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
2. Department of Medicine, University of California, San Francisco
3. College of Veterinary Medicine & Biomedical Sciences, Texas A&M University
4. Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University
5. Division of Public Health Sciences, Fred Hutchinson Cancer Center

* Corresponding author
Email: 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=166) and included only dogs that had at least one of those health conditions (n=26,523). 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 blindness and hypertension with chronic kidney disease. 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 occurs before cataracts, which is 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. Our results suggest that comorbidity network analysis is a promising method to enhance clinical knowledge and canine healthcare management.

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=26,259) assembled by the Dog Aging Project (DAP) [22]. The DAP is a long-term longitudinal study of the genetic, environmental and lifestyle determinants of healthy aging in companion dogs. 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. 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 166 unique health conditions and 26,259 dogs, each with at least one of the 166 conditions (Table 1). The cohort has an approximately equal proportion of male and female dogs, but only 7% of dogs are reproductively intact (unspayed/unneutered). Also, there are roughly equal numbers of purebred and mixed-breed dogs. Dogs in this cohort have a median age of 7.8 years and median weight of 50.8 lbs. 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 (Table S1a). Additionally, most dogs have three or fewer health conditions, with just 37% of dogs having four or more health conditions (Fig. S1).

Table 1: Demographic characteristics of the companion dogs included in the comorbidity network analysis.

Characteristic	**N = 26,253**[1]
Sex	
Female, spayed	12, 416 (46%)
Female, unspayed	679 (3%)
Male, neutered	12, 164 (46%)

Male, unneutered	1, 264 (5%)
Breed	
Mixed breed	13, 374 (50%)
Purebred	13, 149 (50%)
Age (years)	7.8 [4.1, 11.0]
Weight (lbs)	50.8 [25.0, 69.0]

[1]n (%); Median [Q1, Q3]

Association between demographic factors and health conditions

The coefficients in the logistic regression models provide insights into factors associated with different health conditions that are intuitive and consistent with previous findings. For example, over 80% of the models in which the age coefficient was statistically significant ($P < 0.05$) had a positive age coefficient (Table S2). This supports the notion that older age is an important risk factor for many health conditions [31]. 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 [32]. Weight (as a measure of body size) showed significant associations with approximately 68% of health conditions, with 57% of these significant associations being positive. This supports previous research showing that some conditions are more prevalent in larger dogs, while others occur more frequently in smaller dogs [33,34]. The sex variable coefficients were not statistically significant for over 90% of the conditions [34]. 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.e.g., cryptorchidism or pyometra) [32]. The breed background coefficient was only significant for around half of the models, and of those, less than 40% were positive (i.e., greater in purebred dogs). Purebred dogs showed significantly higher rates of ear infection, cataracts, and intervertebral disk disease. In comparison, mixed-breed dogs had significantly higher rates of cruciate ligament rupture, being hit by a car, pruritis, and hookworms. 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, aligning with findings from previous work [35,36].

Undirected comorbidity network

The undirected comorbidity network is presented in Fig. 1. A complete table of all health condition pairs and their corresponding P-values, regardless of statistical significance, is provided in Table S3a. Of note, health conditions of the same categories tend to be connected with each other. This network models a variety of well-documented comorbidities: diabetes (node 1307) and blindness (node 101) ($P < 0.001$); atopic dermatitis (node 405) and allergies (nodes 412) ($P < 0.001$); and Cushing's disease (node 1305) and alopecia (node 404) ($P < 0.001$) [32]. 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 [32,37]. In addition to these skin- and eye-related health condition pairs, we also see some other well-known health condition pairs like hypertension (node 514) and chronic kidney disease (CKD, node 908) ($P < 0.001$) [38,39]. The network demonstrates the central role of ear infections (node 202) with a high degree of connectivity (11 connections), linking it to multiple dermatological conditions like atopic dermatitis or chronic and recurrent skin infections (nodes 405 and 407); various forms of allergies like food/medicine allergies affecting the skin or

seasonal allergies (nodes 412 or 426); and other complications like hearing loss and keratoconjunctivitis sicca (nodes 205 and 104). This pattern of connectivity quantitatively supports clinical observations that ear infections often occur as part of a broader pattern of allergic and inflammatory conditions, particularly those affecting the skin [40]. 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 [40]. The network also reveals a notable cluster of parasite and parasite-related health conditions, which supports the observation that parasites often co-occur [41–43].

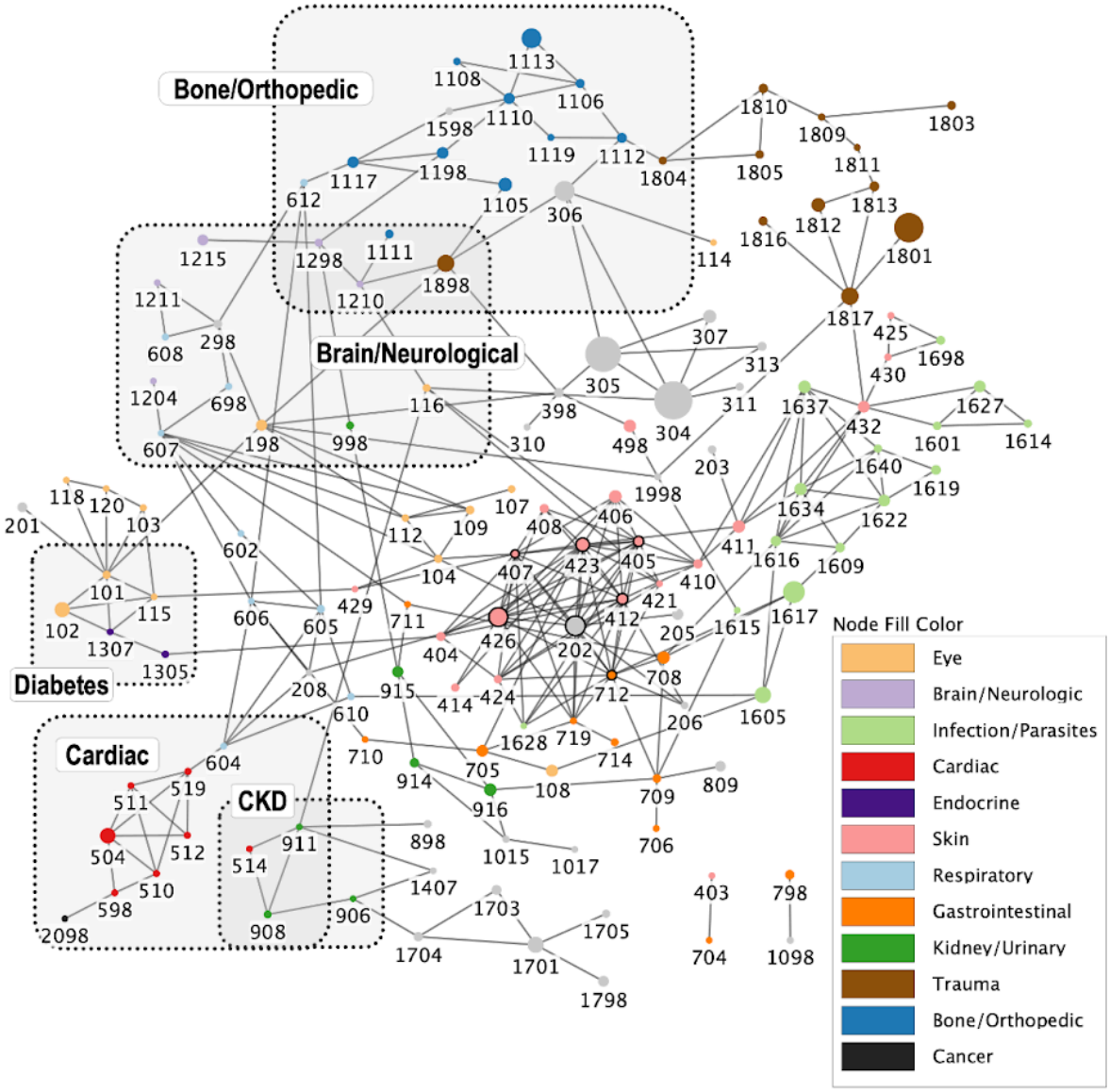

Figure 1. Undirected comorbidity network. Nodes are health conditions, and edges are statistically significant comorbidity connections (at a Bonferroni adjusted P-value < 0.001). Node sizes represent lifetime prevalence, with more common health conditions having larger nodes. Nodes with degrees greater than 10 are outlined in bold. Additionally, several health conditions or health condition groups of interest in geroscience, as outlined by the National Institute of Aging, as well as their first-order neighbors are highlighted with shaded boxes [44]. These health condition groups as well as the five health condition categories with the highest number of affected dogs have colored nodes. All other categories have gray nodes. Health condition categories were created using the categorization system in the DAP survey and the majority of the categories highlighted in the network can best be described as organ systems (i.e., respiratory, cardiac) as opposed to pathophysiological processes (i.e., trauma, infection) [45].

Refer to Table S1b for the corresponding health condition name associated with each numeric code.

In addition to well-known comorbidities, this network also supports less well-studied health condition pairs. For example, this network suggests an association between proteinuria (node 911) and anemia (node 1407) ($P < 0.001$). 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 [46]. 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 [47–49]. The network illustrates these relationships by connecting anemia to CKD (node 908) 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 [49–51]. The connections revealed by our network demonstrate its potential in identifying less obvious health condition associations and comorbidities.

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 [52], 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 [53]. The node degree distribution of our network is better fitted with an exponential distribution than a scale-free power-law distribution ($P < 0.001$) (Fig. S2), providing evidence for a random network.

Comparison to stratified networks

We stratified the dataset into four groups by life stage (Puppy: n=158; Young Adult: n=3487; Mature Adult: n=15784; Senior: n=6119) to investigate the extent to which comorbidity networks varied across life stages. As expected, the four networks constructed using the stratified approach differed greatly from each other and also from the unstratified network (**Fig. 2**). Notably, no significant comorbidity patterns were detected for the puppy age group, which comprised only 2% of the cohort. This likely reflects the relatively low frequency of multiple health conditions in young dogs rather than simply a statistical power limitation.

Ten comorbidity pairs were identified across all four networks: roundworms and tapeworms (nodes 1634 and 1637), hookworms and roundworms (nodes 1622 and 1634), chronic/recurrent diarrhea and gastrointestinal food/medicine allergies (nodes 708 and 712), chronic/recurrent diarrhea and chronic/recurrent vomiting (nodes 708 and 709), atopic dermatitis and pruritus (nodes 405 and 423), pruritis and seasonal allergies (nodes 423 and 426), extracted teeth and fractured teeth (nodes 305 and 306), skin food/medicine allergies and gastrointestinal food/medicine allergies (nodes 412 and 712), skin food/medicine allergies and seasonal allergies (nodes 412 and 426), and fleas and ticks (nodes 411 and 432). Conversely, there were 14 unique comorbidity pairs identified by the stratified networks but not by the unstratified network. Ten of these pairs came from the mature adult group and two from the senior group (Fig. S4). On the other hand, the unstratified network identified 113 comorbidities that were not detected by *any* of the stratified networks. However, this could be due to lower statistical power in the stratified analysis.

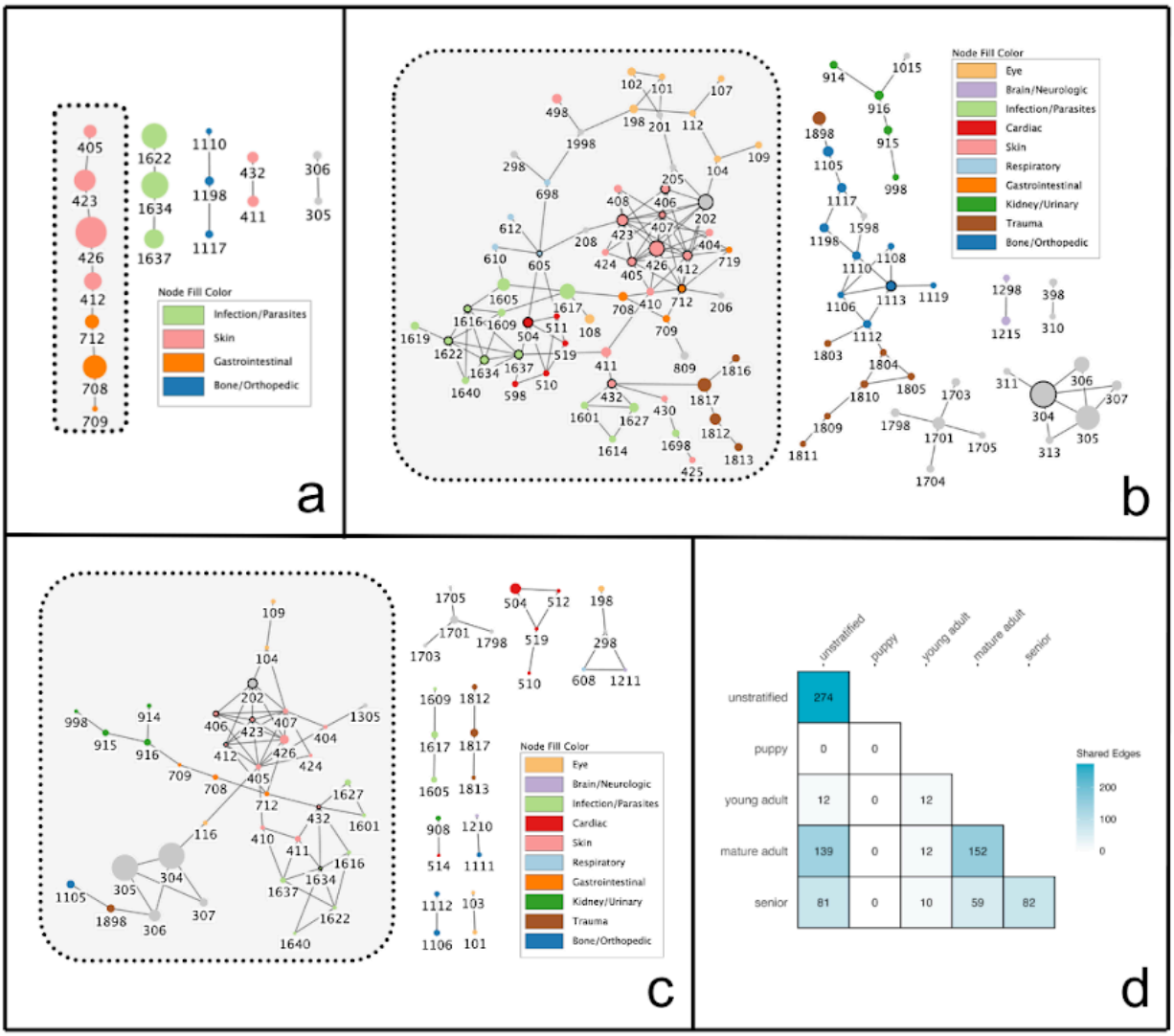

Figure 2. Age-stratified comorbidity networks. (a) Young Adult, (b) Mature Adult, and (c) Senior dog networks. In each network, nodes represent health conditions with node size proportional to condition prevalence, and edges represent statistically significant comorbidity associations (Bonferroni-adjusted $P < 0.01$). Nodes with the highest degree in each network are indicated with bold outlines, and the largest connected subnetwork is highlighted with a dotted shaded box. No network is shown for the Puppy stratum as no significant comorbidities were detected in this age group. (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 edges per network.

Network topology analysis of the age-specific networks revealed distinct patterns in disease interconnectedness (Table 2). 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 2. 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	0.032	0.043
Clustering Coefficient	N/A	0	0.473	0.484
Betweenness Centrality	N/A	0.0604	0.125	0.136
Degree Centrality	N/A	0.0417	0.0831	0.108

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). A time-directed comorbidity network is presented in Fig. 3. This makes the network more informative, as the directed network gives suggestions about health condition progression. For instance, the

directed network suggests that diabetes (node 1307) occurs before cataracts (node 102) ($P < 0.001$). This is consistent with the literature: it is estimated that around 75% of canines diagnosed with diabetes will have blinding cataracts within the following two years [54]. Other examples of directional pairings supported by existing literature can be found in Fig. S5.

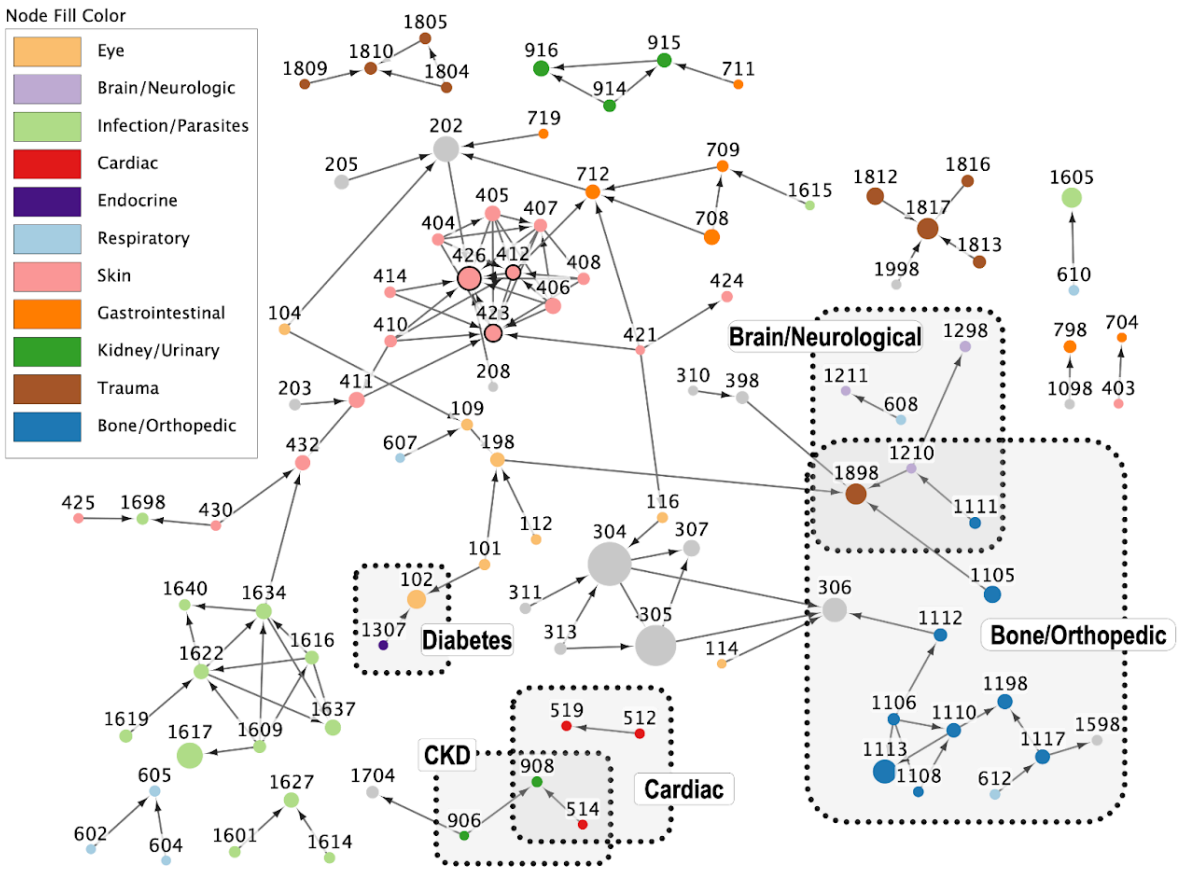

Figure 3. Time-directed network using a window size of 12 months. Nodes are health conditions, and edges are statistically significant comorbidity connections (at a Bonferroni-adjusted P-value < 0.01). Arrowheads point from the health condition that occurs earlier in time to the health condition that occurs later. Node sizes represent prevalence, with more common health conditions having larger nodes. Nodes with total node degree greater than 10 are outlined in bold. Additionally, several health conditions or health condition groups of interest in geroscience, as well as their first-order neighbors are highlighted with shaded boxes [44]. These health condition categories as well as the top five best-represented health condition categories also have colored nodes. All other categories have gray nodes. Refer to Table S1b for the corresponding health condition name associated with each numeric code.

As in the undirected network (**Fig. 1**), the directed network also exhibits a cluster of parasite and parasite-related health conditions. However, the mechanisms of multiparasitism, especially in companion dogs, are poorly understood because previous work focuses on a one host-one parasite paradigm [42]. The addition of temporal information in this network could support future research endeavors in this area.

Materials and Methods

Dog Aging Project owner survey data

The DAP is a nationwide, long-term, longitudinal study on the biological, environmental, and lifestyle determinants of healthy aging in companion dogs [22]. All data are made publicly available (other than identifiable owner information) through the Terra platform at the Broad Institute at MIT and Harvard. 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 technically cross-sectional in nature

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, Table S1b) 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 [31,33–35,55,56]. 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 [57]. Despite being an imprecise summary, breed background has previously been reported to be associated with several health conditions [35]. 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 $P_{y \in d}$:

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

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:

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

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

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:

$$P_{\{y,z\}\subset d} = P_{y\in d} \times P_{z\in d}. \tag{4}$$

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 [18,58]. 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:

$$\mu_{y,z} = \sum_{d} P_{\{y,z\}\subset d}. \tag{5}$$

The variance is

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

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 Equation (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

$$P_{y \to z}^{\{y,z\} \subset d} = \frac{P_{y \in d}}{P_{y \in d} + P_{z \in d}}. \tag{9}$$

We also incorporated a time window size W and the length of the dog's medical history as done in Lemmon *et al*. [18]. 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. S3). Knowing the span of each dog's medical history allows us to study directed comorbidities across prespecified time windows. Here we chose a window of 12 months. Now,

the probability that y, z occur within W given that the canine already has the two health conditions is:

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

$$P_{\{y,z\}\subset d\ in\ W} = 1\ if\ otherwise. \tag{10}$$

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

$$P^{\{y,z\}\subset d}_{y\to z\ in\ W} = P^{\{y,z\}\subset d}_{y\to z} \times P_{\{y,z\}\subset d\ in\ W}. \tag{11}$$

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.

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, breed, and sex in addition to age [59]. 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 (Fig. S2).

Discussion

This paper leverages data from the DAP, which encompasses a large epidemiological cohort, to study the canine comorbidity network among 166 different health conditions. Key demographic characteristics including age, weight, sex, sterilization status, and breed background were adjusted for using LRMs before constructing the comorbidity network. Our LRMs support the contribution of key demographic characteristics to the risk of health conditions [31,33–35,55,56].

The undirected canine comorbidity network reveals both well-known comorbidities and suggests potential connections between health conditions that have been less studied [32,37–39,46,49]. This illustrates the potential for using health condition association networks to aid with novel comorbidity discovery. 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 [60,61].

When comparing the unstratified network to the four networks created after stratification by life stage, the unstratified network identifies many associations not identified by stratified networks.

This makes sense due to the reduction in statistical power after stratification. On the other hand, stratification revealed important age-specific patterns. For instance, within the stratified networks, only the senior network identified comorbidities in the cardiac and neurological categories. Both categories are of particular interest in aging research. 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. Such a combined approach allows for a comprehensive understanding of both overall comorbidity patterns and age-specific associations.

Finally, we also constructed a time-directed network using a window size of one year. Shorter window spans are more adept at capturing comorbidities that occur sequentially [1], but could result in power loss in comorbidity discovery since shorter window sizes encompass fewer dogs. The ideal window size should be guided by clinical knowledge and research objectives [18]. Incorporating the time of reported condition onset provides us with some insights into health condition progression, which could be beneficial for canine health management. For instance, this time-directed network exhibits a cluster of parasite-parasite interactions. While interest in co- or multiparasitism has increased in recent years, more work is needed to understand the mechanisms of these interactions [41–43,62,63]. The temporal information encoded in the directed network may suggest potential synergistic relationships between parasites [42]. Understanding and controlling companion dog parasitism is becoming increasingly important as pet ownership, particularly of dogs, continues to rise globally [64]. This growing human-animal interface elevates the significance of our findings beyond veterinary medicine, potentially informing public health strategies related to zoonotic health conditions [64–66].

There are a few limitations in this study. First, the use of owner-reported lifetime prevalence collected at a single point in time introduces risk of recall bias. Owners may not accurately recall all of the dog's health conditions, or the age at occurrence of each condition. This limitation will be overcome with time as the DAP continues longitudinal data collection; repeating the survey-based inventory of health conditions each year for each dog shortens the interval required for owner recall to one year. Second, rare health conditions are susceptible to overfitting by LRMs, which may result in an inaccurate representation of comorbidity between rare and prevalent health conditions in the network. One example of this is pododermatitis (node 421, n=65), a skin condition, and allergies (node 412, n=1015). General clinical knowledge suggests that allergies are a common cause of pododermatitis [67]. However, this pairing is not observed in the network. Third, owners may misreport health status information. The list of health conditions in the DAP surveys was designed to make the survey easier for owners to complete, so some of the health conditions may 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. This is directly observed in the network as there are several instances where one pathological process is instead represented as a comorbid pair, like laryngeal paralysis duplicated in both respiratory and brain/neurological conditions (nodes 608 and 1211). Beyond general recall accuracy, the temporal sequence of conditions may be particularly challenging for owners to report accurately, as concurrent or gradually developing conditions can be difficult to order chronologically. The owner-reported dates of condition onset do not necessarily represent the pathophysiological course of a health condition, or may themselves be inaccurate. This issue likely explains some counterintuitive results like tick-borne health conditions preceding ticks in the time-directed network (Fig. 4). Fourth, 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, in dogs without kidney disease, blood pressure may not be measured at all, meaning that lack of owner knowledge/reporting of their dog's hypertension is not equivalent to absence of hypertension in their dog. Furthermore, at least at the time of enrollment, dogs in the Dog Aging Project 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 [68–70]. Finally, the owners in our study cohort are predominantly white, highly educated, and wealthier than the US population, which may raise concern on the generalizability of the identified comorbidities (Fig. S6).

Acknowledging these limitations, this study demonstrates the power of the DAP data, 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, 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.

Acknowledgements

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

Conflicts of Interest

DP is a paid consultant of WndrHLTH Club, Inc. and Infinity Research Labs.

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

All statistical analyses were done in R and network visualizations came from Cytoscape. The code for reproducing the results described in this study can be found at https://github.com/aeyfang/pbc-dap-network/.

Funding

This research is based on publicly available data collected by the Dog Aging Project, under U19 grant AG057377 (PI: Daniel Promislow) from the National Institute on Aging, a part of the National Institutes of Health, and by additional grants and private donations, including generous support from the Glenn Foundation for Medical Research, the Tiny Foundation Fund at Myriad Canada, the WoodNext Foundation, and the Dog Aging Institute. DP received support from USDA cooperative agreement USDA/ARS 58-8050-9-004. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

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;7:357–63. https://doi.org/10.1370/afm.983.

[2] Feinstein AR. The pre-therapeutic classification of co-morbidity in chronic disease. J Chronic Dis 1970;23:455–68. https://doi.org/10.1016/0021-9681(70)90054-8.

[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;74:1046–54. https://doi.org/10.1136/thoraxjnl-2017-210822.

[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;15:1968–78. https://doi.org/10.1109/TCBB.2018.2827029.

[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;101:103396. https://doi.org/10.1016/j.ijnurstu.2019.103396.

[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;4:1–12. https://doi.org/10.21926/obm.geriatr.2003127.
[8] Divo MJ, Martinez CH, Mannino DM. Ageing and the epidemiology of multimorbidity. Eur Respir J 2014;44:1055–68. https://doi.org/10.1183/09031936.00059814.
[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;62:275–80. https://doi.org/10.1093/gerona/62.3.275.
[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;70:332–8. https://doi.org/10.1093/gerona/glu113.
[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;67:124–32. https://doi.org/10.1016/j.critrevonc.2008.01.013.
[12] Folino F, Pizzuti C, Ventura M. A Comorbidity Network Approach to Predict Disease Risk. In: Khuri S, Lhotská L, Pisanti N, editors. Inf. Technol. Bio- Med. Inform. ITBAM 2010, Berlin, Heidelberg: Springer; 2010, p. 102–9. https://doi.org/10.1007/978-3-642-15020-3_10.
[13] Goh K-I, Cusick ME, Valle D, Childs B, Vidal M, Barabási A-L. The human disease network. Proc Natl Acad Sci U S A 2007;104:8685–90. https://doi.org/10.1073/pnas.0701361104.
[14] Hidalgo CA, Blumm N, Barabási A-L, Christakis NA. A Dynamic Network Approach for the Study of Human Phenotypes. PLOS Comput Biol 2009;5:e1000353. https://doi.org/10.1371/journal.pcbi.1000353.
[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;7:e1002141. https://doi.org/10.1371/journal.pcbi.1002141.
[16] Fotouhi B, Momeni N, Riolo MA, Buckeridge DL. Statistical methods for constructing disease comorbidity networks from longitudinal inpatient data. Appl Netw Sci 2018;3:46. https://doi.org/10.1007/s41109-018-0101-4.
[17] Moni MA, Xu H, Liò P. CytoCom: a Cytoscape app to visualize, query and analyse disease comorbidity networks. Bioinformatics 2015;31:969–71. https://doi.org/10.1093/bioinformatics/btu731.
[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;1:694–702. https://doi.org/10.1038/s43588-021-00141-9.
[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;11:4952. https://doi.org/10.1038/s41467-020-18682-4.
[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;12:177. https://doi.org/10.1186/s12920-019-0629-x.
[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 2023;6. https://doi.org/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;602:51–7. https://doi.org/10.1038/s41586-021-04282-9.

[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;10:2962. https://doi.org/10.1038/s41598-020-59837-z.

[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;229:1591–601. https://doi.org/10.2460/javma.229.10.1591.

[25] Dinwoodie I, Dwyer B, Zottola V, Gleason D, Dodman N. Demographics and Comorbidity of Behavior Problems in Dogs. J Vet Behav 2019;32. https://doi.org/10.1016/j.jveb.2019.04.007.

[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;59:681–90. https://doi.org/10.1111/jsap.12904.

[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;177:306–15. https://doi.org/10.1136/vr.103360.

[28] Prudic RA, Saba CF, Lourenço BN, Bugbee AC. Prevalence of proteinuria in a canine oncology population. J Small Anim Pract 2018. https://doi.org/10.1111/jsap.12840.

[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;45:254–60. https://doi.org/10.1111/vsu.12444.

[31] 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:e90501. https://doi.org/10.1371/journal.pone.0090501.

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

[33] 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;33:89–99. https://doi.org/10.1111/jvim.15367.

[34] 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;19:e0295840. https://doi.org/10.1371/journal.pone.0295840.

[35] 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 2023;10. https://doi.org/10.3389/fvets.2023.1140417.

[36] 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. https://doi.org/10.2460/javma.242.11.1549.

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

[38] Acierno MJ, Labato MA. Hypertension in renal disease: diagnosis and treatment. Clin Tech Small Anim Pract 2005;20:23–30. https://doi.org/10.1053/j.ctsap.2004.12.004.

[39] Syme H. Hypertension in small animal kidney disease. Vet Clin North Am Small Anim Pract 2011;41:63–89. https://doi.org/10.1016/j.cvsm.2010.11.002.

[40] Bajwa J. Canine otitis externa — Treatment and complications. Can Vet J 2019;60:97–9.
[41] 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;116:727–35. https://doi.org/10.1093/trstmh/trac002.
[42] 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;8:545. https://doi.org/10.1186/s13071-015-1167-9.
[43] 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;13:e0007455. https://doi.org/10.1371/journal.pntd.0007455.
[44] NIA Strategic Directions 2020-2025. Natl Inst Aging n.d. https://www.nia.nih.gov/about/aging-strategic-directions-research (accessed June 20, 2024).
[45] Fleming J m., Creevy K e., Promislow D e. l. Mortality in North American Dogs from 1984 to 2004: An Investigation into Age-, Size-, and Breed-Related Causes of Death. J Vet Intern Med 2011;25:187–98. https://doi.org/10.1111/j.1939-1676.2011.0695.x.
[46] Harley L, Langston C. Proteinuria in dogs and cats. Can Vet J 2012;53:631–8.
[47] Polzin DJ. Chronic Kidney Disease in Small Animals. Vet Clin North Am Small Anim Pract 2011;41:15–30. https://doi.org/10.1016/j.cvsm.2010.09.004.
[48] 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 2023:1–6. https://doi.org/10.2460/ajvr.23.05.0097.
[49] Sannamwong N, Buranakarl C, Sutayatram S, Trisiriroj M, Dissayabutra T. The first study on urinary loss of iron and transferrin in association with proteinuria in dogs with chronic kidney disease. Vet World 2023;16:154–60. https://doi.org/10.14202/vetworld.2023.154-160.
[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;100:1227–39. https://doi.org/10.1016/j.kint.2021.08.024.
[51] Drüeke TB, Massy ZA. Role of proteinuria in the anemia of chronic kidney disease. Kidney Int 2021;100:1160–2. https://doi.org/10.1016/j.kint.2021.09.016.
[52] Freeman LC. Centrality in social networks conceptual clarification. Soc Netw 1978;1:215–39. https://doi.org/10.1016/0378-8733(78)90021-7.
[53] Newman MEJ. Networks. Second edition. Oxford: Oxford University Press; 2018.
[54] Williams DL. Effect of Oral Alpha Lipoic Acid in Preventing the Genesis of Canine Diabetic Cataract: A Preliminary Study. Vet Sci 2017;4:18. https://doi.org/10.3390/vetsci4010018.
[55] Hedhammar ÅA, Malm S, Bonnett B. International and collaborative strategies to enhance genetic health in purebred dogs. Vet J Lond Engl 1997 2011;189:189–96. https://doi.org/10.1016/j.tvjl.2011.06.018.
[56] 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;17:69. https://doi.org/10.1186/s12917-021-02775-3.
[57] King G, Zeng L. Logistic Regression in Rare Events Data. Polit Anal 2001;9:137–63. https://doi.org/10.1093/oxfordjournals.pan.a004868.
[58] Hong Y. On computing the distribution function for the Poisson binomial distribution. Comput Stat Data Anal 2013;59:41–51. https://doi.org/10.1016/j.csda.2012.10.006.
[59] 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. https://doi.org/10.2460/javma.23.07.0366.

[60] Brunson JC, Agresta TP, Laubenbacher RC. Sensitivity of comorbidity network analysis. JAMIA Open 2020;3:94–103. https://doi.org/10.1093/jamiaopen/ooz067.
[61] Broido AD, Clauset A. Scale-free networks are rare. Nat Commun 2019;10:1017. https://doi.org/10.1038/s41467-019-08746-5.
[62] Souza JBB, Silva ZM de A, Alves-Ribeiro BS, Moraes I de S, Alves-Sobrinho AV, Saturnino KC, et al. Prevalence of Intestinal Parasites, Risk Factors and Zoonotic Aspects in Dog and Cat Populations from Goiás, Brazil. Vet Sci 2023;10:492. https://doi.org/10.3390/vetsci10080492.
[63] Vanden Broecke B, Bernaerts L, Ribas A, Sluydts V, Mnyone L, Matthysen E, et al. Linking Behavior, Co-infection Patterns, and Viral Infection Risk With the Whole Gastrointestinal Helminth Community Structure in Mastomys natalensis. Front Vet Sci 2021;8:669058. https://doi.org/10.3389/fvets.2021.669058.
[64] Giannelli A, Schnyder M, Wright I, Charlier J. Control of companion animal parasites and impact on One Health. One Health 2024;18:100679. https://doi.org/10.1016/j.onehlt.2024.100679.
[65] Day MJ. One health: the importance of companion animal vector-borne diseases. Parasit Vectors 2011;4:49. https://doi.org/10.1186/1756-3305-4-49.
[66] Overgaauw PAM, Vinke CM, van Hagen MAE, Lipman LJA. A One Health Perspective on the Human–Companion Animal Relationship with Emphasis on Zoonotic Aspects. Int J Environ Res Public Health 2020;17:3789. https://doi.org/10.3390/ijerph17113789.
[67] Bajwa J. Canine pododermatitis. Can Vet J 2016;57:991–3.
[68] 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;165:874–81. https://doi.org/10.1093/aje/kwk075.
[69] 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;49:783–90. https://doi.org/10.1016/0895-4356(96)00009-1.
[70] Ederer F, Church TR, Mandel JS. Sample sizes for prevention trials have been too small. Am J Epidemiol 1993;137:787–96. https://doi.org/10.1093/oxfordjournals.aje.a116739.

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, including original DAP database codes, reference codes used in network analyses, standardized condition names, body system classifications, and frequency of occurrence in the study cohort.

S2 Table. Logistic regression model parameters. Results from the logistic regression models examining associations between demographic factors and health conditions. For each health condition, coefficients and corresponding P-values are shown for age (years), weight (pounds), breed background (purebred vs mixed-breed), and sex/reproductive status. The intercept term represents the log odds when all predictors are zero. Values in bold indicate positive coefficients with statistical significance at $P < 0.05$.

S3 Table. Health condition pair associations. (a) Complete results from the unstratified, undirected comorbidity network analysis. (b) Results from directed network analysis showing temporal relationships between conditions occurring within a 12-month window. (c-f) Results from age-stratified analyses for Puppy, Young Adult, Mature Adult, and Senior dogs respectively. For all analyses, health condition pairs are shown with their numerical codes, names, raw P-values and Bonferroni-adjusted P-values. For undirected networks (a, c-f), source and target labels are arbitrary designations. For the directed network (b), source indicates the condition that occurred first. No significant associations were found in the Puppy stratum (c).

S1 Fig. Top: Histogram representing number of dogs with different numbers of diseases. Bottom: Histogram representing top 20 most frequently occurring diseases.

S2 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.

S3 Fig. Distribution of medical history length. Box plot comparing each dog's age at last reported health condition with their total medical history span (calculated as time between first and last reported condition.)

S4 Fig. Health condition pairs uniquely identified in age-stratified networks. Health condition pairs that were significant in only age-stratified analyses but not in the unstratified network. Each pair is shown with condition names, numerical codes, and the age stratum where the association was detected.

S5 Fig. 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.

S6 Fig. 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 (N = 26,523) compared to 2021 US Census data. Note that owners could select multiple racial categories, so percentages may sum to more than 100%. Income data from owners was collected in categorical ranges.