The missing link: Cross-species EHR data linkage offers new opportunities for improving One Health

Kathleen R. Mullen[1], Nadia T. Saklou[2], Adam Kiehl[2], G. Joseph Strecker[2], Tracy Webb[2], Susan VandeWoude[2], Ian M. Brooks[3], Toan C. Ong[3], Sabrina Toro[1], Melissa A. Haendel[1]
[1]University of North Carolina at Chapel Hill,[2] Colorado State University, [3] University of Colorado Anschutz Medical Campus

Background

Accelerating precision medicine within the learning healthcare system requires shifting how we structure, share, and collaboratively analyze data in clinical and translational science. Data from veterinary patients provide an underutilized resource for accelerating translational science and precision medicine across species. Because people and their companion animals (dogs, cats, and horses) coexist in the same household, understanding the factors driving disease trends in animals may help reveal shared risk factors, disease pathways, and novel interventions. There is growing awareness of the relevancy of cross-species data, the importance of the human-animal bond, and the need for One Health approaches to health problems.[1] One Health is an integrative multidisciplinary effort focused on achieving optimal health for people, animals, and their shared environments; studies showing the shared risk of disease, such as diabetes in dog owners and their pets,[2] underscore the importance of these efforts.

Driven by sedentary lifestyles and diet, companion animals, like humans, face an obesity epidemic. An estimated 25-50% of companion animals are overweight or obese.[3] Body mass index of adult dog owners and body condition score of their dogs were positively correlated in one study where nearly 58% of dog owners and 63% of dogs were overweight or obese highlighting the cross-species prevalence and need for successful prevention strategies.[4] Overweight body condition and obesity are the major causes of metabolic syndrome (MetS), a cluster of phenotypic features that includes abdominal obesity, hypertension, insulin resistance, and dyslipidemia. MetS is associated with increased risk of cardiovascular disease and Type 2 Diabetes (T2D), with an estimated prevalence of 30-34% among the United States adult population.[5] Metabolic-like syndromes (MetS-like) are clusters of phenotypic features centered around obesity and insulin dysregulation and have been described for horses (equine metabolic syndrome - EMS), cats (feline metabolic syndrome - FMS), and dogs (obesity-related metabolic dysfunction - ORMD).[3,6,7] MetS-like phenotypes have severe health consequences ranging from osteoarthritis, cancer, cardiovascular disease, T2D, hyperinsulinemia-associated laminitis, pregnancy complications, and shortened life span.[6–13] Disease definitions and proposed pathophysiology for MetS and MetS-like are similar across species, highlighting their translational relevance (Figure 1).

An Observational Medical Outcomes Partnership (OMOP) pet-patient registry for geographically overlapping human and veterinary medical institutions enables investigation of the co-occurrence of conditions and environmental influences, including MetS in humans and related phenotypes in companion animals, in coincident households. To accurately capture the rich phenotypic features of MetS and MetS-like, we are integrating open, community-driven, cross-species-enabled ontologies (e.g., from the OBO Foundry[14]) into the OMOP common data model (CDM).

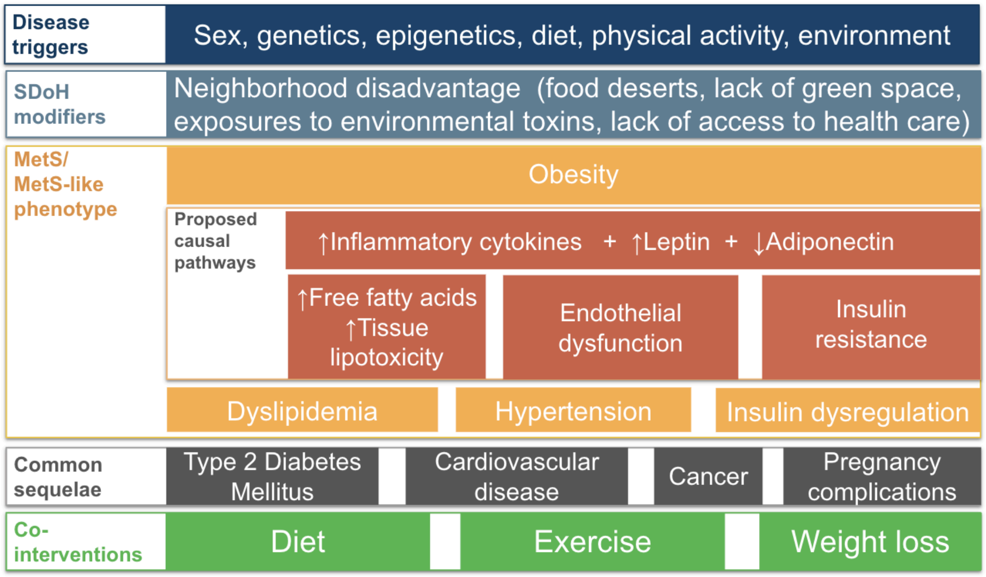

Figure 1: Metabolic syndrome (MetS) in humans and metabolic-like syndromes (MetS-like) in companion animals have similar disease triggers, social determinants of health (SDoH) modifiers, pathophysiology, common sequelae, and potential co-interventions.

Methods

To investigate the feasibility of creating an OMOP pet-patient registry for secure analysis of health data from University of Colorado (UCHealth) patients and their pets who received veterinary care at Colorado State University (CSU-VTH), we matched non-medical personally identifying information (PII) and determined the total number of patients who had a pet visit CSU-VTH.

PII (first name, last name, street address, city, 5-digit zip code, phone number, and email address) from the CSU-VTH records was delivered to the Google Cloud-based Health Data Compass (HDC) research data warehouse (RDW) at CU Anschutz using encrypted file transfer and the HDC HIPAA compliant virtual machine (VM) gateway.[15] CSU-VTH PII was matched to patient records in HDC using a python package hybrid deterministic and probabilistic record linkage method, CU Record Linkage (CURL) (Table 1).[16]

Table 1. Record linkage methods.

Method	Linkage variable
Deterministic 1	First6ofFN + First6ofLN + phone
Deterministic 2	First6ofFN + First6ofLN + email
Deterministic 3	First6ofFN + First6ofLN + street_address
Probabilistic	FN, LN, Street address, Phone ● Blocking scheme 1: FN_Soundex + LN_Soundex ● Blocking scheme 2: First3ofFN + First3ofLN + state

Abbreviations: FN, first name; LN, last name

Linkages identified using deterministic method 1 were removed before deterministic method 2 was executed, and linkages identified using deterministic method 2 were removed before deterministic method 3 was executed. Probabilistic blocking schemes 1 and 2 were sequentially executed after deterministic method 3. The threshold to declare a link was set to 90; therefore, linked pairs with a match score lower than 90 were not considered matches and labeled as "no-matches."[16] The weight redistribution method was applied to the probabilistic linkage method in the presence of missing linkage data.[17] The linkage methods were adopted from existing methods currently used by HDC.

We queried the CSU-VTH veterinary electronic health record (vEHR) system agnostic of the linkage for the frequency of phenotypic features related to MetS-like in companion animals and mapped them to standard vocabularies including SNOMEDCT and SNOMEDCT_VET and the OBO Foundry ontologies.[14] Statistical differences in the prevalence of MetS-like key indicators for each companion animal species were evaluated by the chi-squared test using R.[18]

Results

Data from 41,081 CSU-VTH clients from 2019-2024 were cross-joined with 3,282,860 UCHealth patients from 2015-2024 within the HIPAA-compliant RDW. We identified 12,115 matches, indicating 29% of CSU-VTH clients were UCHealth patients (Figure 2). Fifty-three percent of all UCHealth human patients and 48% of CSU-VTH animal patients were female.

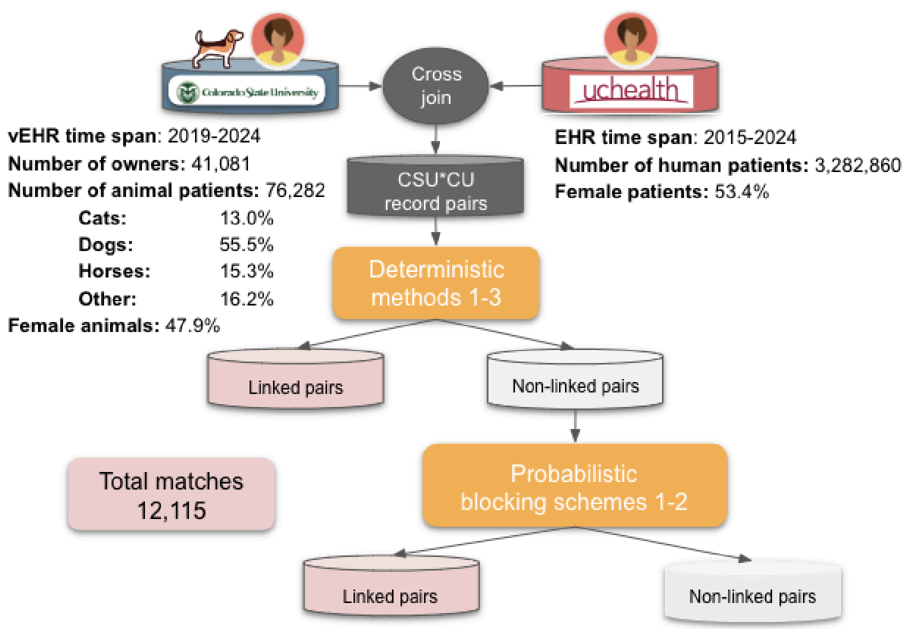

Figure 2: Data linked from geographically adjacent human and veterinary medical academic institutions. The total number of matches (UCHealth patients who took their animal to the CSU-VTH) is 12,115. Approximately 29% of pet owners are UCHealth patients. See Table 1 for details about deterministic and probabilistic record linkage methods.

Query of the CSU-VTH vEHR identified a MetS-like key term in the records of 13,682 (43.9%) of canine, 3,048 (51.0%) of feline, and 1,028 (11.8%) of equine patient records, with significant differences in prevalence among species ($p<0.001$). Medical summary documents (one medical summary document per animal per visit) were queried for phenotypic features and disease sequelae associated with MetS-like are shown in Table 2.

Some of the key terms associated with MetS-like (e.g., over-conditioned, cresty, and cresty neck) were not available in standard vocabularies including SNOMED CT and SNOMEDCT_VET; OBO Foundry[14] ontologies (e.g., the Ontology for Biological Attributes[19]) were used in these cases.

Table 2: Discovery rate and number of animal patient visits from the CSU-VTH with phenotypic features and disease sequelae associated with MetS-like.

Key term	Discovery rate (%)	Number of animal patient visits
Elevated BCS	19.18	33,282
Overweight	3.94	6,830
Diabetes	2.14	3,710
Obesity	1.26	2,186
Over-conditioned*	1.21	2,095
Diabetic	0.80	1,387
Obese	0.56	964
Cresty*	0.38	657
DKA	0.30	520
Insulin resistance	0.24	415
Metabolic syndrome	0.19	325
Equine metabolic syndrome	0.17	293
EMS	0.07	114
Cresty neck*	0.04	75

*Term not present in SNOMED CT or SNOMED Veterinary Extension (SNOMEDCT_VET). BCS, body condition score; EMS, equine metabolic syndrome; DKA, diabetic ketoacidosis.

Conclusion

We have demonstrated the feasibility of developing an OMOP pet-patient registry utilizing EHR data for geographically overlapping human and veterinary medical institutions. We found significant overlap between individuals who were UCHealth patients and had a pet-patient of the CSU-VTH: 29% of CSU-VTH clients were UCHealth patients.

The MetS and MetS-like use-case is constructive to evaluate the utility of a pet-patient registry given the high prevalence of these conditions in humans and companion animals and the need for One Health interventions. Using open, community-driven ontologies within the OMOP pet-patient registry will allow integration of species-specific terms e.g., cresty neck score, which is an independent predictor of insulin dysregulation in ponies.[20]

Next steps include creating a HIPAA-compliant pet-patient registry to explore the dynamics of diseases shared by people and companion animals in coincident households. Top priorities include data governance, stakeholder engagement, and preserving data security and privacy throughout the research cycle.

References

1. Mackenzie JS, Jeggo M. The One Health Approach-Why Is It So Important? Trop Med Infect Dis [Internet]. 2019 May 31;4(2). Available from: http://dx.doi.org/10.3390/tropicalmed4020088

2. Delicano RA, Hammar U, Egenvall A, Westgarth C, Mubanga M, Byberg L, et al. The shared risk of diabetes between dog and cat owners and their pets: register based cohort study. BMJ. 2020 Dec 10;371:m4337.

3. Ragno VM, Zello GA, Klein CD, Montgomery JB. From Table to Stable: A Comparative Review of Selected Aspects of Human and Equine Metabolic Syndrome. J Equine Vet Sci. 2019 Aug;79:131–8.

4. Linder DE, Santiago S, Halbreich ED. Is There a Correlation Between Dog Obesity and Human Obesity? Preliminary Findings of Overweight Status Among Dog Owners and Their Dogs. Front Vet Sci. 2021 Jul 9;8:654617.

5. Kim EG, Kaelber DC. Phenotypic prevalence of obesity and metabolic syndrome among an underdiagnosed and underscreened population of over 50 million children and adults. Front Genet. 2022 Sep 6;13:961116.

6. Tvarijonaviciute A, Ceron JJ, Holden SL, Cuthbertson DJ, Biourge V, Morris PJ, et al. Obesity-related metabolic dysfunction in dogs: a comparison with human metabolic syndrome. BMC Vet Res. 2012 Aug 28;8:147.

7. Okada Y, Kobayashi M, Sawamura M, Arai T. Comparison of Visceral Fat Accumulation and Metabolome Markers among Cats of Varying BCS and Novel Classification of Feline Obesity and Metabolic Syndrome. Front Vet Sci. 2017 Feb 14;4:17.

8. Manfredi JM, Jacob SI, Boger BL, Norton EM. A one-health approach to identifying and mitigating the impact of endocrine disorders on human and equine athletes. Am J Vet Res [Internet]. 2022 Dec 27;84(2). Available from: http://dx.doi.org/10.2460/ajvr.22.11.0194

9. Salt C, Morris PJ, Wilson D, Lund EM, German AJ. Association between life span and body condition in neutered client-owned dogs. J Vet Intern Med. 2019 Jan;33(1):89–99.

10. Montoya M, Morrison JA, Arrignon F, Spofford N, Charles H, Hours MA, et al. Life expectancy tables for dogs and cats derived from clinical data. Front Vet Sci. 2023 Feb 21;10:1082102.

11. Marchi PH, Vendramini THA, Perini MP, Zafalon RVA, Amaral AR, Ochamotto VA, et al. Obesity, inflammation, and cancer in dogs: Review and perspectives. Front Vet Sci. 2022 Oct 3;9:1004122.

12. Thengchaisri N, Theerapun W, Kaewmokul S, Sastravaha A. Abdominal obesity is associated with heart disease in dogs. BMC Vet Res. 2014 Jun 13;10:131.

13. Hallman I, Karikoski N, Kareskoski M. The effects of obesity and insulin dysregulation on mare reproduction, pregnancy, and foal health: a review. Front Vet Sci. 2023 Apr 20;10:1180622.

14. Jackson R, Matentzoglu N, Overton JA, Vita R, Balhoff JP, Buttigieg PL, et al. OBO Foundry in 2021: operationalizing open data principles to evaluate ontologies. Database [Internet]. 2021 Oct 26;2021. Available from: http://dx.doi.org/10.1093/database/baab069

15. Health Data Compass [Internet]. [cited 2024 Jul 31]. Available from: https://www.healthdatacompass.org/

16. Ong TC, Duca LM, Kahn MG, Crume TL. A hybrid approach to record linkage using a combination of deterministic and probabilistic methodology. J Am Med Inform Assoc. 2020 Apr 1;27(4):505–13.

17. Ong TC, Mannino MV, Schilling LM, Kahn MG. Improving record linkage performance in the presence of missing linkage data. J Biomed Inform. 2014 Dec 1;52:43–54.

18. The R Project for Statistical Computing [Internet]. [cited 2024 Oct 2]. Available from: https://www.r-project.org/

19. Stefancsik R, Balhoff JP, Balk MA, Ball RL, Bello SM, Caron AR, et al. The Ontology of Biological Attributes (OBA)-computational traits for the life sciences. Mamm Genome. 2023 Sep;34(3):364–78.

20. Fitzgerald DM, Anderson ST, Sillence MN, de Laat MA. The cresty neck score is an independent predictor of insulin dysregulation in ponies. PLoS One. 2019 Jul 24;14(7):e0220203.