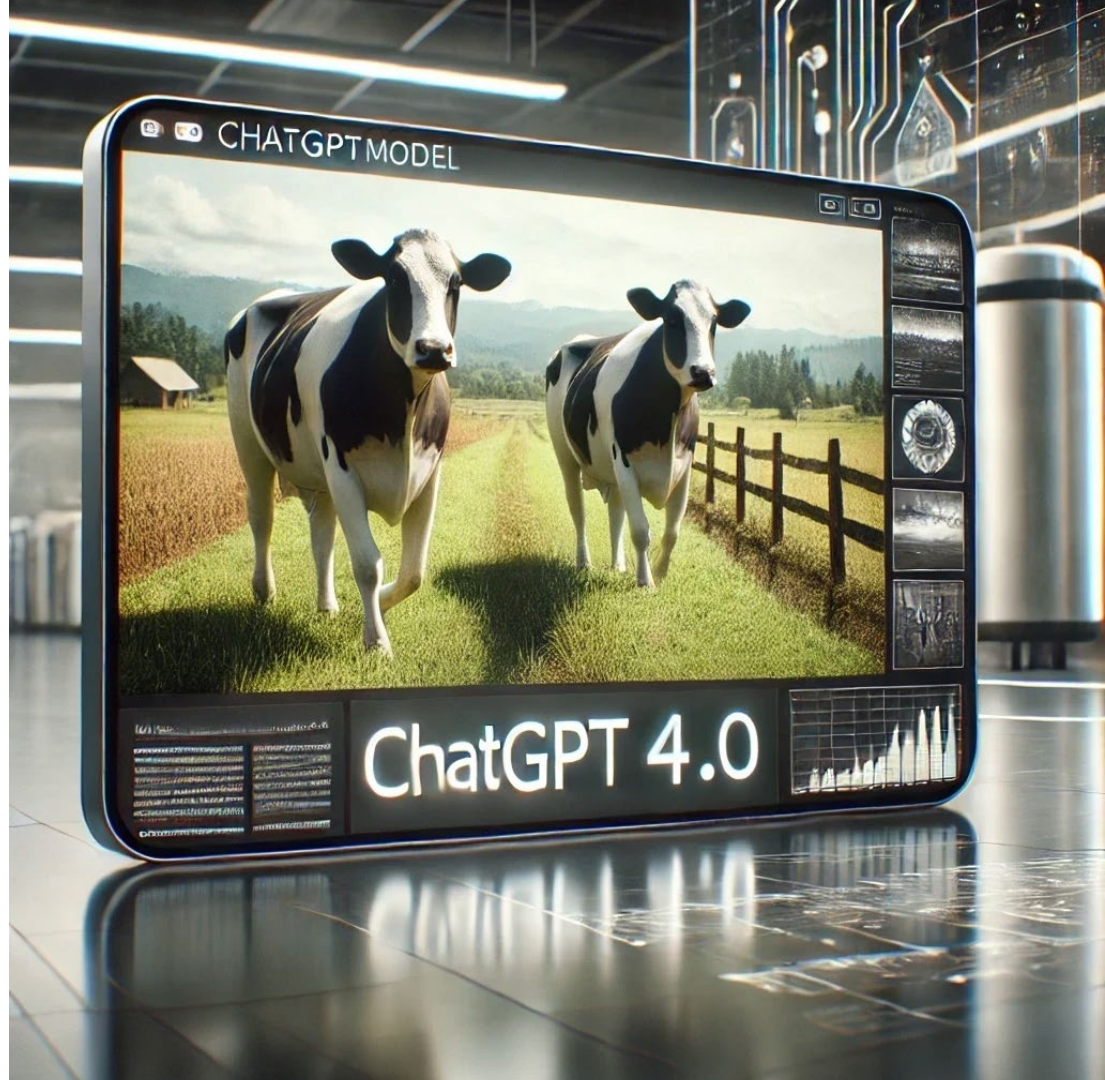
What's next? – Longitudinal tracking of lameness within each individual



What's next? – Validation from hoof pathology and cows' response to treatment



**What's next? –
AI?**



Take home message

- We constructed lameness hierarchy by comparing which cow is more lame between every 2 cows
- Experienced assessors show higher interobserver reliability (ICC = 0.81) when using lameness hierarchy approach compared to traditional gait scoring (ICC = 0.44)
- The lameness hierarchy by crowd workers is in high agreement with that from experienced assessors (ICC = 0.85)





THE UNIVERSITY OF BRITISH COLUMBIA

Animal Welfare Program

Faculty of Land and Food Systems

Questions?



sky.sheng@ubc.ca