Constructing the first comorbidity networks in companion dogs in the Dog Aging Project

Antoinette Fang[1], Lakshin Kumar[2], Kate E Creevy[3], the Dog Aging Project Consortium, Daniel E.L. Promislow[4], & Jing Ma[5]*

1. Department of Mathematics, University of Chicago, Chicago, Illinois, United States of America
2. Department of Medicine, University of California, San Francisco, California, United States of America
3. College of Veterinary Medicine & Biomedical Sciences, Texas A&M University, College Station, United States of America
4. Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University, Boston, Massachusetts, United States of America
5. Division of Public Health Sciences, Fred Hutchinson Cancer Center, Seattle, Washington, United States of America

* jingma@fredhutch.org

Abstract

Comorbidity and its association with age are of great interest in geroscience. However, there are few model organisms that are well-suited to study comorbidities that will have high relevance to humans. In this light, we turn our attention to the companion dog. The companion dog shares many morbidities with humans. Thus, a better understanding of canine comorbidity relationships could benefit both humans and dogs. We present an analysis of canine comorbidity networks from the Dog Aging Project, a large epidemiological cohort study of companion dogs in the United States. We included owner-reported health conditions that occurred in at least 60 dogs (n=160) and included only dogs that had at least one of those health conditions (n=26,614). We constructed an undirected comorbidity network using a Poisson binomial test, adjusting for age,

sex, sterilization status, breed background (i.e., purebred vs. mixed-breed), and weight. The comorbidity network reveals well-documented comorbidities, such as diabetes with cataracts and blindness, and hypertension with chronic kidney disease (CKD). In addition, this network also supports less well-studied comorbidity relationships, such as proteinuria with anemia. A directed comorbidity network accounting for time of reported condition onset suggests that diabetes precedes cataracts, elbow/hip dysplasia before osteoarthritis, and keratoconjunctivitis sicca before corneal ulcer, which are consistent with the canine literature. Analysis of age-stratified networks reveals that global centrality measures increase with age and are the highest in the Senior group compared to the Young Adult and Mature Adult groups. Only the Senior group identified the association between hypertension and CKD. Our results suggest that comorbidity network analysis is a promising method to enhance clinical knowledge and canine healthcare management.

Author Summary

Companion dogs age alongside humans and suffer many of the same diseases, making them an ideal “real-world” model for human health. Using owner-reported data from 26,614 dogs enrolled in the nationwide Dog Aging Project, we built the first large-scale maps—called comorbidity networks—that show which canine diseases tend to appear together and in what order. The networks correctly highlighted well-known pairings such as diabetes with cataracts and blindness, and hypertension with chronic kidney disease. They also revealed under-appreciated links—for example, protein loss in urine associated with anaemia—suggesting new avenues for veterinary research and care. By adding the reported date of diagnosis, we could infer likely sequences of the diseases: diabetes generally preceded cataracts, hip dysplasia came before osteoarthritis, and dry-eye disease often led to corneal

ulcers. When we split the data by life stage, we saw disease webs become denser and more centred on a few key conditions as dogs grew older, echoing patterns seen in people. Together, these findings show that network analysis of large pet-health datasets can guide clinicians, inform breeding and prevention strategies, and ultimately improve the wellbeing of both dogs and humans.

Introduction

Comorbidity refers to the simultaneous presence of two health conditions in an individual organism (1,2). The understanding of comorbidities can offer insights into health condition associations and progression, which can in turn improve predictive and diagnostic tools in the clinic (3–6). The strong positive association between measures of comorbidity and age has led those in the geroscience community to explore the mechanisms by which comorbidity might be a cause or consequence of aging (7–11). Comorbidities are often studied in the form of a network (12–18), where each node in the network represents a health condition and each edge represents a statistically significant association between a pair of health conditions. Many health conditions are related through a variety of biological mechanisms, and comorbidity networks can help us better understand these underlying connections. Thus far, there are several large studies that use a variety of methods investigating comorbidities and comorbidity networks in humans (14,15,18–21). The same is not true for dogs, for which there is a limited understanding of comorbidities. Notably, dogs are a good model organism because they share the same environments and many of the same health conditions as humans, they have an almost equally sophisticated healthcare system, and show similar manifestations of aging as humans (22). Existing dog comorbidity studies are limited in sample size and diversity (23–26), or focus on comorbidities in relation to an index health condition (26–30). While some previous studies controlled for demographic covariates when examining dog comorbidities (26), many studies

either do not report adjusting for such variables or do not explicitly describe their methods for handling dog demographics and other important covariates (23–25,27–30).

In this study, we address these limitations by constructing a comorbidity network for more than 150 canine morbidities using data from a large cohort (n=26,614) assembled by the Dog Aging Project (DAP). The DAP is a long-term longitudinal study of the genetic, environmental and lifestyle determinants of healthy aging in companion dogs (22). As of November 2024, more than 50,000 participants across the United States have signed up their dogs. Participating dog owners fill out an annual survey that includes extensive information on both owner demographics and dog signalment, the home and external environment, and the dog's physical activity, behavior, diet, medications, and of direct relevance to the present study, health status. Having completed this first survey, the dog becomes a member of the DAP "Pack" cohort, and all participants are invited to continue to provide additional data about their dogs over time. The data we used in this study come from the first year of DAP, and are therefore cross-sectional. Through analysis of owner-reported health data, we identify both well-documented and novel health condition associations, demonstrate age-dependent patterns in health condition co-occurrence, and investigate temporal relationships between health conditions. These findings contribute to veterinary informatics and health policy, expanding opportunities for evidence-based medicine in veterinary practice.

Results

Study cohort

We included health conditions that occurred in at least 60 dogs, and we only included dogs that had at least one of these conditions. The final cohort encompasses 160 unique health

conditions and 26,614 dogs, each with at least one of the 160 conditions. Sex and reproductive status were relatively balanced, with spayed females comprising 12,457 (46.81%) of the sample, and neutered males accounting for 12,196 (45.83%). Intact females (687, 2.58%) and intact males (1,274, 4.79%) were less common, with only 7.37% of dogs in the cohort being reproductively intact overall. The breed distribution was approximately equal, with 13,407 (50.38%) mixed breed dogs and 13,207 (49.62%) purebred dogs. The median age of dogs in the study was 7.75 (interquartile range [IQR]: 4.08-11.00 years), and the median weight was 50.90 lbs (IQR: 25.00-69.0 lbs [11.34-31.30 kg]). The health conditions are grouped into 20 body system categories. The skin category contains 21 health conditions (the highest number), while some other categories (congenital, immune-mediated, and hematopoietic) contain only one health condition (Sheet A in S1 Table). Additionally, 28.4% of dogs have two or fewer conditions, while a substantial 71.6% exhibit multimorbidity with three or more concurrent health conditions (Fig 1B). The most frequently observed condition is fractured bone, followed by dental calculus, extracted teeth, and dog bite (Fig 1C). The difference between each dog's most recent date of reported condition onset and earliest date of reported condition onset was used as a proxy for its medical history span (Fig 1A). Unsurprisingly, dogs' medical history increases as they age, except for those over 21 years old. Knowing the span of each dog's medical history allows us to study directed comorbidities across prespecified time windows.

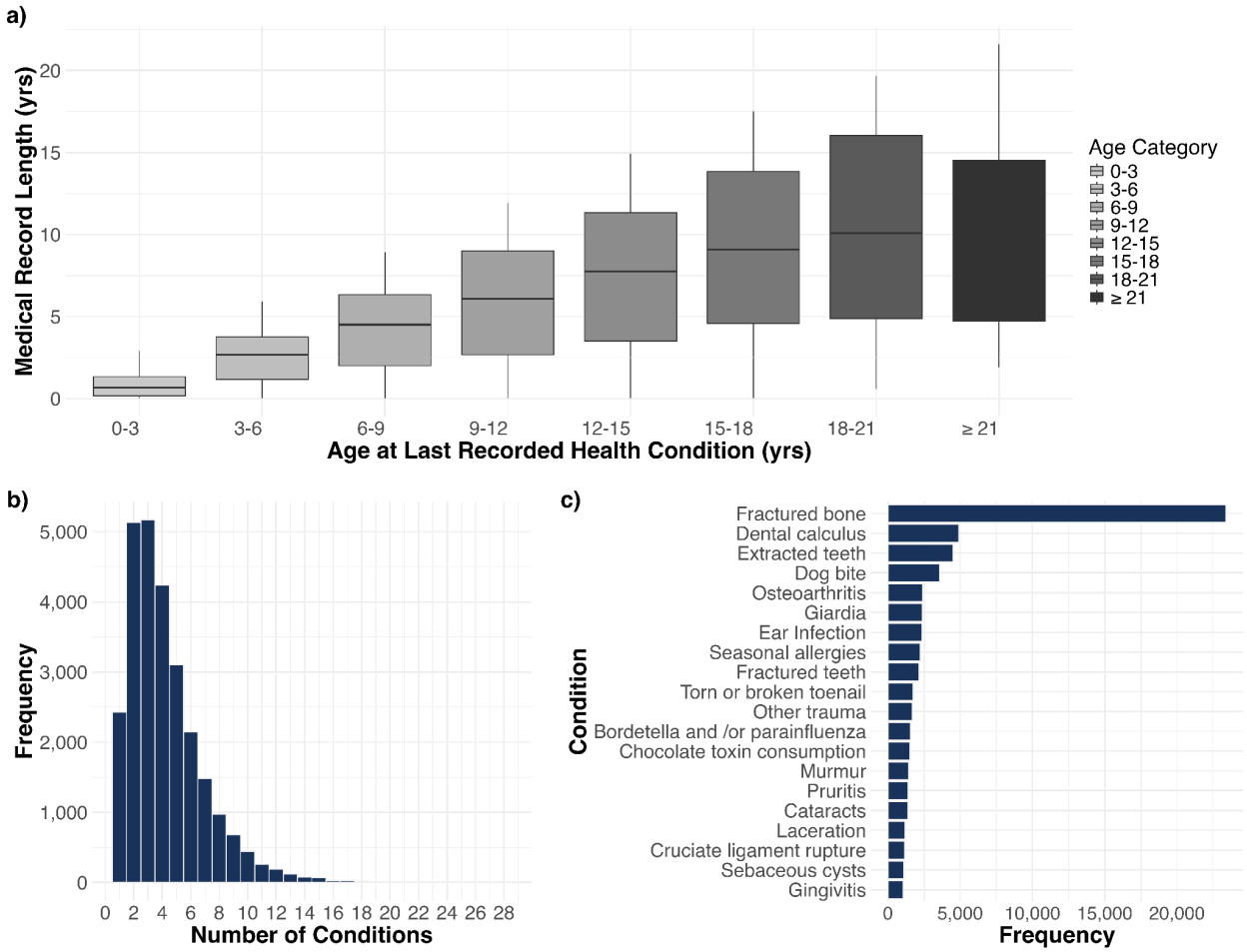

Fig 1. Summary of health condition distribution and medical history characteristics. **(a)** Boxplot comparing each dog's age at last reported health condition with the total span of their medical history (i.e., time between first and last reported condition). **(b)** Histogram representing number of dogs with different numbers of conditions. **(c)** Bar chart displaying the 20 most frequently occurring health conditions in the study population.

Association between demographic factors and health conditions

Over 80% of the models in which the age coefficient was statistically significant ($P < 0.05$) had a positive age coefficient (Fig 2). The age effect is particularly strong for conditions in the

bone/orthopedic, brain/neurologic, cancer, cardiac, and liver/pancreas categories. Conditions that had a significant negative coefficient for age were conditions we would expect to occur in younger dogs, such as retained baby teeth, cryptorchidism, or coccidia (31). Weight (as a measure of body size) showed significant associations with approximately 70% of health conditions, with 58% of these significant associations being positive. All cancer types showed significant positive associations for age ($P < 0.001$) and weight ($P < 0.05$). The sex variable coefficients were not statistically significant for over 90% of the conditions. There were several health conditions for which correlation coefficients differed by sex. However, these health conditions also tend to be those that affect the reproductive organs of male or female dogs (e.g., cryptorchidism or pyometra) (31). The breed background coefficient was only significant for around 47% of the models, and of those, less than 40% were positive (i.e., greater in mixed breed dogs). Purebred dogs showed significantly higher rates of ear infection, cataracts, and intervertebral disk disease (IVDD). In comparison, mixed-breed dogs had significantly higher rates of cruciate ligament rupture, being hit by a car, and many conditions in the infection/parasites category. No significant differences were found between purebred and mixed breed dogs for conditions such as hip dysplasia, patellar luxation, hypoadrenocorticism, hyperadrenocorticism, and mast cell tumors.

	Age (yrs)	Weight (lbs)	Mixed breed	Female, spayed	Male, unneutered	Male, neutered	
Spondylosis	0.29***	0.01***	-4.13e-03	-0.07	0.51	-0.31	
Rheumatoid arthritis	0.33***	0.02***	0.07	14.71	-11.75	14.41	
Patellar luxation	3.81e-03	-0.01***	0.09	1.87***	1.4	2***	
Other orthopedic condition	0.42***	1.51e-03	-0.04	0.39	1.03	0.35	
Osteoarthritis	0.33***	0.02***	-1.89e-03	1.11***	0.31	1.21***	Bone/Orthopedic
Lameness	0.12***	0.01***	0.07	0.55	0.98	0.26	
Hip dysplasia	0.1***	0.02***	-0.03	0.24	-0.31	0.07	
Elbow dysplasia	0.05**	0.01***	-0.29	0.13	0.77	0.31	
Cruciate ligament rupture	0.14***	0.02***	0.17*	1.87***	0.81	1.64*	
Vestibular disease	0.45***	0.01***	0.02	0.1	-0.72	-0.06	
Seizures (including epilepsy)	0.11***	-3.21e-03*	-0.19*	0.31	0.61	0.61	
Other neurologic condition	0.12***	1.04e-03	-0.26	0.11	0.67	0.79	
Laryngeal paralysis	0.45***	0.04***	0.36***	0.04	0.51	0.1	Brain/Neurologic
Intervertebral disk disease (IVDD)	0.19***	-0.02***	-0.51***	13.67	13.9	13.92	
Dementia or senility	0.48***	2.05e-03	-0.28	0.11	-1.27	0.21	
Degenerative myelopathy	0.38***	0.02***	0.2	14.2	-0.92	14	
Unknown cancer	0.25***	0.01***	0.03	0.39	0.92	0.48	
Soft tissue sarcoma	0.27***	0.02***	0.24	12.64	12.07	12.08	
Other cancer	0.23***	0.01*	0.06	-0.28	-0.05	-0.01	
Melanoma	0.24***	0.01***	-0.16	12.89	12.39	12.7	Cancer
Mast cell tumor	0.15***	0.01***	0.01	1.28	0.51	0.84	
Lipoma	0.2***	0.01***	0.31*	12.91	11.51	13.01	
Carcinoma	0.27***	0.01***	0.25	13.45	13.71	13.37	
Valve disease	0.2***	-0.02***	-0.23	0.01	0.11	0.09	
Other Cardiac	0.15***	-6.96e-03	-0.57*	-0.28	0.03	-0.67	
Murmur	0.17***	-0.01***	-0.14*	-0.45	0.05	-0.75	
Hypertension	0.28***	0.01*	0.18	12.79	11.89	12.78	Cardiac
Congestive heart failure	0.29***	-0.03***	-0.11	13.00	13.25	13.35	
Cardiomyopathy	0.12***	5.30e-03	-0.16	0.39	-0.29	0.5	
Arrhythmia	0.17***	4.63e-03	-0.1	-0.28	0.28	0.22	
Rhinitis	0.06	0.01*	0.13	0.4	0.59	0.28	
Other ear condition	0.05*	1.91e-03	-0.10	-0.17	-0.51	-0.11	
Hematoma	0.19***	0.02***	0.08	0.23	0.69	0.12	
Hearing loss	0.46***	-5.53e-03*	-0.23	0.31	-0.2	0.33	Ear/Nose/Throat
Ear Mites	-0.02	-1.16e-03	0.01	0.89	0.44	0.91	
Ear infection	0.04**	0.01***	-0.28***	0.1	-0.16	0.22	
Deafness	0.53***	8.45e-03***	-0.12	0.14	0.13	0.18	
Hypothyroidism (low thyroid function)	0.21***	0.01***	-0.09	0.35	0.67	0.42	
Diabetes mellitus	0.17***	0.01***	-0.28	-0.56	-1.46	-0.25	Endocrine
Cushing's disease	0.3***	-7.40e-03*	0.08	13.45	12.17	13.75	
Addison's disease	0.09**	6.30e-03	0.21	13.41	13.26	13.54	
Uveitis	0.19***	1.25e-03	-0.61*	-0.27	-0.18	-0.29	
Third eyelid prolapse (cherry eye)	-0.02	-7.77e-03**	-0.1	-0.58	-0.4	-0.76	
Progressive retina atrophy	0.18***	-0.98e-03	-0.11	13.13	12.98	13.18	
Other eye condition	0.09***	7.07e-04	0.38***	0.29	0.24	0.23	
Nuclear sclerosis	0.13***	-5.72e-03	0.1	0.79	0.73	0.75	
Keratoconjunctivitis sicca	0.15***	-2.02e-03**	-0.06	0.66	0.5	0.72	
Juvenile cataracts	0.02	5.63e-03	-0.93**	0.2	0.64	-0.06	Eye
Iris cyst	0.18***	0.02***	-0.37	-0.18	-0.75	-1.67	
Glaucoma	0.27***	-1.52e-03*	-0.31	-0.96	-2.11	-0.6	
Entropion (eyelid rolled in)	0.04	0.02***	0.78***	0.5	0.94	0.98	
Corneal ulcer	0.14***	-0.01***	-0.62**	-0.07	0.09	0.22	
Conjunctivitis	-0.02***	-7.02e-04	0.05	-0.05	-0.08	-0.02	
Cataracts	0.38***	-2.58e-03***	-0.26***	0.45	0.48	0.67	
Blindness	0.24***	-0.02***	-0.36**	-0.53	-8.86e-03	-0.3	
Other gastrointestinal condition	2.08e-03	-1.51e-03	-0.14	0.55	0.89	0.55	
Other allergies	3.38e-03	1.65e-03	-0.46*	0.01	-2.05	0.12	
Hemorrhagic gastroenteritis or stress colitis	0.05**	-4.22e-03	-0.08	0.22	0.47	0.94	
Foreign body ingestion or blockage	-0.01	4.65e-03***	-0.07	0.31	0.28	0.5	
Food/medicine allergies	3.66e-03	5.38e-03***	0.17*	0.02	8.65e-03	0.2	
Fecal incontinence	0.23**	0.01**	-0.31	-0.41	-15.16	-0.41	Gastrointestinal
Constipation	0.03	0.01***	0.1	-0.03	0.17	0.11	
Chronic/recurrent vomiting	-5.29e-03	-7.44e-03***	-0.13	0.3	0.1	0.6	
Chronic/recurrent diarrhea	-0.02	2.49e-03	-0.12	-0.39	-0.23	-0.17	
Bloat with torsion (GDV)	0.18**	0.02***	-0.54	0.61	0.41	0.34	
Bilious vomiting syndrome	0.01	-5.47e-03	0.09	0.13	3.74e-03	0.03	
Anal sac impaction	0.09***	-5.85e-03***	-0.09	0.35	0.26	0.28	
Anemia	0.15***	9.08e-04	-0.07	0.08	-0.2	0.15	Hematopoietic
Other Immune	0.04	3.03e-03	-0.45*	1.11	1.76	1.29	Immune-mediated
Whipworms	-0.03	4.29e-03	0.51***	0.14	-0.15	-0.19	
Tapeworms	-5.08e-03	2.81e-03*	0.42***	-0.07	-0.07	-0.02	
Roundworms	0.68***	4.73e-03***	0.22*	0.4	0.15	0.97**	
Parvovirus	-0.02***	2.74e-03	0.89***	-0.01	-0.71	-0.58	
Other infectious disease	-0.02	2.69e-03	0.23**	0.28	0.2	0.46	
MRSA/MRSP	0.1*	0.01**	1.22***	0.31	0.62	0.42	
Lyme disease	0.1***	7.94e-04**	0.19	0.05	0.61	0.11	
Hookworms	0.11***	2.27e-03	0.17*	0.21	0.46	0.16	
Heartworm infection	0.01	4.53e-03*	1.06***	14.41	13.31	14.35	Infection/Parasites
Giardia	-0.11***	5.00e-04	-0.08	0.05	-0.13	0.03	
Gastrointestinal parasites	-0.02***	2.43e-03	0.48**	0.22	0.06	0.2	
Fever of unknown origin	0.02	1.71e-03	-0.03	-0.28	0.18	-0.57	
Ehrlichiosis	0.01	9.45e-03***	0.42**	0.62	0.81	0.73	
Dermatophytosis (ringworm)	-0.03	2.15e-03	0.52***	0.79	0.9	0.98	
Coccidia	-0.13***	8.21e-04	0.34***	0.26	0.1	0.44	
Bordetella/parainfluenza	-0.02**	3.30e-03**	0.2**	0.41	0.13	0.43	
Anaplasmosis	0.07***	5.87e-03**	-1.31e-03	-0.52	-0.04	-0.51	
Urinary tract infection	0.07***	4.63e-03***	0.31***	0.21	0.72	1.35***	
Urinary incontinence	0.13***	0.01***	-0.17	1.21**	0.1	1.37	
Urinary crystals/stones	0.08***	-0.02***	-0.32	0.53	0.07	0.35	
Proteinuria	0.19***	3.58e-03	-0.38	13.25	13.41	12.81	Kidney/Urinary
Other kidney condition	0.05***	-1.13e-03	-0.02	0.42	0.26	-0.09	
Chronic kidney disease	0.33***	-5.71e-03	-0.09	13.81	11.7	13.33	
Acute kidney failure	0.28***	6.32e-03	-0.18	12.64	13.11	12.38	
Pancreatitis	0.15***	-0.01***	-0.24**	0.55	0.16	0.56	
Other liver condition	0.17***	-8.33e-03**	-0.1	-0.09	-0.6	-0.18	Liver/Pancreas
Chronic inflammatory liver disease	0.27***	-6.52e-03	-0.19	12.97	12.55	12.89	
Underbite	-0.06*	-0.02***	0.63***	0.39	-0.03	0.52	
Retained deciduous teeth	-0.04***	-0.01***	-0.32***	0.15	0.45	0.48	
Overbite	-0.04	-0.01***	-0.42	0.03	-0.38	0.14	
Other oral condition	0.07***	-3.73e-03	-0.35*	0.19	0.21	0.04	Mouth/Dental/Oral
Gingivitis	0.15***	0.01***	-0.13	0.73	0.94	0.98	
Fractured teeth	0.11***	2.07e-03*	0.11**	0.2	0.07	0.02	
Extracted teeth	0.19***	-0.02***	-0.19**	0.81***	0.43	0.9***	
Dental calculus	0.15***	-0.02***	-0.04	0.24	0.04	0.32	
Other congenital disorder	9.34e-03	1.39e-03	0.59**	0.19	0.48	0.09	Other congenital disorder
Vaginitis	-0.1***	2.30e-03	-0.41	-0.72*	-18.72	-18.27	
Recessed vulva	-0.1***	1.07***	-0.19	-0.16	-18.97	-18.44	
Pyometra	-0.1**	6.72e-03	1.41***	0.61	-18.98	-18.67	Reproductive
Other reproductive condition	0.06**	1.37e-03	1.08***	0.33	0.1	0.13	
Cryptorchid	-0.14***	-3.02e-03	-1.52***	1	18.08	18.04	
Tracheal stenosis	0.17***	-0.02***	-0.38	13.22	12.88	13.31	
Tracheal collapse	0.18**	-0.02***	0.09	0.11	0.06	0.48	
Pneumonia	0.02	5.60e-03*	0.2	0.73	1.19	1.03	
Other respiratory condition	0.07**	-3.47e-03	-0.19	14.16	13.94	14.01	Respiratory
Elongated soft palate	-0.03	-0.02***	-1.39***	0.17	0.31	0.53	
Chronic/recurrent cough	0.18***	-0.02***	0.04	0.25	0.05	0.73	
Chronic/recurrent bronchitis	0.15***	-8.62e-03	-0.23	-0.48	-1.49	-0.6	
Chronic or recurrent rhinitis	0.13***	-3.40e-03	-0.22	12.98	12.2	13.04	
Umbilical hernia	-0.09***	-5.40e-03*	-0.07*	-0.51	8.21e-03	-0.42	
Ticks	0.04***	3.11e-03*	0.28***	-0.23	0.35	-0.25	
Systemic demodectic mange	-0.18**	8.10e-03	0.75	0.08	-0.93	0.61	
Seborrhea or sebomrhic dermatitis	0.13***	-1.19e-03	-0.51	13.38	13.57	13.32	
Sebaceous cysts	0.19***	4.55e-03***	0.11	0.59	0.01	0.4	
Seasonal allergies	0.01	3.12e-03**	0.06	0.81**	0.45	0.64	
Sarcoptic mange	-7.55e-04	1.21e-04	0.21***	14.19	13.8	14.04	
Pyoderma or bacterial dermatitis	-9.73e-03	7.93e-03***	-0.35**	0.15	0.26	-0.07	
Pruritis	-0.01	-3.14e-04	0.16*	0.49	0.19	0.59	
Pododermatitis	1.80e-03	6.13e-03	0.32	0.08	0.28	0.8	
Other skin condition	0.03*	3.24e-03	0.04	0.26	0.23	0.26	Skin
Non-specific dermatosis	2.00e-03	7.27e-03**	0.17	0.70	0.14	0.03	
Lick granuloma	0.05*	-0.01***	0.1	-0.21	0.7	-0.12	
Food/medicine allergies affecting skin	-0.01	6.93e-03***	-0.05	0.17	-0.15	0.45	
Fleas	0.05***	-4.57e-03***	0.92***	-0.43	-0.12	-0.32	
Flea allergy dermatitis	0.08**	-3.50e-03	0.36**	0.62	0.38	0.5	
Contact dermatitis	-0.43e-03	4.21e-03*	-0.06	0.53	0.03	0.5	
Chronic/recurrent skin infections	0.09***	8.31e-03***	-0.27*	0.03	0.18	0.21	
Chronic/recurrent hot spots	0.05**	-0.01***	-0.12	0.41	0.55	0.83	
Atopic dermatitis	0.02	2.89e-03	-0.01	0.45	0.41	0.54	
Alopecia	0.05***	-2.45e-04	-0.28**	0.11	-0.35	0.01	
Overdose of medications prescribed to the dog	0.05	3.84e-03	0.22	1.16	0.65	0.97	
Other Toxin/Consumption	-0.02	-3.83e-03*	-0.15	-0.28	-0.41	-0.23	
Mouse or rat bait/poison (Warfarin)	0.04	4.45e-03	-0.74	0.67	-0.35	0.64	
Mouse or rat bait/poison (Other/Unknown)	0.07***	3.36e-03	0.09	0.75	0.77	0.56	Toxin consumption
Ingestion of recreational drugs	-0.02	-8.55e-03**	0.09	0.07	-0.02	-0.07	
Ingestion of human medications	-0.02	-2.90e-03	0.17	0.74	0.09	0.65	
Grapes or raisins	-0.02	4.59e-04	0.08	0.46	0.04	0.84	
Chocolate	0.04***	-5.09e-03***	0.74***	0.76*	0.47	0.88**	
Torn or broken toenail	0.04***	0.23e-03***	0.17**	0.43	0.41	0.39*	
Tail injury	0.04**	3.01***	0.14**	-0.21	0.1	-0.38	
Snakebite	0.06*	8.79e-03***	0.32	0.06	0.83	0.11	
Penetrating wound	0.02	1.05***	0.06	-0.2	-0.32	-0.43	
Other trauma	0.03***	2.86e-03**	0.08	0.04	-0.02	0.17	
Laceration	2.69e-03	3.74e-03*	0.27**	0.08	0.13	0.14	
Kicked by horse or other large animal	0.03	5.37e-03	0.3	0.75	0.64	1.00	Trauma
Hit by car or other vehicle	0.05***	2.40e-03	0.62***	-0.03	-0.03	0.05	
Head trauma	0.06**	-1.79e-03	-0.04	-0.44	-0.27	-0.4	
Fractured bone	0.02**	5.18e-03**	0.12**	0.16	0.19	0.15	
Fall from height	0.11***	-4.00e-03	-0.22	-0.51	-0.1	-0.33	
Dog bite	0.05***	4.75e-03***	0.23**	0.08	0.03	0.03	
Bite wound from another animal	0.07***	1.75e-03	0.26*	-0.11	-0.14	-0.05	

color Negative, not significant Negative, significant Positive, not significant Positive, significant

Fig 2. Associations between demographic factors and health conditions, stratified by condition category. Coefficients from logistic regression models are shown for each health condition, with age, weight, breed background, and sex/reproductive status as predictors. Statistical significance is denoted by asterisks (* $p<0.05$, ** $p<0.01$, *** $p<0.001$).

Undirected comorbidity network

The undirected comorbidity network is presented in Fig 3. A complete table of all health condition pairs and their corresponding P-values, regardless of statistical significance, is provided in Sheet A in S2 Table. Of note, health conditions of the same categories tend to be connected with each other. This network highlights a variety of well-known comorbidities: diabetes and blindness ($P < 0.001$); diabetes and cataracts ($P < 0.001$); cataracts and blindness ($P < 0.001$); Cushing's disease and alopecia ($P < 0.001$); and atopic dermatitis and various allergies— specifically skin food/medicine allergies ($P < 0.001$), seasonal allergies ($P < 0.001$), and gastrointestinal food/medicine allergies ($P < 0.001$). All of these pairs include eye or skin health conditions, and these health conditions are commonly used by veterinarians as indicators of systemic diseases like diabetes (31,32). In addition to these skin- and eye-related health condition pairs, we also see some other health condition pairs like hypertension and chronic kidney disease (CKD) ($P < 0.001$); hypertension and proteinuria ($P < 0.001$); and proteinuria and anemia ($P <0.001$). The network also demonstrates the central role of ear infections with a high degree of connectivity (11 connections), linking it to multiple dermatological conditions like atopic dermatitis or chronic and recurrent skin infections; various forms of allergies like food/medicine allergies affecting the skin or seasonal allergies; and other complications like hearing loss and keratoconjunctivitis sicca. Lastly, the network reveals a notable cluster of parasite and parasite-related health conditions, which supports the observation that parasites often co-occur (33–35).

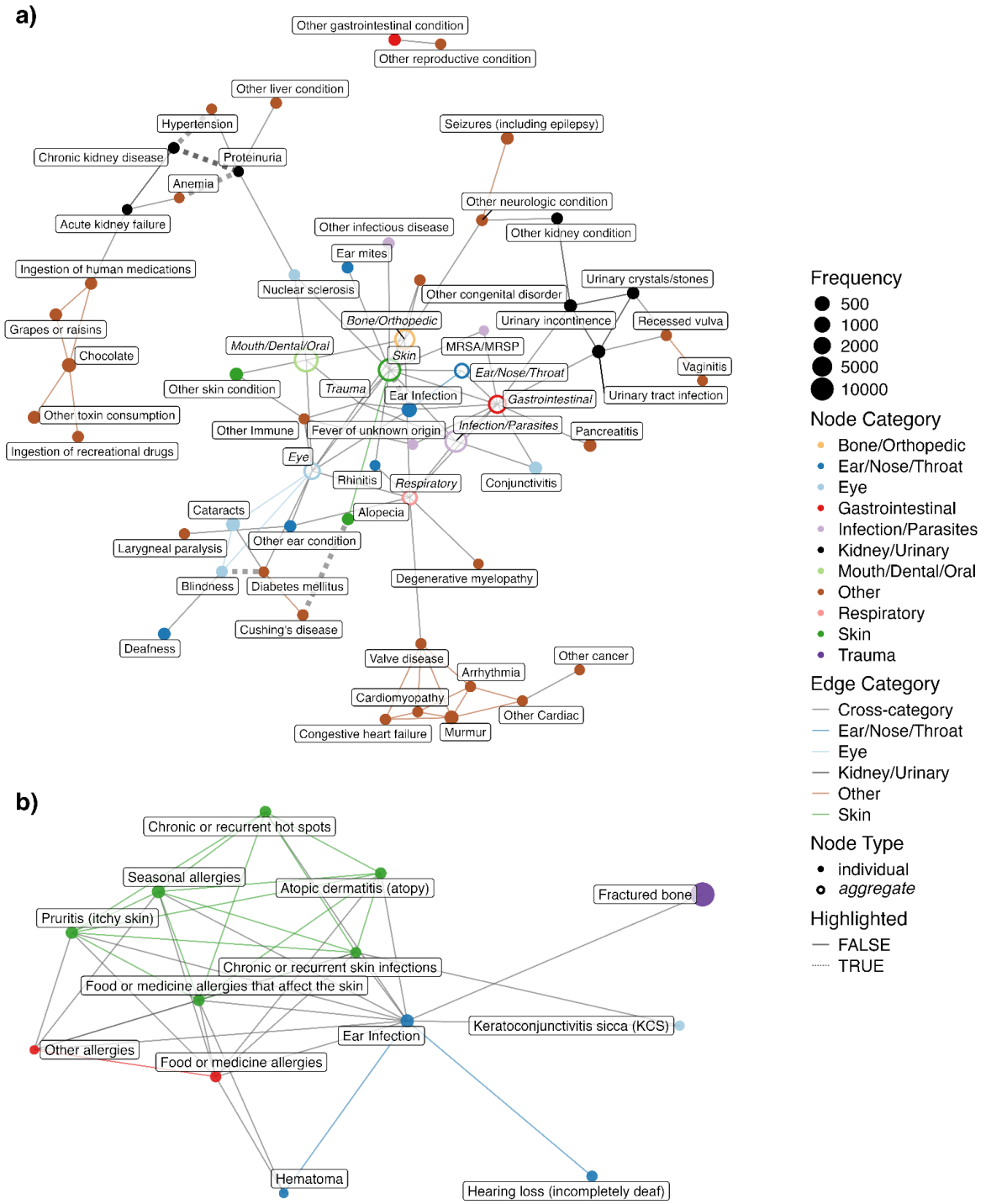

Fig 3. Undirected comorbidity network. Nodes are health conditions with node size proportional to condition prevalence, and edges are statistically significant comorbidity connections (at a Bonferroni adjusted P-value < 0.001). Colors indicate health condition

categories. **(a)** Full network view with Bone/Orthopedic, Ear/Nose/Throat, Eye, Gastrointestinal, Infection/Parasites, Mouth/Dental/Oral, Skin, and Respiratory categories collapsed into single aggregate nodes. **(b)** Zoomed-in view of the Skin conditions, including their first-degree neighbors to ear infection.

Next, we examined the node degree distribution to see if the network can be classified as random or scale-free. Node degree, which is defined as the number of edges connected to each node (36), is one measure of node centrality. The degree distribution is one of the most fundamental properties of networks and provides important information about a network's structure (37). The node degree distribution of our network is better fitted with an exponential distribution than a scale-free power-law distribution ($P < 0.001$) (S1 Fig), providing evidence for a random network.

Comparison to stratified networks

We stratified the dataset into four groups by life stage (Puppy: n=476; Young Adult: n=3931; Mature Adult: n=15895; Senior: n=6155) to investigate the extent to which comorbidity networks varied across life stages (see Methods). As expected, the four networks constructed using the stratified approach differed greatly from each other and also from the unstratified network (Fig 4). Notably, no significant comorbidity patterns were detected for the puppy age group, which comprised less than 2% of the cohort. Puppies had only three conditions meeting the inclusion threshold and a much lower proportion of dogs with two concurrent conditions (28%) compared to other age groups. The young adult, mature adult, and senior networks have, respectively, 19, 97, and 57 conditions. A complete table of all health condition pairs and their corresponding P-values in these 3 networks, regardless of statistical significance, is provided in Sheets B, C, and D in S2 Table. The mature adult network has the most number of edges with a large

connected component (Fig 4B). In all three networks, conditions of the same type tend to be connected with each other, an observation also seen in the unstratified network.

Thirteen comorbidity pairs were identified across all four networks: extracted teeth – fractured teeth, skin food/medicine allergies – gastrointestinal food/medicine allergies, chronic or recurrent diarrhea – chronic or recurrent vomiting, chronic or recurrent diarrhea – gastrointestinal food/medicine allergies, hookworms – roundworms, roundworms – tapeworms, atopic dermatitis – pruritis, pruritis – seasonal allergies, fleas – ticks, skin food/medicine allergies – seasonal allergies, seasonal allergies – gastrointestinal food/medicine allergies, atopic dermatitis – skin food/medicine allergies, and chocolate consumption – grapes/raisins consumption. Conversely, there were 12 unique comorbidity pairs identified by the stratified networks but not by the unstratified network. Eleven of these pairs came from the mature adult group and one from the senior group (Sheets A and B in S3 Table). The pair unique to the senior group is between ticks and gastrointestinal food/medicine allergies. Many pairs unique to the mature adult group are expected, such as elbow dysplasia / lameness / spondylosis – osteoarthritis, murmur / cardiomyopathy – chronic or recurrent cough, and contact dermatitis – flea-allergy dermatitis. However, there is also one pair between other respiratory conditions and other immune conditions, and pairs that are not intuitive, e.g., Conjunctivitis – Giardia and cataracts – deafness. On the other hand, the unstratified network identified 110 comorbidities that were not detected by *any* of the stratified networks. However, this could be due to lower statistical power in the stratified analysis. Among the stratified networks, only the senior network identified the association between hypertension and CKD, while only the mature adult network identified the comorbid pair of blindness and cataracts.

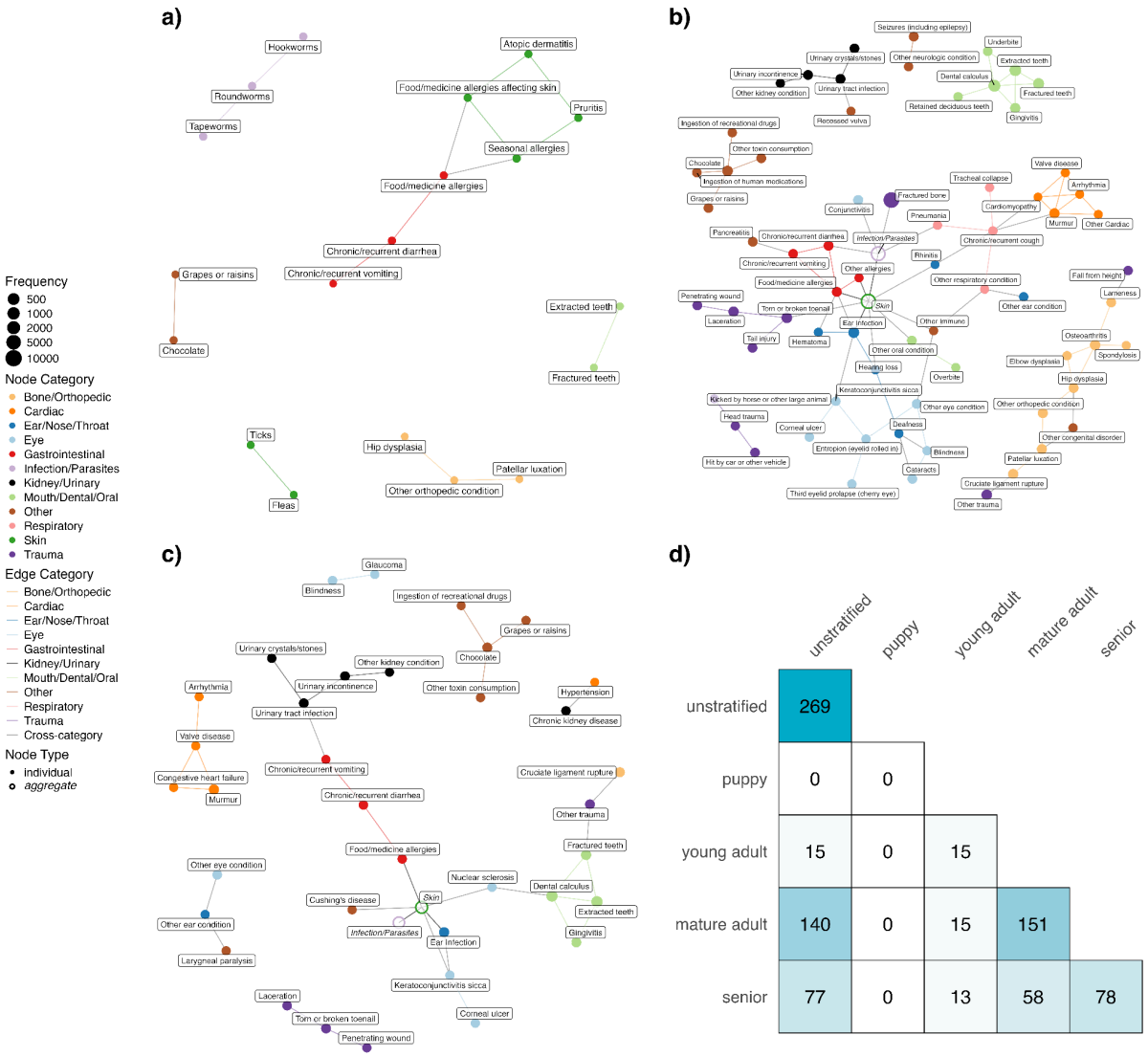

Fig 4. Age-stratified comorbidity networks. Undirected comorbidity networks in Young adult (**a**), Mature adult (**b**) and Senior (**c**). In each network, nodes represent health conditions with node size proportional to condition prevalence, and edges represent statistically significant comorbidity associations (at a Bonferroni-adjusted P-value < 0.01). Colors indicate health condition categories. No network is shown for the Puppy stratum as no significant comorbidities were detected in this age group. Conditions in the Skin and Infection/Parasites categories are presented as aggregate nodes in (b) and (c) to minimize cluttering. (**d**) Heatmap displaying edge overlap between disease networks across age groups. Numbers indicate shared edges between a pair of networks, with diagonal values showing total number of edges per network.

Network topology analysis of the age-specific networks revealed distinct patterns in disease interconnectedness (Table 1). Edge density displayed a different trajectory than other measures,

where it was the highest in the young adult network. By contrast, all other measures– clustering coefficient, betweenness centrality, and degree centrality– increased with age. This suggests a progression from a less structured and more diffuse network in the young adult life stage to an increasingly modular network with a few hub nodes as dogs age. Together, these results support the idea that age is a critical factor influencing comorbidity development.

Table 1. Network topology measures across life stages: edge density refers to the number of edges relative to the maximum number of possible edges in each graph; clustering coefficient measures the degree to which nodes in each graph cluster together; betweenness and degree centrality refer to the normalized graph level centralization measure based on the respective node-level centrality scores**.**

	Puppy	**Young Adult**	**Mature Adult**	**Senior**
Edge Density	N/A	0.088	0.034	0.046
Clustering Coefficient	N/A	0.214	0.491	0.480
Betweenness Centrality	N/A	0.047	0.133	0.149
Degree Centrality	N/A	0.088	0.083	0.113

Adding temporal P-values to understand health condition progression

Finally, we examined temporal relationships between conditions by computing the probability of one condition preceding another within a time window of 12 months and testing its significance using the Poisson binomial approach to comorbidity (see Methods for details). Shorter window spans are more adept at capturing comorbidities that occur sequentially (18), but could result in

power loss in comorbidity discovery since shorter window sizes encompass fewer dogs. Additionally, in shorter time windows, condition development, appearance of relevant signs and symptoms, provider diagnosis, and the recording of said diagnosis may not be accurately, meaningfully captured. The ideal window size should be guided by clinical knowledge and research objectives (18). In our case, we chose a 12-month window to balance capturing meaningful temporal associations while minimizing potential confounding by intervening events or treatments.

A time-directed comorbidity network is presented in Fig 5. A complete table of all health condition pairs and their corresponding P-values in these 3 networks, regardless of statistical significance, is provided in Sheet E in S2 Table. Similar to the undirected network case, conditions of the same category tend to be connected with each other with a dense cluster of skin-related conditions. The directed network suggests that diabetes occurs before cataracts ($P < 0.001$), keratoconjunctivitis sicca (KCS) before corneal ulcer, elbow / hip dysplasia before osteoarthritis, hypertension before CKD, pneumonia before bordetella/parainfluenza, congestive heart failure before valve disease, and blindness before cataracts. A full list of directional pairings supported by existing literature can be found in S4 Table.

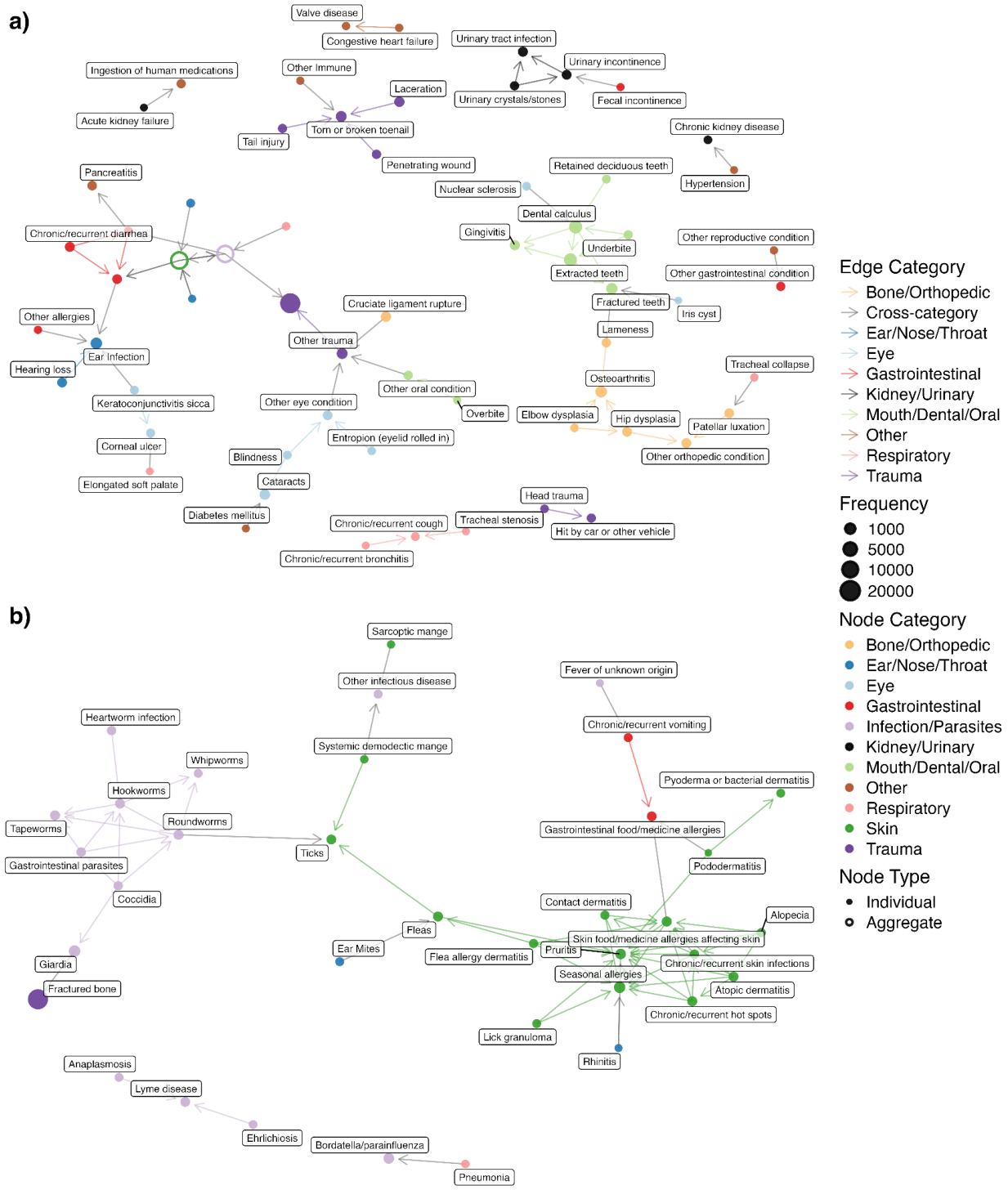

Fig 5. Time-directed network using a window size of 12 months. Nodes are health conditions, and edges are statistically significant comorbidity connections (at a Bonferroni-adjusted P-value < 0.01). Colors indicate health condition categories. Arrowheads point from the health condition that occurs earlier in time to the health condition that occurs later. **(a)** Full network view with Skin and Infection/Parasites categories collapsed into single

aggregate nodes. **(b)** Zoomed-in view of the Skin and Infection/Parasites conditions, including their first-degree neighbors.

Discussion

Aging and comorbidities are linked in a self-reinforcing feedback loop (8). Biological aging leads to increased risk of various diseases, which increases the risk of comorbidities. The presence of these comorbidities contributes to further physiological decline which accelerates biological aging. The dog is a great model to explore this relationship. Through the use of canine comorbidity networks, this study represents an important step towards understanding the relationship between aging and comorbidity. Our work also lays a critical foundation for incorporating future DAP genetic and other molecular data to elucidate the mechanisms underlying the association between aging and comorbidity.

Our logistic regression analyses provide a comprehensive overview of demographic and biological factors associated with the occurrence of various canine health conditions. Our results reaffirm that age is the primary risk factor for many health conditions, while body size and breed background modulate disease risk in condition-specific manners. In our models, 80% of the coefficients in which age is significant are positive, confirming that morbidity rises steeply with older age. Other large epidemiological studies support this pattern (38,39). Weight is significant in about 70% of the conditions, with 57% positive associations. Separate investigation using DAP data agreed that larger dogs had higher lifetime prevalence of cancer and orthopedic disorders but lower odds of ocular, cardiac, hepatic and respiratory disease (40). For cancer specifically, heavier dogs were diagnosed several years earlier than lighter dogs (38). These

external findings corroborate our weight coefficients and clarify that “size risk” is condition-specific rather than universally detrimental. Our models did not consider the interaction between age and body size, which may be included to further refine disease risks (40). The effects of breed background are highly condition-specific and not uniformly disadvantageous to purebreds. Consistent with our findings, a study based on the UK VetCompass cohort showed that ear infections were more common in purebreds (39). A US study based on electronic health records of 27k dogs found purebreds over-represented in cataracts and IVDD, but no difference in hip dysplasia, patellar luxation, hypoadrenocorticism, hyperadrenocorticism, and mast cell tumors, while mixed-breed dogs had excess cruciate-ligament rupture and being hit by a car (41), which are aligned with our discoveries. On the other hand, their results on elbow dysplasia (bone/orthopedic) and hypothyroidism (endocrine), both over-represented in purebreds, are distinct from ours. This could be due to our use of a multivariate logistic regression model, where the effect of breed background is after adjusting for other covariates. We found that mixed-breed dogs had higher rates of infectious and parasitic diseases. This pattern does not appear to stem from poorer husbandry: mixed-breed owners actually reported flea- and tick-control use more often than purebred owners and had comparable rates of heartworm prevention. Although mixed-breed owners were more likely to be in lower-income brackets, income alone therefore does not explain the difference. A plausible alternative is acquisition history: mixed-breed dogs are more often adopted as adults—particularly from shelters—where they may have contracted infections or parasites before joining their current households. Lastly, it is also worth noting that breed background is very imprecise. Future work should incorporate genetic data to fully disentangle the effect of breed on disease risk.

The undirected canine comorbidity network supports clinical observations that ear infections often occur as part of a broader pattern of allergic and inflammatory conditions, particularly those affecting the skin (42). The multiple connections to allergic conditions align with previous findings that hypersensitivity diseases are among the most common primary factors leading to otitis in dogs (42). The connection between Cushing's disease and alopecia is well-known: the disease leads to overproduction of the hormone cortisol, which affects hair growth and can cause the hair to fall out (43). The conditions diabetes, blindness and cataracts are associated with each other in our study, which are aligned with clinical observations that most dogs with diabetes mellitus will acquire cataracts in 1-2 years (44). In addition, the network highlights the association between anemia, proteinuria and CKD. Proteinuria describes the presence of protein in the urine and may be associated with acute or chronic kidney disease, as well as with certain systemic health conditions (40). Anemia, a deficiency of red blood cells, is also often present in advanced CKD due to the combination of impaired red blood cell survival deficiency of production of erythropoietin by the kidney, and impaired bone marrow responsiveness to erythropoietin (45–47). The network illustrates these relationships by connecting anemia to CKD through proteinuria. While our undirected network cannot imply causal relationships or mediation effects, it does highlight potential associations that warrant further investigation. A recent study conducted with a small cohort of 37 dogs suggests that proteinuria may contribute to anemia in dogs with CKD, a relationship that has also been suggested in other model organisms (48–50). The connections revealed by our network demonstrate its potential in identifying less obvious health condition associations and comorbidities. The network's degree distribution is better modeled with an exponential distribution when compared with a scale-free power-law distribution. Had we observed support for the power-law distribution, this could have been interpreted as suggesting that the risk of having an additional morbidity is directly proportional to the number of morbidities a dog already has. By contrast, the exponential distribution is consistent with a model where the probability of developing a morbidity is

independent of the number of morbidities a dog already has. However, we note that since this is a small network, observing a power-law degree distribution was unlikely to begin with (51,52).

Analysis stratified by life stage revealed important age-specific patterns in canine comorbidities. The mature adult network includes more conditions and edges than the young adult and senior networks due to the larger sample sizes, but the senior network is slightly more complex in terms of edge density and centrality measures. Interestingly, only the senior network identified the comorbid pair of hypertension and CKD. Both diseases, as well as the condition categories they belong to (cardiac and kidney/urinary) are of particular interest in aging research.

When comparing the unstratified network to the age-stratified versions, the unstratified network revealed numerous associations that were absent in the stratified analyses—likely due to reduced statistical power from smaller sample sizes. Conversely, the age-stratified networks uncovered several associations that did not appear in the overall network, highlighting patterns that may be specific to particular life stages. The ticks – gastrointestinal food/medicine allergies pair is unique to the senior network. Neither condition is strongly associated with age, yet their comorbidity is only represented in the senior age group. The association between tick bites and alpha-gal syndrome (AGS) is well-documented in humans: the lone star tick is the primary source in transmitting the alpha-gal sugar that triggers AGS—an allergy to red meats and other mammalian products (53). In dogs, tick bites can also induce alpha-gal antibodies, but no clinical signs of food allergy were observed (54). Further investigation is warranted to understand this comorbidity pair. The conjunctivitis – giardia pair unique to the mature adult network is unexpected. Given that both conditions involve mucosal inflammation, it could be that one or more genetic factor(s) make dogs susceptible to both conditions. Similarly, there may be a genetic component underlying the association between cataracts and deafness. For example,

breeds carrying the merle allele (Australian Shepherd, Collie, Shetland Sheepdog, Great Dane, Dachshund, Catahoula Leopard Dog, etc.) have strict breeding recommendations that avoid mating two merles, preventing the high-risk M/M genotype that is associated with sensorineural deafness and ocular anomalies (55). Most of the common comorbidities identified by all networks regardless of stratification involve skin and/or gastrointestinal conditions that tend to have weaker—or even negative—associations with age than do cancer, endocrine, cardiac or brain/neurologic conditions. This suggests that while the unstratified network is valuable for broad discovery, combining it with stratification can provide further insights when focusing on specific conditions or specific groups of health conditions.

Finally, incorporating the time of reported condition onset provides us with some insights into health condition progression, which could be beneficial for canine health management. Our results show that diabetes occurs before cataracts, which is consistent with the literature: it is estimated that around 75% of dogs diagnosed with diabetes will have blinding cataracts within the following two years (44). KCS, also known as dry eye, is a frequent canine ophthalmic disease and can lead to corneal ulcers in dogs (56). Osteoarthritis is a disease involving the entire joint organ and all of its supporting tissues, though it is most commonly marked by the degeneration and dysfunction of articular cartilage (57). Joint dysplasia—typically affecting the hip or elbow—results from abnormal joint development and can advance to osteoarthritis (58). Our network identifies this causal cascade by linking elbow dysplasia to hip dysplasia and finally osteoarthritis. Moreover, our network reveals a directional pairing from hypertension to CKD. This not only reaffirms the association identified in the undirected network, but also corroborates with existing literature which found systemic hypertension causes renal injury to dogs (59). On the other hand, the time-directed network also reveals several counter-intuitive pairings: congestive heart failure → valve disease, pneumonia → bordetella/parainfluenza, and blindness

→ cataracts. One possible explanation for all three cases is reporting lag: owners may be informed of one condition and learn the specific cause later, reversing the apparent time-stamp. In the case of pneumonia → bordetella/parainfluenza, some dogs receive the bordetella/parainfluenza vaccine during hospital recovery, creating a false post-pneumonia "exposure" entry.

This study has several limitations. First, health condition data were based on owner-reported lifetime prevalence collected at a single point, which introduces recall bias. Owners may not accurately recall all of the dog's health conditions or misremember the order in which the conditions occurred. Furthermore, the cross-sectional design is subject to survivor bias: only dogs who have survived to the point of entry into the study are included. This could distort the observed comorbidity patterns, especially for conditions like cancer that are associated with increased mortality, as dogs that died from these conditions prior to recruitment are not represented in our sample. As DAP collects longitudinal data, these issues may be partially addressed, but they remain a limitation of the current analysis. Second, owners may misreport health status information. The owner-reported dates of condition onset do not necessarily represent the pathophysiological course of a health condition, or may themselves be inaccurate. The survey was designed to be user-friendly, but some of the health conditions overlap or are medically imprecise. Also, health condition categories better align with organ systems, so a single pathological process that is implicated in multiple systems could end up being described in the survey as two or more health conditions. We did not analyze the free-text responses provided by participants in the "other, please describe" fields. These free-text fields were used by about 10% respondents and may contain valuable information. Future work will involve the systematic categorization and analysis of these responses. Third, veterinarians may only recommend, or owners may only elect to proceed with, certain diagnostic tests based on prior

knowledge of a dog's current diagnoses, which could introduce confirmation bias into the network. For instance, blood pressure testing is likely to be recommended in dogs with diagnosed kidney disease, due to the known association of hypertension with kidney disease. However, dogs without kidney disease may not be tested, so lack of an owner-reported condition does not imply its true absence. Furthermore, dogs in DAP might be healthier than average if owners are disinclined to sign up a very unhealthy dog for a study on the determinants of healthy aging, similar to the well-known "healthy volunteer" effect in human studies (60–62). Fourth, breed classification (purebred vs. mixed-breed) is also owner-reported. While this binary classification eases statistical analysis, we cannot verify the accuracy of breed designations without genetic testing, which is only available for a subset of dogs in our cohort. Fifth, rare health conditions pose challenges for logistic regression models, increasing the risk of overfitting and potentially misrepresenting the relationship between rare and common health conditions. For example, although allergies are a common cause of pododermatitis (63), this relationship was not detected in our network—likely due to the small number of dogs with pododermatitis (n=65) compared to allergies (n=1015). Finally, our study cohort is not demographically representative. The owners in our study are predominantly white, highly educated, and wealthier than the US population (S5 Table), which may limit the generalizability of our findings. Additionally, we only considered biological factors (age, sex, sterilization status, breed background, and weight) in disease risk prediction. Environmental and social factors can also affect disease risk (64) and should be incorporated in future analyses.

Acknowledging these limitations, this study demonstrates the power of comorbidity discovery from DAP epidemiological surveys, which has great potential to uncover further insights. First, while this report focuses on positively correlated health condition pairs, our analysis also identified several negatively correlated pairs. A systematic investigation of these negative

associations could reveal potential antagonistic relationships between conditions, though careful interpretation would be needed to distinguish true biological relationships from competing risk effects. Second, DAP has access to veterinary electronic medical records (VEMRs) that can be mined using natural language processing and analyzed using the network methods. It would be interesting to compare a network learned from VEMRs with the one inferred from owner-reported data. Furthermore, while our analysis focuses primarily on pairwise disease associations (comorbidities), many dogs in our cohort (71.6%) exhibit multimorbidity with three or more conditions. Future work could extend these network approaches to better understand complex interactions between multiple diseases simultaneously. This direction would align with the growing interest in multimorbidity research in human medicine. Lastly, we have already collected genetic profiles for over 7000 dogs in DAP. Future work could use gene-relatedness learned from these genetic profiles to control for genetic differences among dogs, potentially surpassing the precision of our current binary classification of breed background. The inclusion of genetic profiles would better reflect the subtle but important differences between various dog breeds and help elucidate the molecular mechanism underlying canine comorbidities.

Methods

Dog Aging Project owner survey data

In the Dog Aging Project, data on signalment, husbandry and health status of the dogs in the study are collected from all participants through a comprehensive owner-reported survey; subsets of participants provide other types of information as well. The Health section of the baseline Health and Life Experiences Survey (HLES) captures lifetime prevalence of health conditions, defined as the proportion of individuals in a population who have experienced a particular condition or event at any point in their lives up to the time of the survey or study.

Owners follow survey logic to identify their dogs' health conditions or clinical signs, organized by process (e.g., trauma) or body system (e.g., cardiac). Owners are presented with more than 300 different health conditions or clinical signs, with an option for "other, please describe" in each body system, and also provide the time of onset for each reported condition. This study utilized the 2021 DAP Curated Data Release, which contains all survey responses obtained by 12/31/2021. As such, while the data we analyze are cross-sectional because they were reported at a single point in time, the survey items are constructed to capture lifetime prevalence as defined above.

Of note, the following information from this dataset was used: (1) age, sex, sterilization status, breed background, and weight of each dog, and (2) list of health conditions (out of 365 possible, distinct health conditions, Sheet B in S1 Table) and the date of reported condition onset (if any) for each dog. Note that the list of health conditions here reflects the health condition options presented to dog owners on the survey.

Logistic regression modeling

To estimate the individualized probabilities for each health condition, we fit logistic regression models (LRM). Previous veterinary work suggests that factors such as age, sex, spay/neuter status, weight, and breed background are associated with health condition development in dogs (39,40,45–47,65). Thus, each LRM included the following demographic characteristics as covariates: age, weight, sex, and whether the dog is purebred or mixed-breed. Sex is defined as a four-category variable (intact female, spayed female, intact male, neutered male). Thus, there are in total six predictors in each LRM. We used breed background as opposed to breeds as a covariate because there are a large number of breeds and many health conditions are very rare. Fitting LRMs on rare events data with a large number of covariates can lead to large bias in

estimation and statistical inference (66). Despite being an imprecise summary, breed background has previously been reported to be associated with several health conditions (46). The response variable is binary indicating whether the dog has a given health condition. The LRMs estimate each dog's probability for having a certain health condition $P_{y \in d}$:

$$P_{y \in d} = \frac{1}{1+\exp(-(\beta_0+\beta_1 x_1+\ldots+\beta_6 x_6))}. \qquad \text{(Eq 1)}$$

To minimize the risk of overfitting, we applied the one-in-ten rule of thumb for logistic regression analysis, which suggests studying one predictive variable for every ten events (67). Thus, we only included health conditions that occurred in at least 60 dogs, given the six covariates in our model.

Comorbidity P-values

The null hypothesis is that a pair of health conditions y, z are independently distributed: $H_0 = y \perp z$. A significant P-value indicates evidence that the health condition pair co-occur more often than by chance. Previous canine comorbidity literature largely relies on methods that assume that the probability of seeing health conditions y, z in a dog $P_{\{y,z\} \in d}$ is equal to the product of the population incidence rate for each health condition under the null hypothesis:

$$P_{y \in d} = \frac{C_y}{N}, \; P_{z \in d} = \frac{C_z}{N} \qquad \text{(Eq 2)}$$

$$P_{\{y,z\} \subset d} = \frac{C_y}{N} \times \frac{C_z}{N} \qquad \text{(Eq 3)}$$

where C_y, C_z represents the prevalence of health conditions and N the total number of dogs. A comorbidity P-value can then be calculated under the binomial distribution. However, this method does not account for important covariates that could impact individual disease risks. A standard approach to deal with covariates is to stratify the population into homogeneous subgroups (15,18,19). The downside of this approach is that it greatly decreases statistical power (18). To overcome this issue, we used the Poisson binomial approach to comorbidity (PBC) (18) method to calculate P-values for each health condition pair. We opt for the PBC method as opposed to other notions of health condition associations (16), because PBC does not require sample stratification and has been shown to have higher power in comorbidity discovery (18). The probability that two health conditions occur by chance in a dog is simply the product of the two individualized health condition probabilities given by the LRMs described previously:

$$P_{\{y,z\} \subset d} = P_{y \in d} \times P_{z \in d}. \qquad \text{(Eq 4)}$$

In short, the Poisson binomial distribution is a generalization of the binomial distribution in which every dog has its own probability of having a health condition. Since the cumulative distribution function for the Poisson binomial is computationally unwieldy, we used the normal approximation, which is accurate but also faster (18,68). To determine a P-value approximately,

we estimated the mean and variance under the normal distribution as follows. The mean in this case represents the expected number of dogs with both health conditions:

$$\mu_{y,z} = \sum_{d} P_{\{y,z\}\subset d}. \qquad \text{(Eq 5)}$$

The variance is

$$\sigma^2_{y,z} = \sum_{d} \sigma^2_{P_{\{y,z\}\in d}} = \sum_{d} P_{\{y,z\}\subset d}\left(1 - P_{\{y,z\}\subset d}\right). \qquad \text{(Eq 6)}$$

Initial analyses showed that accounting for the uncertainties in the predicted individual probabilities through the delta method did not meaningfully alter the comorbidity relationships identified. Therefore, we proceed with the variance presented in Eq 6.

Finally, as described previously, we can find the comorbidity P-values using the normal approximation of the Poisson binomial distribution. Health condition pairs with Bonferroni-adjusted $P < 0.001$ were included in the final network. Comorbid pairs were also verified by finding the Pearson correlation coefficient for the individualized health condition probabilities for each pair. Only pairs with positive coefficients were included.

Reported onset dates and directed networks

We extended our analysis to examine temporal relationships between conditions by analyzing the owner-reported dates when each condition first occurred in their dog. For each pair of health conditions y, z, we analyzed the temporal relationship using only dogs reported to have both

conditions, comparing their individual dates of reported condition onset to determine which condition manifested first. Unlike population-level approaches that compare average age of condition onset across all cases, our method examines the sequence of reported conditions within individual dogs' health histories. This tells us whether a health condition is more likely to occur before another health condition. The probability that the reported onset of health condition y occurs before the reported onset of health condition z in a dog that has both conditions y, z is

$$P_{y \to z}^{\{y,z\} \subset d} = \frac{P_{y \in d}}{P_{y \in d} + P_{z \in d}}. \qquad \text{(Eq 7)}$$

We also incorporated a time window size W and the length of the dog's medical history as done in Lemmon *et al*. (18). Now, the probability that y, z occur within W given that the canine already has the two health conditions is:

$$P_{\{y,z\} \subset d \text{ in } W} = \frac{W}{Span}\left(2 - \frac{W}{Span}\right) if\ Span > W\ ,$$

$$P_{\{y,z\} \subset d \text{ in } W} = 1\ if\ otherwise. \qquad \text{(Eq 8)}$$

We combined these two probabilities to get the probability that y occurs within W before z as:

$$P_{y \to z \text{ in } W}^{\{y,z\} \subset d} = P_{y \to z}^{\{y,z\} \subset d} \times P_{\{y,z\} \subset d \text{ in } W}. \qquad \text{(Eq 9)}$$

Finally, as described previously, we can find the direction P-values using the normal approximation of the Poisson binomial distribution. Note that for each health condition pair y, z, we calculated both $P^{\{y,z\}\subset d}_{y\to z\ in\ W}$ and $P^{\{y,z\}\subset d}_{z\to y\ in\ W}$. Health condition pairs with Bonferroni adjusted P < 0.01 were included in the final network.

In our analysis we chose a window of 12 months. Other window sizes were considered, though the results are similar.

Comorbidity network with stratification

We stratified the DAP dogs into four life-stage cohorts (from youngest to oldest): puppy, young adult, mature adult, and senior. These life stages were constructed using DAP veterinary guidelines that incorporate weight and breed in addition to age (S6 Table) (69). Within each stratum, we selected health conditions that at least 60 dogs had and included only dogs with at least one of those selected health conditions, similar to the criteria for the unstratified network. We again used the PBC method described previously to identify potential comorbidity relationships within each age stratum. Health condition pairs with Bonferroni adjusted P < 0.01 were included in the final networks, with one network for each stratum.

Statistical Analysis

All statistical analyses were performed using R version 4.5.0 (2025-04-11) in RStudio. We used the *tidyverse* package for data manipulation and visualization, and *igraph* for network visualization. Additional specialized packages included *fastDummies* for creating dummy

variables, *furrr* for parallel processing, and *poweRlaw* for degree distribution analysis. Bonferroni correction was applied for multiple testing throughout our analyses.

Acknowledgements

The Dog Aging Project thanks study participants, their dogs, and community veterinarians for their important contributions.

Ethics Statement

The University of Washington IRB deemed that recruitment of dog owners for the Dog Aging Project, and the administration and content of the DAP Health and Life Experience Survey (HLES), are human subjects research that qualifies for Category 2 exempt status (IRB ID no. 5988, effective 10/30/2018). No interactions between researchers and privately owned dogs occurred; therefore, IACUC oversight was not required. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Data Availability

This study uses data in the Health and Life Experiences Survey (HLES) from the 2021 DAP Curated Data Release. These data are housed on the Terra platform at the Broad Institute of MIT and Harvard. Access to these data can be applied for free at https://dogagingproject.org/data-access.

Code Availability

The code for reproducing all analyses and visualizations described in this study can be found at https://doi.org/10.5281/zenodo.15833083 .

References

1. Valderas JM, Starfield B, Sibbald B, Salisbury C, Roland M. Defining Comorbidity: Implications for Understanding Health and Health Services. Ann Fam Med. 2009 Jul 1;7(4):357–63.

2. Feinstein AR. The pre-therapeutic classification of co-morbidity in chronic disease. J Chronic Dis. 1970 Dec;23(7):455–68.

3. Lone NI, Lee R, Salisbury L, Donaghy E, Ramsay P, Rattray J, et al. Predicting risk of unplanned hospital readmission in survivors of critical illness: a population-level cohort study. Thorax. 2019 Nov;74(11):1046–54.

4. Wang H, Cui Z, Chen Y, Avidan M, Abdallah AB, Kronzer A. Predicting Hospital Readmission via Cost-Sensitive Deep Learning. IEEE/ACM Trans Comput Biol Bioinform. 2018 Nov 1;15(6):1968–78.

5. Facchinetti G, D'Angelo D, Piredda M, Petitti T, Matarese M, Oliveti A, et al. Continuity of care interventions for preventing hospital readmission of older people with chronic diseases: A meta-analysis. Int J Nurs Stud. 2020 Jan;101:103396.

6. Atashi A, Sarbaz M, Marashi S, Hajialiasgari F, Eslami S. Intensive Care Decision Making: Using Prognostic Models for Resource Allocation. Stud Health Technol Inform. 2018;251:145–8.

7. Emmerton D, Khan S, Conway J, Mosby D, Abdelhafiz AH. Ageing, Comorbidity and Frailty-Synergistic Risk Factors for Covid-19 Adverse Outcomes. OBM Geriatr. 2020 Jul;4(3):1–12.

8. Divo MJ, Martinez CH, Mannino DM. Ageing and the epidemiology of multimorbidity. Eur Respir J. 2014 Oct;44(4):1055–68.

9. Yancik R, Ershler W, Satariano W, Hazzard W, Cohen HJ, Ferrucci L. Report of the National Institute on Aging Task Force on Comorbidity. J Gerontol Ser A. 2007 Mar 1;62(3):275–80.

10. Stenholm S, Westerlund H, Head J, Hyde M, Kawachi I, Pentti J, et al. Comorbidity and Functional Trajectories From Midlife to Old Age: The Health and Retirement Study. J Gerontol Ser A. 2015 Mar 1;70(3):332–8.

11. Piccirillo JF, Vlahiotis A, Barrett LB, Flood KL, Spitznagel EL, Steyerberg EW. The changing

prevalence of comorbidity across the age spectrum. Crit Rev Oncol Hematol. 2008 Aug 1;67(2):124–32.

12. Folino F, Pizzuti C, Ventura M. A Comorbidity Network Approach to Predict Disease Risk. In: Khuri S, Lhotská L, Pisanti N, editors. Information Technology in Bio- and Medical Informatics, ITBAM 2010. Berlin, Heidelberg: Springer; 2010. p. 102–9.

13. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabási AL. The human disease network. Proc Natl Acad Sci U S A. 2007 May 22;104(21):8685–90.

14. Hidalgo CA, Blumm N, Barabási AL, Christakis NA. A Dynamic Network Approach for the Study of Human Phenotypes. PLOS Comput Biol. 2009 Apr 10;5(4):e1000353.

15. Roque FS, Jensen PB, Schmock H, Dalgaard M, Andreatta M, Hansen T, et al. Using Electronic Patient Records to Discover Disease Correlations and Stratify Patient Cohorts. PLOS Comput Biol. 2011 Aug 25;7(8):e1002141.

16. Fotouhi B, Momeni N, Riolo MA, Buckeridge DL. Statistical methods for constructing disease comorbidity networks from longitudinal inpatient data. Appl Netw Sci. 2018;3(1):46.

17. Moni MA, Xu H, Liò P. CytoCom: a Cytoscape app to visualize, query and analyse disease comorbidity networks. Bioinformatics. 2015 Mar 15;31(6):969–71.

18. Lemmon G, Wesolowski S, Henrie A, Tristani-Firouzi M, Yandell M. A Poisson binomial-based statistical testing framework for comorbidity discovery across electronic health record datasets. Nat Comput Sci. 2021 Oct;1(10):694–702.

19. Siggaard T, Reguant R, Jørgensen IF, Haue AD, Lademann M, Aguayo-Orozco A, et al. Disease trajectory browser for exploring temporal, population-wide disease progression patterns in 7.2 million Danish patients. Nat Commun. 2020 Oct 2;11(1):4952.

20. Guo M, Yu Y, Wen T, Zhang X, Liu B, Zhang J, et al. Analysis of disease comorbidity patterns in a large-scale China population. BMC Med Genomics. 2019 Dec 12;12(12):177.

21. Zhao B, Huepenbecker S, Zhu G, Rajan SS, Fujimoto K, Luo X. Comorbidity network analysis using graphical models for electronic health records. Front Big Data [Internet]. 2023 Aug 17 [cited 2024 Apr 18];6. Available from: https://www.frontiersin.org/articles/10.3389/fdata.2023.846202

22. Creevy KE, Akey JM, Kaeberlein M, Promislow DEL, Dog Aging Project Consortium. An open science study of ageing in companion dogs. Nature. 2022 Feb;602(7895):51–7.

23. Salonen M, Sulkama S, Mikkola S, Puurunen J, Hakanen E, Tiira K, et al. Prevalence, comorbidity, and breed differences in canine anxiety in 13,700 Finnish pet dogs. Sci Rep. 2020 Mar 5;10(1):2962.

24. Bamberger M, Houpt KA. Signalment factors, comorbidity, and trends in behavior diagnoses in dogs: 1,644 cases (1991-2001). J Am Vet Med Assoc. 2006 Nov 15;229(10):1591–601.

25. Dinwoodie I, Dwyer B, Zottola V, Gleason D, Dodman N. Demographics and Comorbidity of Behavior Problems in Dogs. J Vet Behav. 2019 May 1;32.

26. Hoffman JM, Lourenço BN, Promislow DEL, Creevy KE. Canine hyperadrenocorticism associations with signalment, selected comorbidities and mortality within North American veterinary teaching hospitals. J Small Anim Pract. 2018 Nov;59(11):681–90.

27. Packer RMA, Volk HA. Epilepsy beyond seizures: a review of the impact of epilepsy and its comorbidities on health-related quality of life in dogs. Vet Rec. 2015 Sep 26;177(12):306–15.

28. Prudic RA, Saba CF, Lourenço BN, Bugbee AC. Prevalence of proteinuria in a canine oncology population. J Small Anim Pract. 2018 Apr 2;

29. Leal-Lima A, Ferreira TC, Ferreira TMV, Coelho PC, Nunes-Pinheiro DCS. Successful treatment and management of canine ehrlichiosis-leishmaniosis-heartworm comorbidity. Acta Sci Vet Impr. 2021;Pub. 687-Pub. 687.

30. Bookbinder LC, Flanders J, Bookbinder PF, Harvey HJ, Barry JS, Cheetham J. Idiopathic Canine Laryngeal Paralysis as One Sign of a Diffuse Polyneuropathy: An Observational Study of 90 Cases (2007-2013). Vet Surg VS. 2016 Feb;45(2):254–60.

31. Aiello SE, Moses MA, Allen DG, editors. The Merck veterinary manual. Eleventh edition. Kenilworth, NJ, USA: Merck & Co., Inc; 2016. 3325 p.

32. Englar RE. Common clinical presentations in dogs and cats. Hoboken: Wiley Blackwell; 2019.

33. Kajero OT, Janoušková E, Bakare EA, Belizario V, Divina B, Alonte AJ, et al. Co-infection of intestinal helminths in humans and animals in the Philippines. Trans R Soc Trop Med Hyg. 2022 Feb 16;116(8):727–35.

34. Vaumourin E, Vourc'h G, Gasqui P, Vayssier-Taussat M. The importance of multiparasitism: examining the consequences of co-infections for human and animal health. Parasit Vectors. 2015 Oct 20;8:545.

35. Donohue RE, Cross ZK, Michael E. The extent, nature, and pathogenic consequences of helminth polyparasitism in humans: A meta-analysis. PLoS Negl Trop Dis. 2019 Jun;13(6):e0007455.

36. Freeman LC. Centrality in social networks conceptual clarification. Soc Netw. 1978 Jan 1;1(3):215–39.

37. Newman MEJ. Networks. Second edition. Oxford: Oxford University Press; 2018. 1 p.

38. Rafalko JM, Kruglyak KM, McCleary-Wheeler AL, Goyal V, Phelps-Dunn A, Wong LK, et al. Age at cancer diagnosis by breed, weight, sex, and cancer type in a cohort of more than 3,000 dogs: Determining the optimal age to initiate cancer screening in canine patients. PLOS ONE. 2023 Feb 1;18(2):e0280795.

39. O Neill DG, Church DB, McGreevy PD, Thomson PC, Brodbelt DC. Prevalence of disorders recorded in dogs attending primary-care veterinary practices in England. PloS One. 2014;9(3):e90501.

40. Nam Y, White M, Karlsson EK, Creevy KE, Promislow DEL, McClelland RL. Dog size and patterns of disease history across the canine age spectrum: Results from the Dog Aging Project. PLOS ONE. 2024 Jan 17;19(1):e0295840.

41. Bellumori TP, Famula TR, Bannasch DL, Belanger JM, Oberbauer AM. Prevalence of inherited disorders among mixed-breed and purebred dogs: 27,254 cases (1995–2010). 2013 Jun 1 [cited 2024 Nov 22]; Available from: https://avmajournals.avma.org/view/journals/javma/242/11/javma.242.11.1549.xml

42. Bajwa J. Canine otitis externa — Treatment and complications. Can Vet J. 2019 Jan;60(1):97–9.

43. de Bruin C, Meij BP, Kooistra HS, Hanson JM, Lamberts SWJ, Hofland LJ. Cushing's Disease in Dogs and Humans. Horm Res. 2009 Jan 21;71(Suppl. 1):140–3.

44. Beam S, Correa MT, Davidson MG. A retrospective-cohort study on the development of cataracts in dogs with diabetes mellitus: 200 cases. Vet Ophthalmol. 1999;2(3):169–72.

45. 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.

46. Forsyth KK, McCoy BM, Schmid SM, Promislow DEL, Snyder-Mackler N, the DAP Consortium, et al. Lifetime prevalence of owner-reported medical conditions in the 25 most common dog breeds in the Dog Aging Project pack. Front Vet Sci [Internet]. 2023 Nov 3 [cited 2024 Jun 25];10. Available from: https://www.frontiersin.org/journals/veterinary-science/articles/10.3389/fvets.2023.1140417/full

47. O'Neill DG, James H, Brodbelt DC, Church DB, Pegram C. Prevalence of commonly diagnosed disorders in UK dogs under primary veterinary care: results and applications. BMC Vet Res. 2021 Feb 17;17(1):69.

48. Drüeke TB, Massy ZA. Role of proteinuria in the anemia of chronic kidney disease. Kidney Int. 2021 Dec;100(6):1160–2.

49. Borin-Crivellenti S, Crivellenti LZ, Gilor C, Gilor S, Silva DG, Maia SR, et al. Anemia in canine chronic kidney disease is multifactorial and associated with decreased erythroid precursor cells, gastrointestinal bleeding, and systemic inflammation. Am J Vet Res. 2023 Aug 21;1–6.

50. Bissinger R, Nemkov T, D'Alessandro A, Grau M, Dietz T, Bohnert BN, et al. Proteinuric chronic kidney disease is associated with altered red blood cell lifespan, deformability and metabolism. Kidney Int. 2021 Dec;100(6):1227–39.

51. Brunson JC, Agresta TP, Laubenbacher RC. Sensitivity of comorbidity network analysis. JAMIA Open. 2020 Apr 1;3(1):94–103.

52. Broido AD, Clauset A. Scale-free networks are rare. Nat Commun. 2019 Mar 4;10(1):1017.

53. Mitchell CL, Lin FC, Vaughn M, Apperson CS, Meshnick SR, Commins SP. Association between lone star tick bites and increased alpha-gal sensitization: evidence from a

prospective cohort of outdoor workers. Parasit Vectors. 2020 Sep 14;13(1):470.

54. Hodžić A, Mateos-Hernández L, Leschnik M, Alberdi P, Rego ROM, Contreras M, et al. Tick Bites Induce Anti-α-Gal Antibodies in Dogs. Vaccines. 2019 Sep 15;7(3):114.

55. Double Merle Eyes: Understanding the Risks and Health Effects in Dogs - Petful [Internet]. 2024 [cited 2025 Jun 6]. Available from: https://www.petful.com/pet-health/double-merle-eyes-health-risks/

56. Dodi PL. Immune-mediated keratoconjunctivitis sicca in dogs: current perspectives on management. Vet Med Res Rep. 2015 Oct 30;6:341–7.

57. Brandt KD. Response of joint structures to inactivity and to reloading after immobilization. Arthritis Rheum. 2003 Apr 15;49(2):267–71.

58. Meeson RL, Todhunter RJ, Blunn G, Nuki G, Pitsillides AA. Spontaneous dog osteoarthritis — a One Medicine vision. Nat Rev Rheumatol. 2019 May;15(5):273–87.

59. Finco DR. Association of systemic hypertension with renal injury in dogs with induced renal failure. J Vet Intern Med. 2004;18(3):289–94.

60. Pinsky PF, Miller A, Kramer BS, Church T, Reding D, Prorok P, et al. Evidence of a healthy volunteer effect in the prostate, lung, colorectal, and ovarian cancer screening trial. Am J Epidemiol. 2007 Apr 15;165(8):874–81.

61. Lindsted KD, Fraser GE, Steinkohl M, Beeson WL. Healthy volunteer effect in a cohort study: temporal resolution in the Adventist Health Study. J Clin Epidemiol. 1996 Jul;49(7):783–90.

62. Ederer F, Church TR, Mandel JS. Sample sizes for prevention trials have been too small. Am J Epidemiol. 1993 Apr 1;137(7):787–96.

63. Bajwa J. Canine pododermatitis. Can Vet J. 2016 Sep;57(9):991–3.

64. McCoy BM, Brassington L, Jin K, Dolby GA, Shrager S, Collins D, et al. Social determinants of health and disease in companion dogs: a cohort study from the Dog Aging Project. Evol Med Public Health. 2023 Jan 1;11(1):187–201.

65. Hedhammar ÅA, Malm S, Bonnett B. International and collaborative strategies to enhance genetic health in purebred dogs. Vet J Lond Engl 1997. 2011 Aug;189(2):189–96.

66. King G, Zeng L. Logistic Regression in Rare Events Data. Polit Anal. 2001 Jan;9(2):137–63.

67. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996 Dec 1;49(12):1373–9.

68. Hong Y. On computing the distribution function for the Poisson binomial distribution. Comput Stat Data Anal. 2013 Mar 1;59:41–51.

69. Pearson EB, Hoffman JM, Melvin RL, McNulty KE, Creevy KE, Ruple A. Analysis of 2,570 responses to Dog Aging Project End of Life Survey demonstrates that euthanasia is

associated with cause of death but not age. 2024 Feb 1 [cited 2024 Dec 6]; Available from: https://avmajournals.avma.org/view/journals/javma/262/2/javma.23.07.0366.xml

Supporting Information

S1 Table. Health condition classifications and frequencies. (A) Distribution of unique health conditions across body system categories, showing the number of distinct conditions within each category. (B) Detailed listing of all health conditions included in the study, standardized condition names, body system classifications, and frequency of occurrence in the study cohort.

S2 Table. Health condition pair associations. (A) Complete results from the unstratified, undirected comorbidity network analysis. (B-D) Results from age-stratified analyses for Young Adult, Mature Adult, and Senior dogs respectively. There is no table for the puppy cohort as no comorbid pairs were identified. (E) Results from directed network analysis showing temporal relationships between conditions occurring within a 12-month window. For all analyses, health condition pairs are shown with their names, corresponding body system classifications, test statistic, expected values and variance under the null, and p-values (unadjusted and adjusted).

S3 Table. Health condition pairs uniquely identified in age-stratified networks. Health condition pairs that were significant in Mature Adult (A) and Senior (B) but not in any other age-stratified network. The "In Unstratified Network" column indicates whether the pair was also detected in the unstratified network analysis. Each pair is shown with condition names, numerical codes, and the age stratum where the unique association was detected.

S4 Table. Time-directed associations supported by previous literature. Significant health condition pairs identified in our network analysis that are supported by existing veterinary literature. Each pair is shown with condition codes, names, unadjusted and Bonferroni-adjusted P-values from our analysis, and corresponding literature citations.

S5 Table. Demographic characteristics of dog owners in study cohort compared to US population. Racial distribution, educational attainment, and household income of dog owners enrolled in the study compared to 2021 US Census data. Note that owners could select multiple

racial categories, so percentages sum to more than 100%. Income data from owners was collected in categorical ranges.

S6 Table. Life stage definitions across age ranges and weight classes. The table presents the age ranges (in years) used to categorize dogs into life stages (puppy, young adult, mature adult, and senior) based on their weight class. These life stage definitions were adapted from the Dog Aging Project veterinary guidelines (69).

S1 Fig. Log scale density distribution of node degrees in the undirected comorbidity network. The overlaid lines represent the density distribution predicted by an exponential and scale-free power-law model.