Link between viral genetics and clinical impact in populations: Can artificial intelligence outperform humans?

Zvonimir Poljak, Associate Professor
Department of Population Medicine
Ontario Veterinary College
University of Guelph, ON, Canada

Big Bug Day
Stratford, ON
December 8th 2021

UNIVERSITY of GUELPH

IMPROVE LIFE.


CORONAVIRUS | News

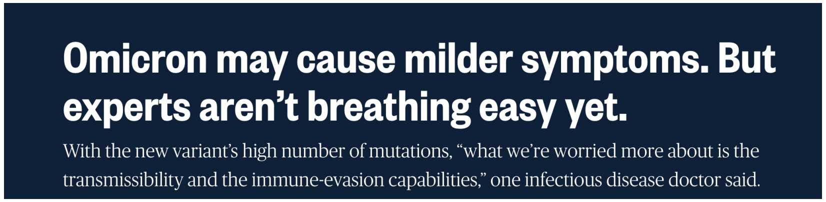
Death rate from COVID-19 variants 60 per cent higher: epidemiologist

Cameron French CTVNews.ca writer
@camjfrench | Contact

Published Friday, March 26, 2021 2:17PM EDT

Coronavirus variant first identified in UK appears to be more deadly, study suggests

By Jacqueline Howard, CNN
Updated 12:13 PM ET, Wed March 10, 2021



Omicron may cause milder symptoms. But experts aren't breathing easy yet.

With the new variant's high number of mutations, "what we're worried more about is the transmissibility and the immune-evasion capabilities," one infectious disease doctor said.

Covid: Delta variant patients twice as likely to need hospital care

3 days ago

'Some evidence' British coronavirus variant causes more deaths

Prime Minister Boris Johnson said evidence suggests vaccines still work against the new variant.

Covid: 'More deadly' UK variant claim played down by scientists

23 January

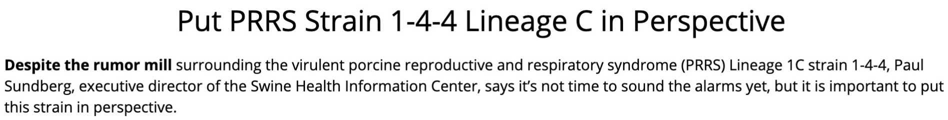
New PRRS Strain Kills Sows, Piglets

PRRS Strain 1-4-4 Results in 'More Dramatic' Hog Mortality

2/8/2021 | 11:44 AM CST

By Russ Quinn , DTN Staff Reporter
Connect with Russ:
@RussQuinnDTN

OMAHA (DTN) -- The last thing the world needs right now is another virus. But U.S. hog producers are facing another health challenge to their herds in the form of a familiar yet slightly different foe. The most recent strain (1-4-4) of the Porcine Reproductive and Respiratory Syndrome (PRRS) virus is nothing producers have not seen before with PRRS, but it is more dramatic, resulting in higher rates of hog mortality, according to swine veterinarians.



Put PRRS Strain 1-4-4 Lineage C in Perspective

Despite the rumor mill surrounding the virulent porcine reproductive and respiratory syndrome (PRRS) Lineage 1C strain 1-4-4, Paul Sundberg, executive director of the Swine Health Information Center, says it's not time to sound the alarms yet, but it is important to put this strain in perspective.

HOG PRODUCTION

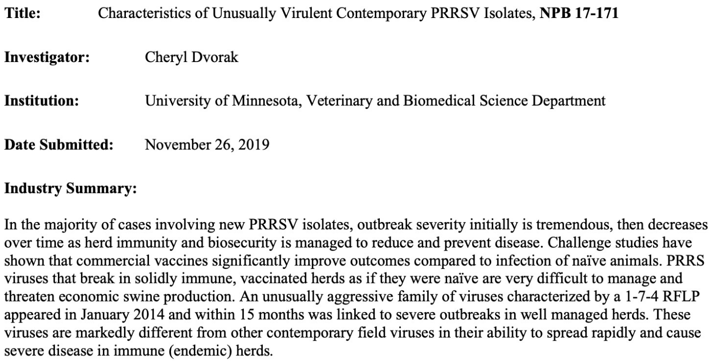
PRRS Strain 1-4-4: The Most Dramatic Strain I've Seen, Yeske Says

---

Genomic sequence and virulence comparison of four Type 2 porcine reproductive and respiratory syndrome virus strains

Susan L. Brockmeier[a,*], Crystal L. Loving[a], Ann C. Vorwald[a], Marcus E. Kehrli Jr.[a], Rodney B. Baker[b], Tracy L. Nicholson[a], Kelly M. Lager[a], Laura C. Miller[a], Kay S. Faaberg[a]

[a] *Virus and Prion Research Unit, National Animal Disease Center, USDA, Agricultural Research Service, Ames, IA, United States*
[b] *Veterinary Diagnostic and Production Animal Medicine, Iowa State University, Ames, IA, United States*



Title: Characteristics of Unusually Virulent Contemporary PRRSV Isolates, **NPB 17-171**

Investigator: Cheryl Dvorak

Institution: University of Minnesota, Veterinary and Biomedical Science Department

Date Submitted: November 26, 2019

Industry Summary:

In the majority of cases involving new PRRSV isolates, outbreak severity initially is tremendous, then decreases over time as herd immunity and biosecurity is managed to reduce and prevent disease. Challenge studies have shown that commercial vaccines significantly improve outcomes compared to infection of naïve animals. PRRS viruses that break in solidly immune, vaccinated herds as if they were naïve are very difficult to manage and threaten economic swine production. An unusually aggressive family of viruses characterized by a 1-7-4 RFLP appeared in January 2014 and within 15 months was linked to severe outbreaks in well managed herds. These viruses are markedly different from other contemporary field viruses in their ability to spread rapidly and cause severe disease in immune (endemic) herds.

Transboundary and Emerging Diseases

ORIGINAL ARTICLE

Emergence of a novel highly pathogenic porcine reproductive and respiratory syndrome virus in China

J.-k. Liu, X. Zhou, J.-q. Zhai, B. Li, C.-h. Wei, A.-l. Dai, X.-y. Yang, M.-l. Luo

First published: 14 February 2017 | https://doi.org/10.1111/tbed.12617 | Citations: 26

OPEN ACCESS Freely available online PLoS one

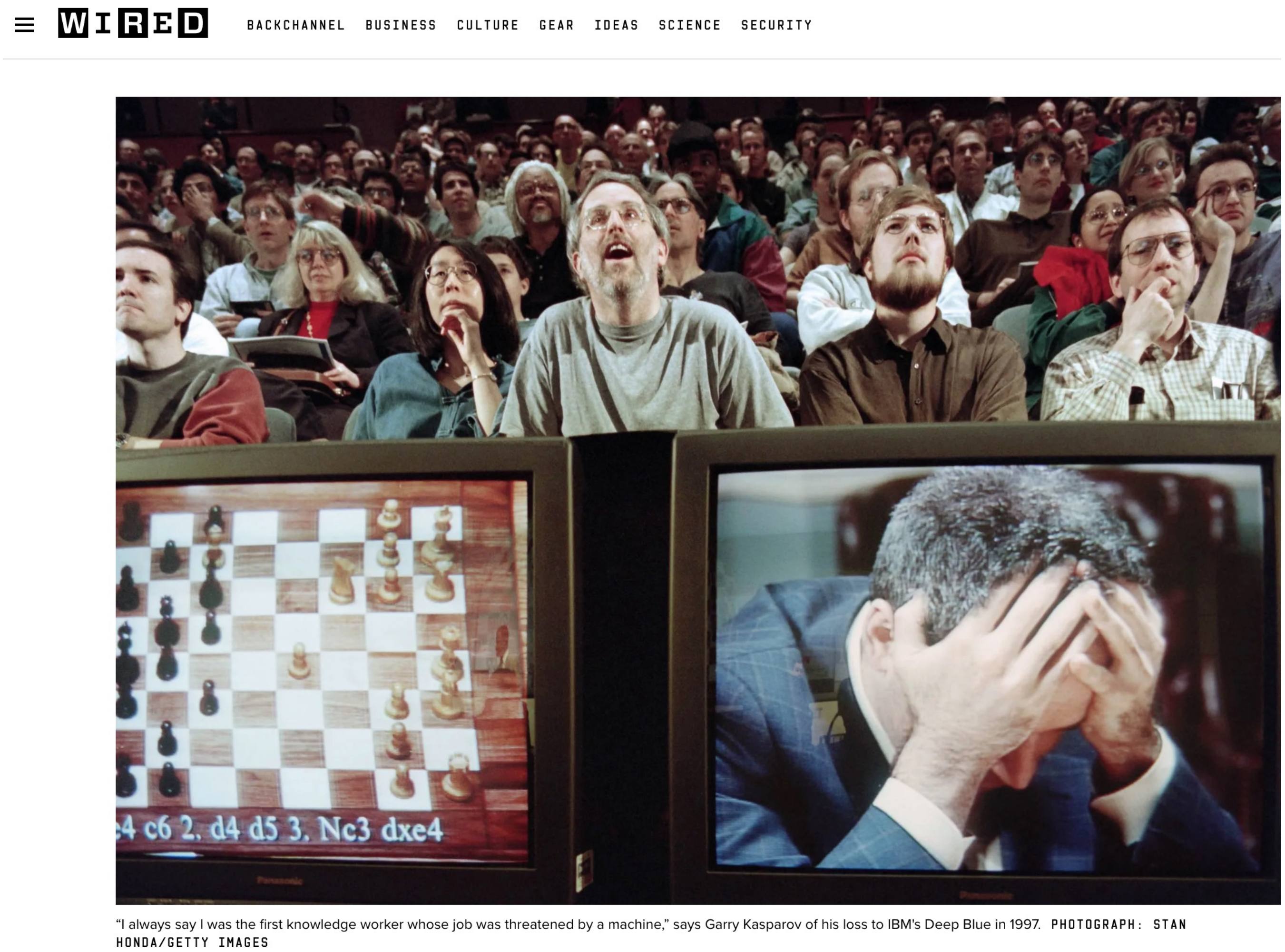
Emergence of Fatal PRRSV Variants: Unparalleled Outbreaks of Atypical PRRS in China and Molecular Dissection of the Unique Hallmark

Kegong Tian[1*], Xiuling Yu[1], Tiezhu Zhao[1], Youjun Feng[2,3], Zhen Cao[1], Chuanbin Wang[1], Yan Hu[1], Xizhao Chen[1], Dongmei Hu[1], Xinsheng Tian[2,3], Di Liu[2], Shuo Zhang[1], Xiaoyu Deng[1], Yinqiao Ding[1], Lu Yang[1], Yunxia Zhang[1,4], Haixia Xiao[2,3], Mingming Qiao[1,4], Bin Wang[1,4], Lili Hou[1,4], Xiaoying Wang[1], Xinyan Yang[1], Liping Kang[1], Ming Sun[1], Ping Jin[1], Shujuan Wang[2,4], Yoshihiro Kitamura[2,5], Jinghua Yan[2], George F. Gao[2,5*]

Long Term Vision

- Create applications that will assist with clinical decision making in the field and will:
  - Utilize accurate and relevant data that are routinely collected in medical and other records (e.g. PigKnows, PigChamp, MetaFarms, etc)
  - Have known accuracy (predictive ability)
  - Be deployed to end users as application only if predictive ability warrants it

Historical motivation (1997)



"I always say I was the first knowledge worker whose job was threatened by a machine," says Garry Kasparov of his loss to IBM's Deep Blue in 1997. PHOTOGRAPH: STAN HONDA/GETTY IMAGES

"So machines will be dominant in the closed systems, whether it's games, or any other world designed by humans." G. Kasparov

H5 influenza study

- 2,368 sequences collected from literature and influenza research database
- Classification based on presence of high-path marker
- Various Machine Learning approaches
- Accuracy

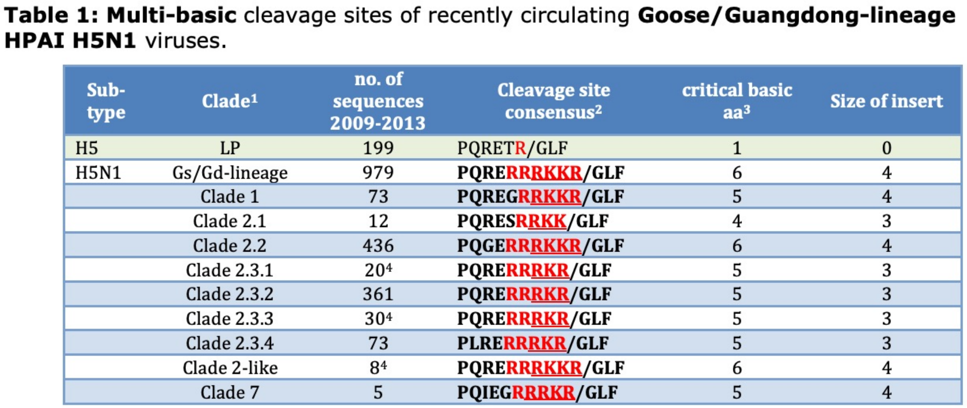


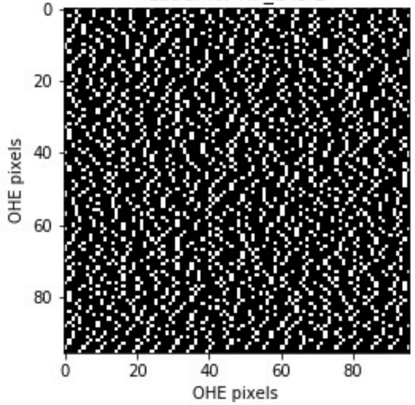
Table 1: Multi-basic cleavage sites of recently circulating **Goose/Guangdong-lineage HPAI H5N1** viruses.

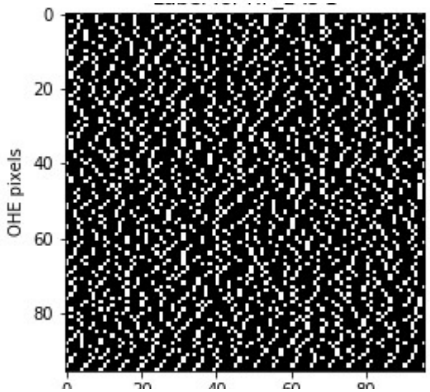
Sub-type	Clade[1]	no. of sequences 2009-2013	Cleavage site consensus[2]	critical basic aa[3]	Size of insert
H5	LP	199	PQRETR/GLF	1	0
H5N1	Gs/Gd-lineage	979	**PQRERRRKKR/GLF**	6	4
	Clade 1	73	**PQREGRRKKR/GLF**	5	4
	Clade 2.1	12	**PQRESRRKK/GLF**	4	3
	Clade 2.2	436	**PQGERRRKKR/GLF**	6	4
	Clade 2.3.1	20[4]	**PQRERRRKR/GLF**	5	3
	Clade 2.3.2	361	**PQRERRRKR/GLF**	5	3
	Clade 2.3.3	30[4]	**PQRERRRKR/GLF**	5	3
	Clade 2.3.4	73	**PLRERRRKR/GLF**	5	3
	Clade 2-like	8[4]	**PQRERRRKKR/GLF**	6	4
	Clade 7	5	**PQIEGRRRKR/GLF**	5	4

Source: Chadha et al, submitted

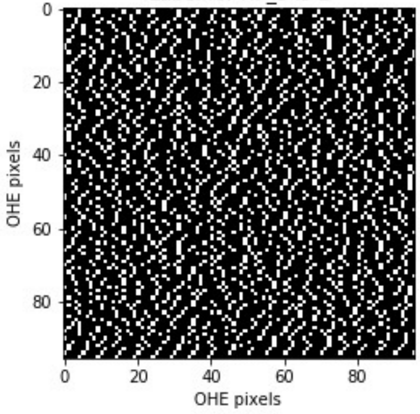
High-pathogenic H5 AI

Label for HP_1 is 1



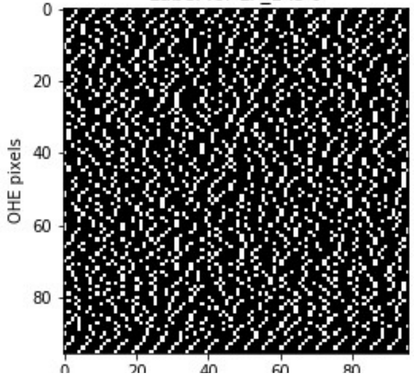


Label for HP_3 is 1

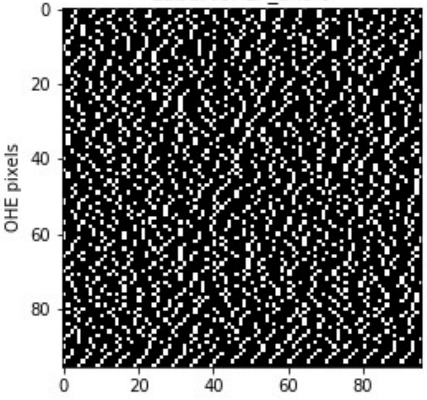


Low-pathogenic H5 AI

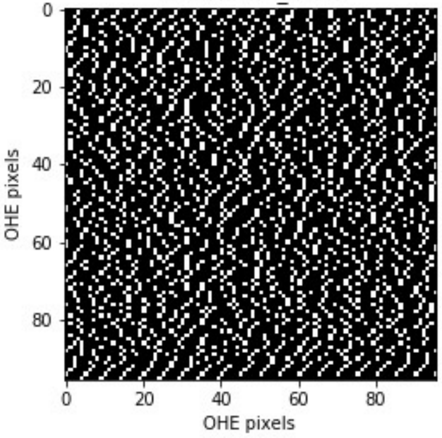
Label for LP_1 is 0



Label for LP_2 is 0

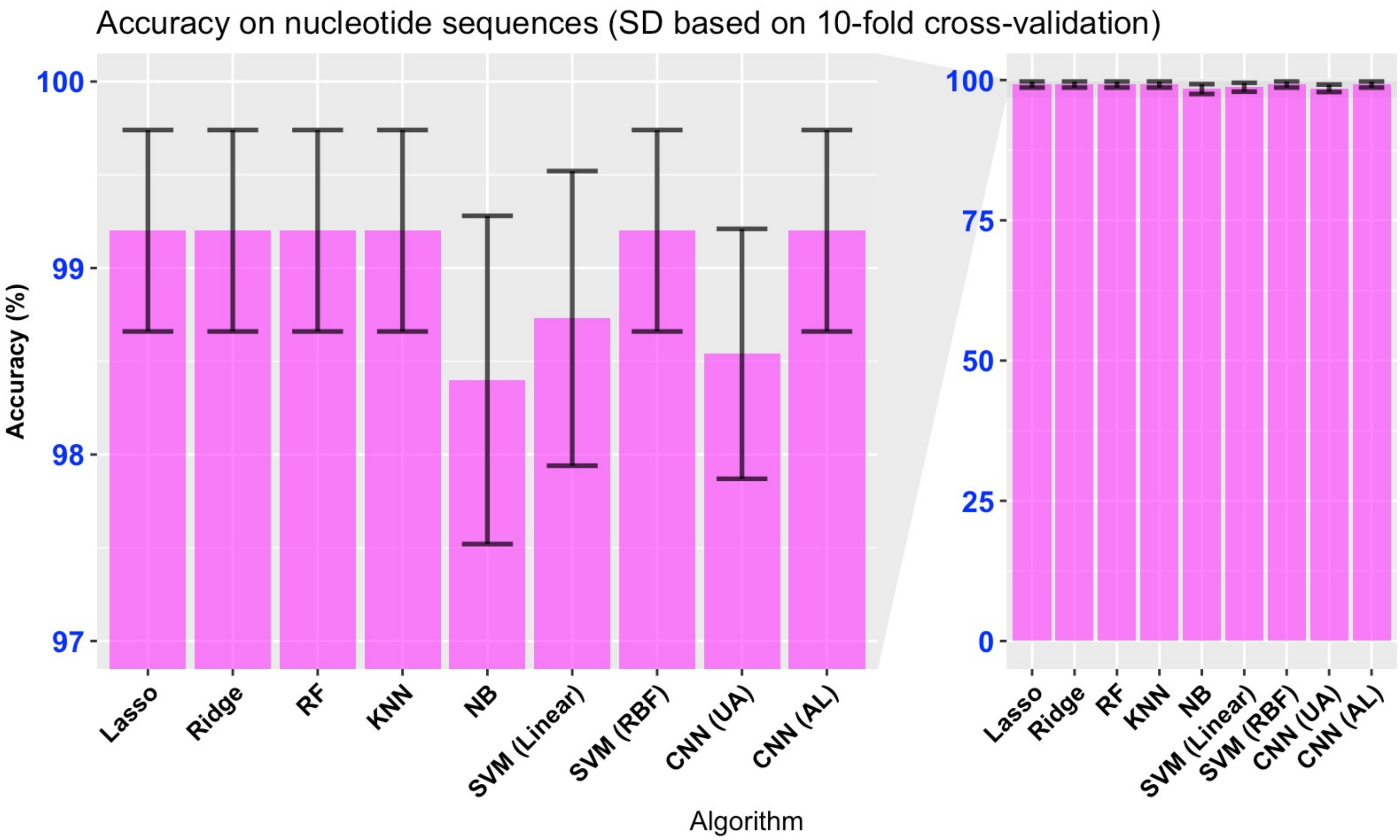


Label for LP_3 is 0



Source: Chadha Akshay

Performance on Avian Influenza Virus - Excellent



~48%

Majority class frequency = 50.76%

Source: Chadha et al, submitted

Can we do the same for PRRS: small data

- Outcome
  - 30 herds
  - Abortion, preweaning and sow mortality over 8 weeks

- Predictors
  - Nucleotide sequences organized into 10 distinct clades (variants)
  - Demographic and management data

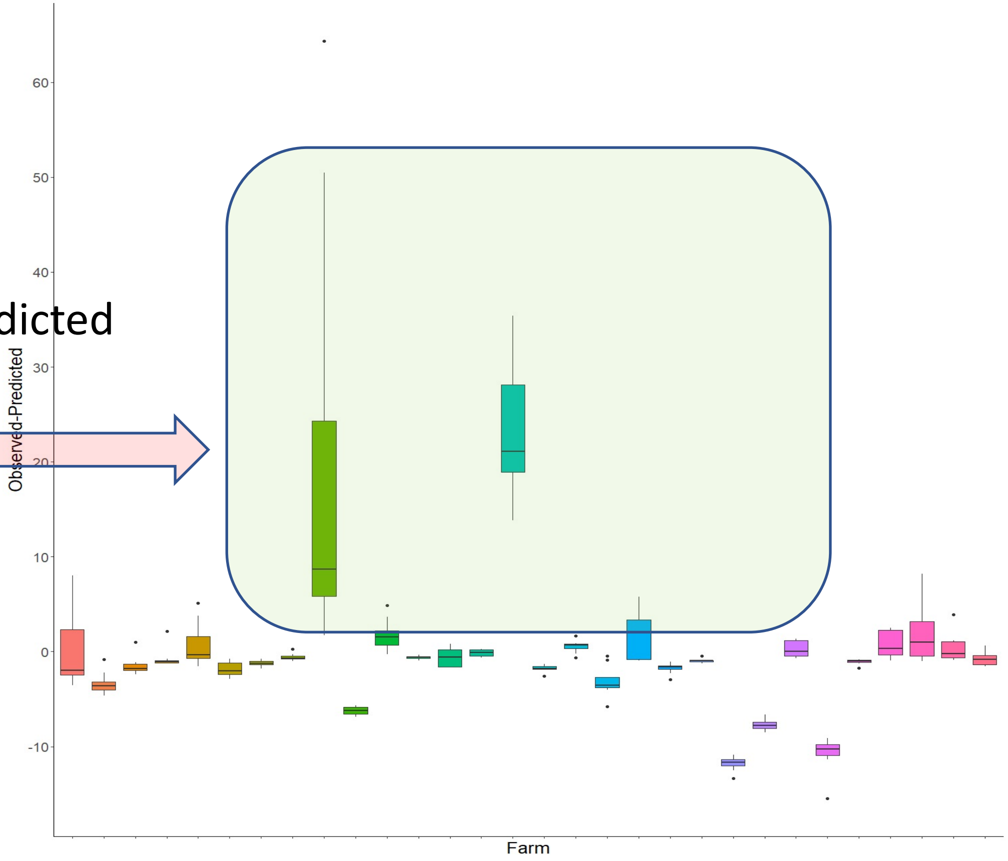
Source: Melmer et al, Prev Vet Med, 2021. Melmer et al, CJVR, 2021

Clinical measure	Model	RMSE	R^2	CCC	CCC 95% CI
Abortions	RepeatedCV	4.78	0.57	0.57	0.56, 0.58
	GroupedCV	6.20	0.07	-0.12	-0.21, 003
Sow mortality	RepeatedCV	2.26	0.17	0.30	0.28, 0.31
	GroupedCV	2.53	0.05	0.11	0.02, 0.19
Pre-weaning mortality	RepeatedCV	11.32	0.49	0.59	0.58, 0.60
	GroupedCV	17.00	0.09	-0.05	-0.17, 0.06

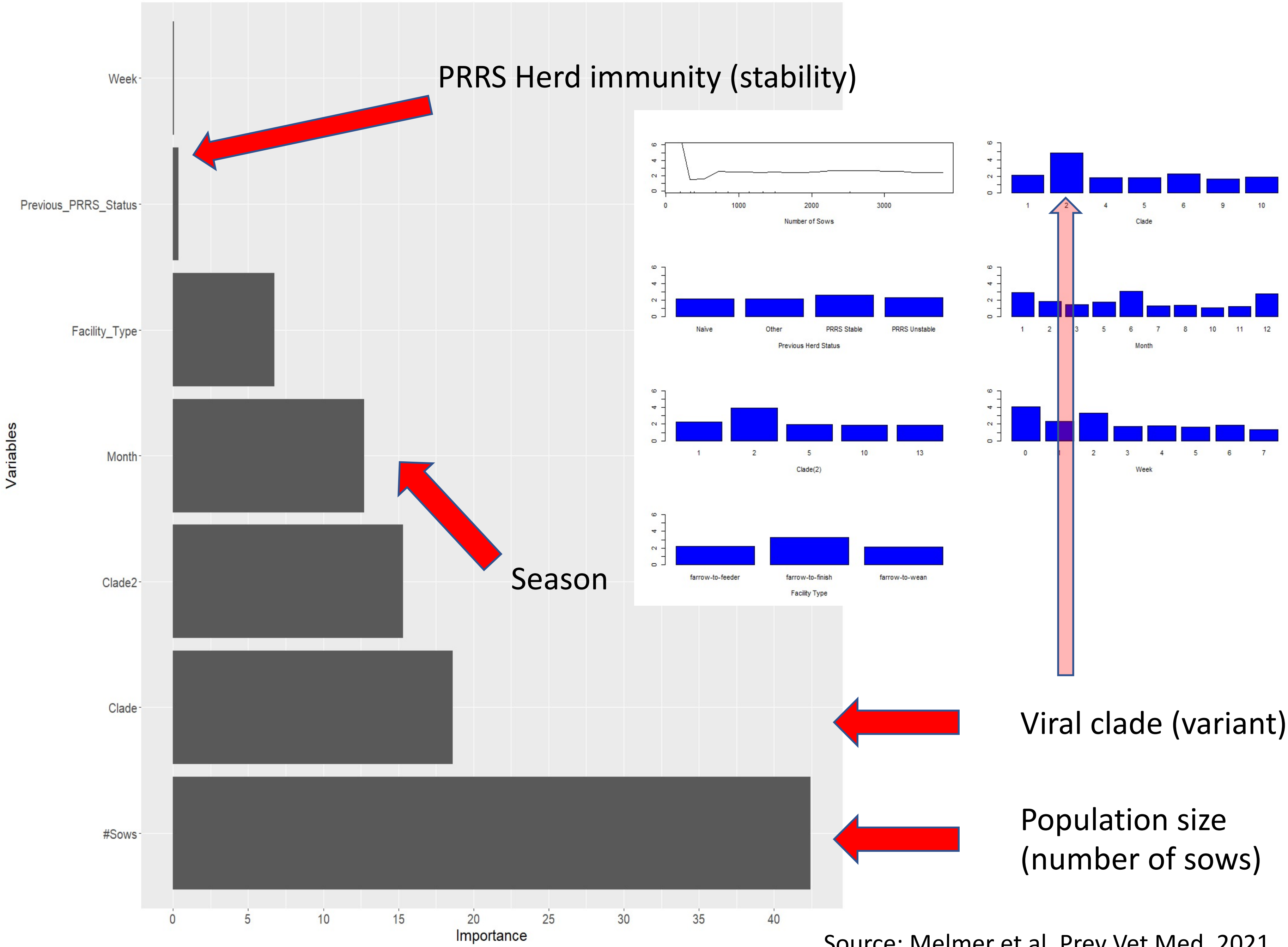
High impact outbreaks always underpredicted

No commonalities

Different variants



Source: Melmer et al, Prev Vet Med, 2021.



PRRS Herd immunity (stability)

Season

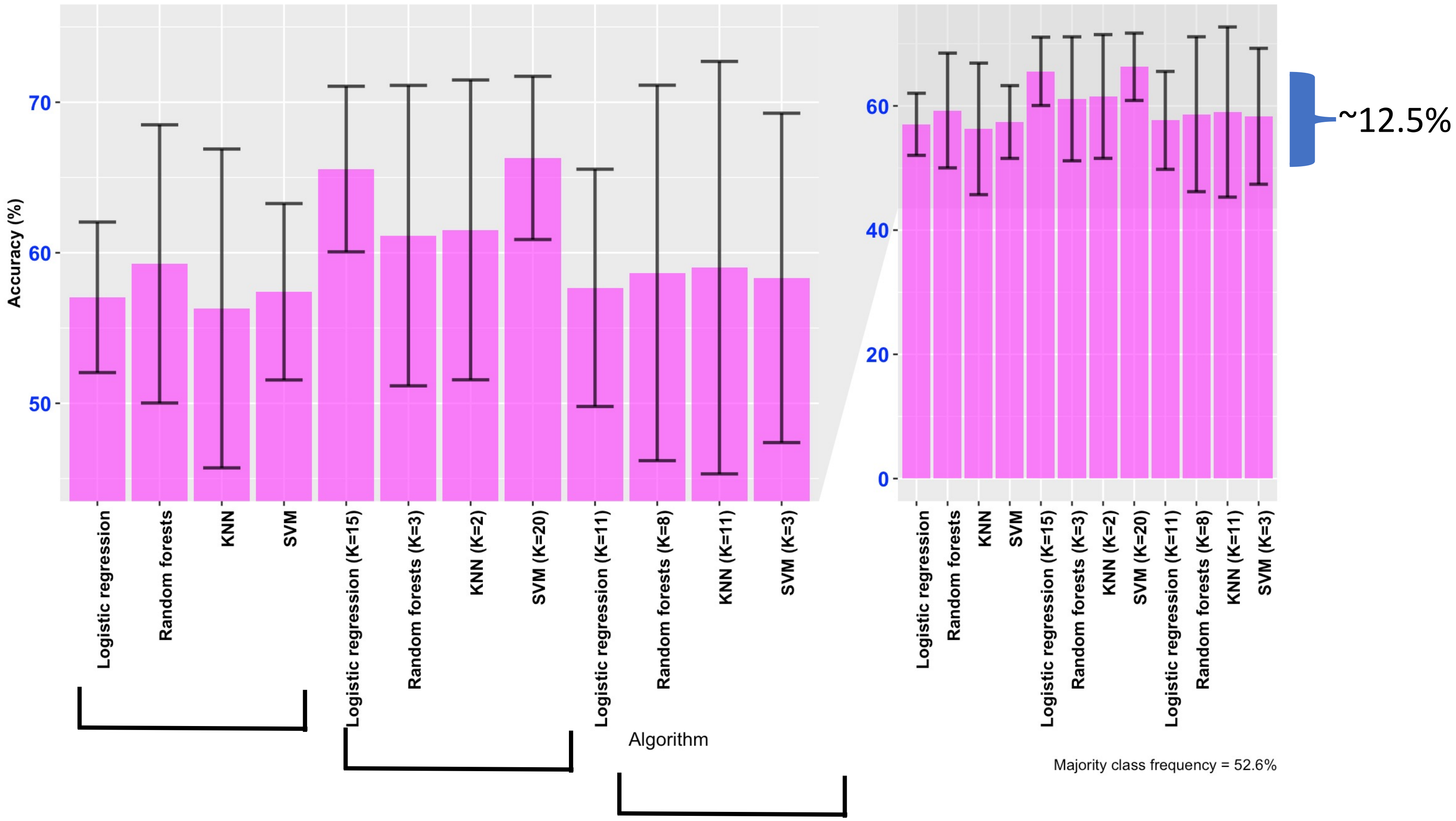
Viral clade (variant)

Population size
(number of sows)

Source: Melmer et al, Prev Vet Med, 2021.

Can we predict PRRS severity: large data

Accuracy on different data types (SD based on 10-fold cross-validation)



Majority class frequency = 52.6%

Demographics

DNA: PCA components

DNA: Kmers

Source: Chadha A, preliminary data 2021

Conclusions

- Avian influenza ->99% accuracy
  - Suitable candidate for deployment
  - Closed system with “known” rules
  - Can perform as good as experts
- PRRS
  - Moderate yield in accuracy
  - Not a closed system with known rules
  - Experts may still perform better than AI
    - understand subtle differences among farms
  - Input **data accuracy!**
  - Whole genome
  - Data volume for training

Acknowledgements

- Students
    - Dylan Melmer, Akshay Chadha
- Collaborators
    - Rozita Dara, Robert Friendship
- Funding
    - NSERC
    - Canada First Research Excellence Fund
        - Food from Thought
    - OMAFRA