



# Causal inference reveals long-term impact of health events on production performance in Holstein cattle

Ao Wang<sup>1</sup>, Liyun Han<sup>2</sup>, Ya-Chun Wang<sup>1\*</sup>

(1 College of Animal Science and Technology, China Agricultural University, Beijing, 100193, China;  
2 College of Animal Science and Technology, Ningxia University, Yinchuan, 750021, China)

## INTRODUCTION & AIM

- Health events represents a key constraint to **farm sustainability** in the dairy industry.
- Previous research tended to focus on **short-term milk loss** during disease periods.
- In the long-term, milk yield is usually not fully recovered even after the disease is **cured**.
- The gold standard for obtaining causal effects is the **randomized experiment**, but it's not always feasible in the real world.
- Due to the **confounding factors**, inferring causal effects between variables from farm-recorded data is challenging.

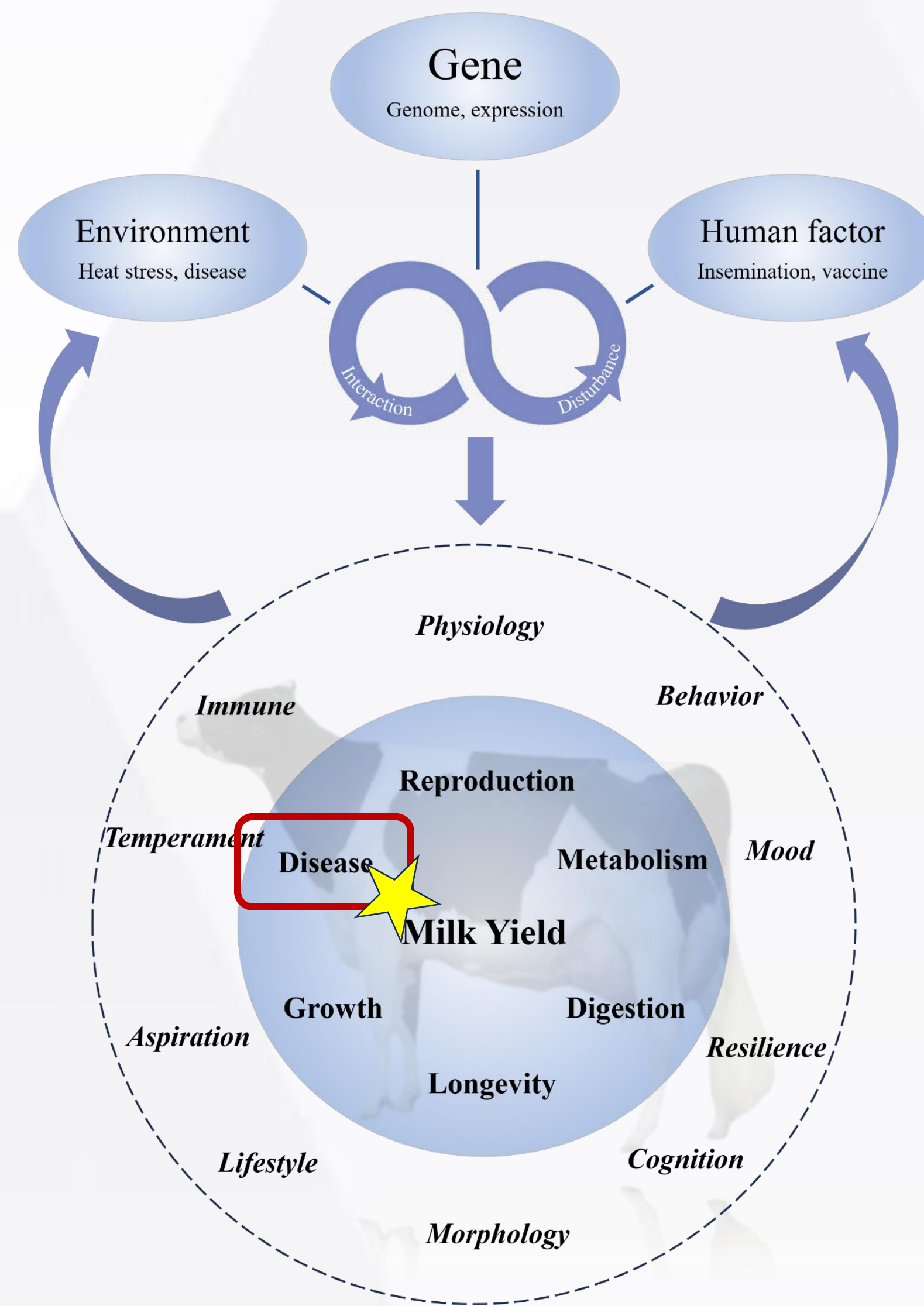


Fig 1. Challenges and responses for dairy cattle

- To quantify the **long-term impacts** of different health events on milk yield through causal inference strategies.
- To investigate the differences in the effects of **single and secondary onset** of diseases on milk yield.

## METHOD

### Milk yield data

- High-throughput session milk yield records for 37,246 Holstein cattle.

### Health events

- 40,699 health event records from January 2020 to April 2024.

- Confounders: herd-year-season, parity, stillbirth, calving ease, the number of inseminations, and the genetic levels of milk yield and resilience.

Propensity score matching (PSM)

Overlap weighting (OW)

Convergent cross-mapping (CCM)

- Causal effects of diseases on **overall** (305d milk yield and variation) and **local** (average daily milk yield and variation within a specific period) lactation features

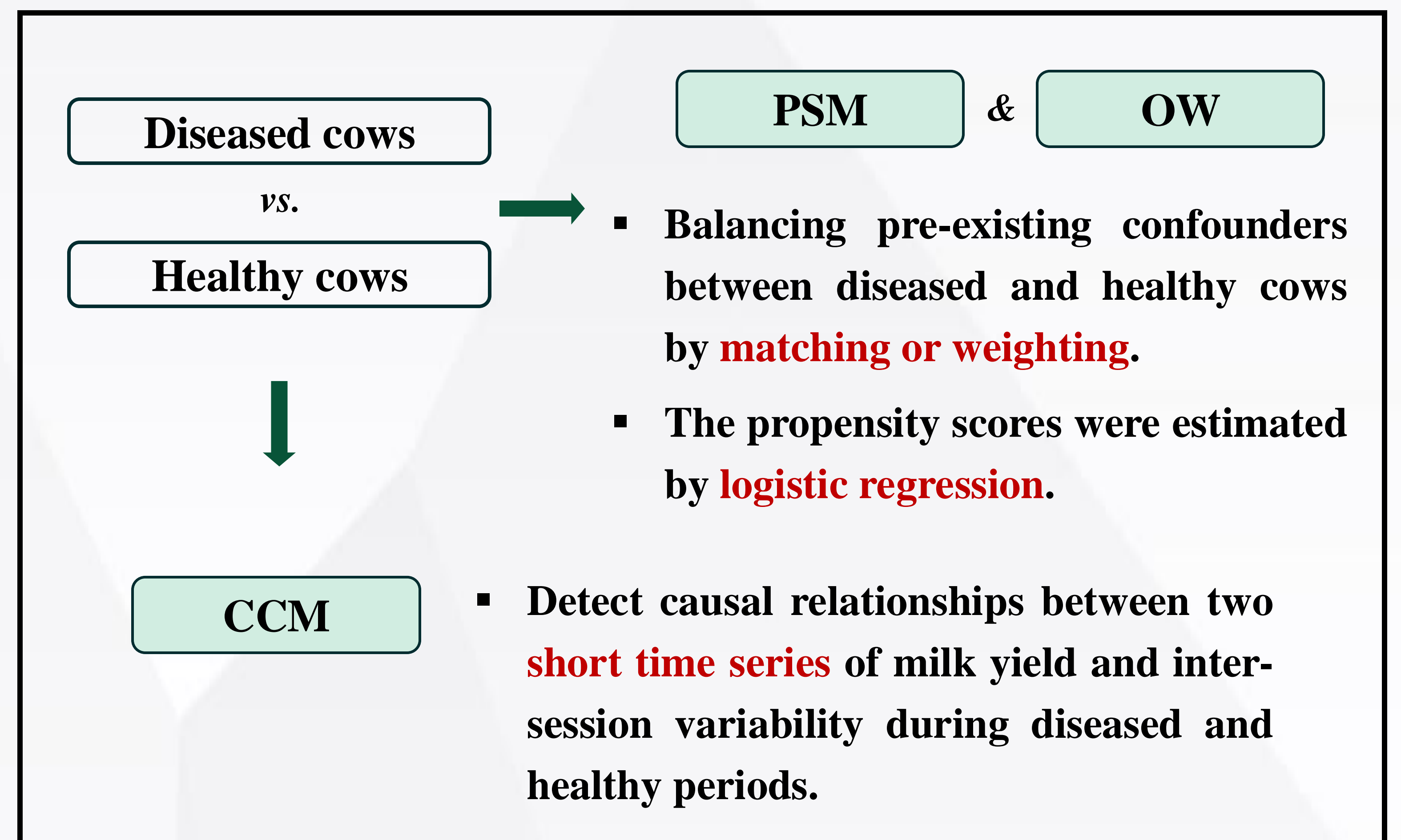


Fig 2. The description of the three causal inference strategies

## RESULTS & DISCUSSIONS

- Five diseases: udder health, reproductive disorders, metabolic disorders, digestive disorders, and hoof health.
- Two lactation features: milk yield and average coefficient of variation for session milk yields within a day.

	First onset		Secondary onset	
	Disease period	After cured	Disease period	After cured
Udder health	13.12%	7.88%	13.46%	7.17%
Reproductive disorders	11.17%	4.99%	12.82%	1.30%
Metabolic disorders	20.78%	4.06%	36.62%	5.53%
Digestive disorders	24.76%	8.42%	25.08%	8.05%
Hoof health	15.53%	8.00%	11.67%	3.80%

- During the period of single disease, daily milk yield decreased by **16.06%**.
- The long-term effect of second disease onset would be **weakened** to 3.28%.
- The coefficient of variation for session milk yields would **fluctuate sharply and significantly** around the onset day.

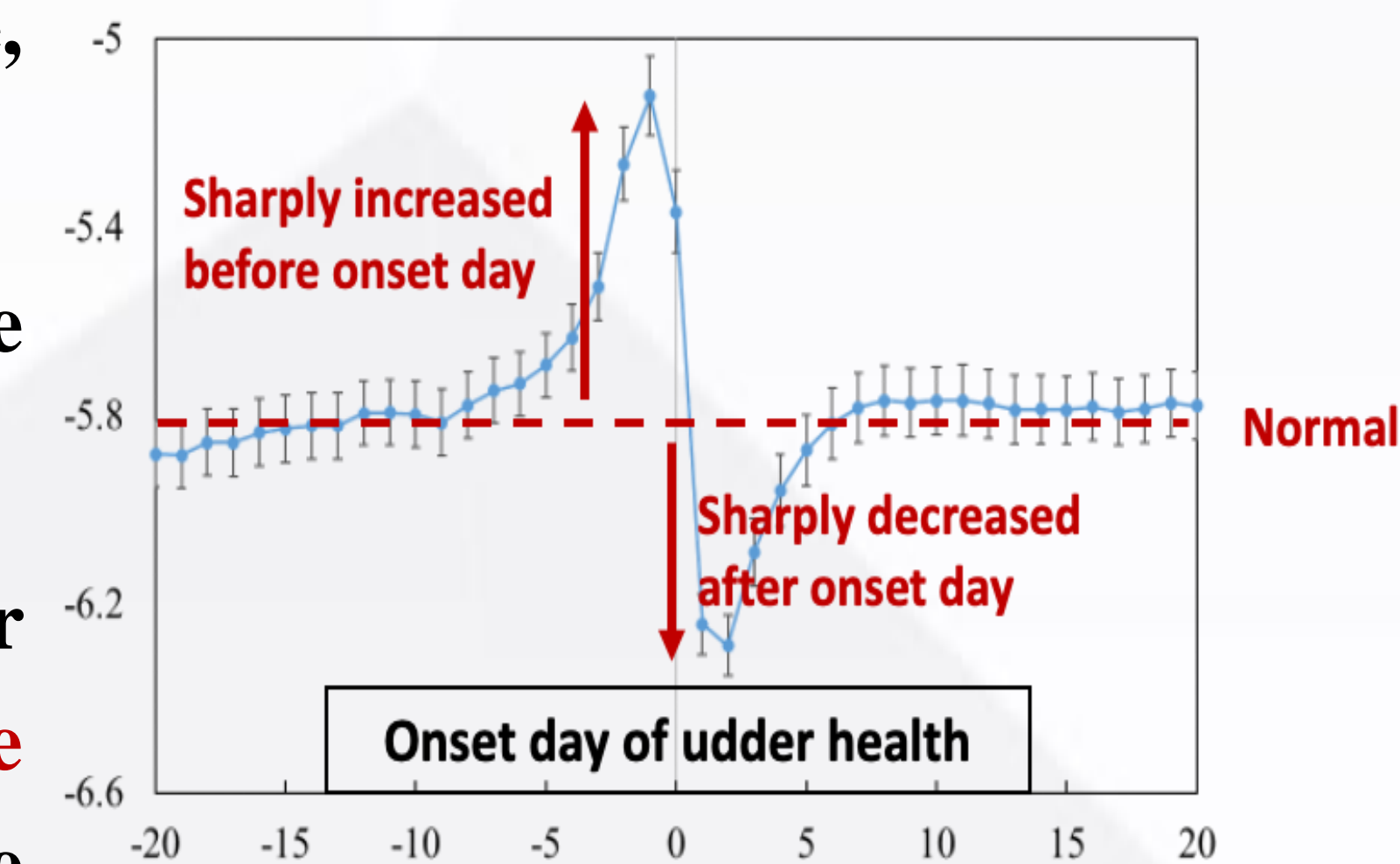


Fig 3. The fluctuation of the coefficient of variation for session milk yield

## CONCLUSION

- This study confirms the **long-term impact** of health events on production performance and shows the potential of **causal inference** to serve as a powerful tool for precision livestock farming.

## Main Reference

- Ferreira, V.C., Thomas, D.L., Valente, B.D., et al. Causal effect of prolificacy on milk yield in dairy sheep using propensity score. 2017. J. Dairy Sci. 100(10), 8443-8450.