```
# install.packages("readr")
# install.packages("dplyr")
```

Team 6 - Data 602

Analysis of Canine Degenerative myelopathy (DM) in purebred and mixed breed pet dogs (Canis familiaris)

Yanwei Yu: 30249113

Sai Krishnam Raju Pusapati: 30231852

Pradeepa Remamurthy: 30270796

Jennifer Stadtfelder: 30250230

Motivation and purpose

With the establishment of organizations such as the Kennel Club and the AKC, dog breeding became subject to stricter regulation. In order to become a member of these prestigious clubs, the animals had to descend from registered parents of the same breed. Today, there are numerous breed clubs around the world that collect information about their members and set clear standards for the appearance and behavior of dogs. At shows organized by these clubs, the dogs are judged according to these criteria. The winners of such competitions develop into sought-after breeding animals, as the aim is to pass on the best characteristics to future generations (Clements et al, 2015) .

By selectively breeding the winning type, only within the same dog breed, the so-called pure dog breeds remain genetically separate from others. Historical events such as wars have also led to purebred dogs being bred in isolation, often through line breeding and the use of popular stud dogs. Problems arise when a breed is based on a small number of parent dogs and one or more of these dogs are carriers of disease genes. Extreme expressions of breed traits that have a genetic bias can also lead to severe health problems and the inheritance of these traits. In such cases, diseases will likely occur more frequently in the offspring. (Clements et al, 2015) Supporting that, in 2019 Donner et al. discovered that there are differences in the mutations a mixed breed is developing compared to a pure breed. He concluded that there were diseases specific to certain breeds or bloodlines so that mixed breeds were more likely to be carriers of common recessive diseases and purebred dogs more likely to have genetic diseases (Donner et al, 2019).

At the same time, the study of genetic diseases in dogs is becoming particularly relevant in the light of the growing interest in dogs as pets and the associated professionalization of the breeding industry. According to data from the American Kennel Club, the number of registered dogs has risen sharply in recent years, which has led to increased interest in pure breeds. (Clements et al, 2015) In Europe, there has been an increase in dog ownership between 1988 and 2022. (Shahbandeh, 2024) Both are reflected in the rising profits of the pet industry and pet insurance companies. (Grand View Research, 2024) In this context, the risks of pure breeds and mixed breeds are particularly relevant for various players. Insurance companies calculate premiums based on the risk of illness in dogs, but owners also want to consider the risks and associated costs. Furthermore, the potential increase in cases of genetically diseased dogs places an additional burden and challenge on veterinarians who, on a statistical basis, conclude that they need to research the disease and its treatment further.

On the other hand, a 2014 Swedish study found that selective breeding and inbreeding were not identified as a major cause of health problems in many purebred dog breeds. (Laikre and Jansson, 2014) These findings were also confirmed 2015. The study concluded that only a low correlation was found between the degree of inbreeding or reduced heterozygosity and the incidence of genetic diseases, suggesting that many harmful genes were acquired before the breed was officially registered and the current breeding patterns began. (Mellanby et al., 2013) Subsequently, Belanger et al. found that there were no differences between purebred

and mixed breeds in terms of certain genetic disorders between 19995 and 20210, indicating a reduction of genetic diseases through targeted breeding practices and gene scans, which were increasingly recommended by breeding clubs to maintain a good pure breed reputation. (Belanger et al., 2013) In addition to that, he has also been disproved that pure breeds are more susceptible to developing and contracting diseases. Vets and dog owners generally exclude the possibility of genetic diseases as a cause of health problems in mixed-breed dogs, as they assume that the mutations in individual genes are not a cause for concern. He, however, argues that there are no differences in the development of genetic diseases between pure breeds and mixed breeds, as mixed breeds develop and inherit just as many mutations as the pure breed. This was also supported by Clements et al., who discovered that mutations seemed to occur in a similar way regardless of breed background (Clements et al., 2015). Supporting that, recent study also did not find that designer crosses have poorer health or shorter lifetime prevalence compared to mixed breeds. (Brand et al., 2024) This, however, is not attributed to advances in veterinary medicine. A higher proportion of purebred dogs than mixed breeds had no owner-reported diseases, concluding that purebreds are not at higher risk of inherited genetic diseases than mixed breeds (Forsyth et al., 2023).

This raises the question of why Donner et al. obtained such a different research result. (Donner, 2019) This difference could be due not only to the fact that the different studies focused on different diseases in purebreds and mixed breeds, but also to the methodology used. With large samples, even a very small effect can become statistically significant. The central limit theorem states that for sufficiently large samples, the distribution of sample means is normalized independently of the population distribution. This leads to a high probability of finding significant differences, even if these differences are irrelevant in practice. (Cohen, 1988) Since not all genetic diseases are analyzed in detail uniformly and appropriate methodology, it is not possible to make a comprehensive assessment of the frequency and prevalence of all genetic diseases in purebred and mixed-breed dogs. This limitation is of particular importance as different breeds have different susceptibilities to different hereditary diseases. Following that, this raises the question, are certain genetic diseases more common in purebreds than in mixed breeds? As a result, we will review the Donner et al. dataset again using a more appropriate methodology: confidence interval and chi-square test. (Donner, 2019) These tests will give more certainty and insights in which of the two hypotheses holds true:

Bellumori et al.: No differences in genetical diseases of pure breeds and mixed breeds

Donner et al.: Purebred dogs have a higher likelihood of genetic diseases

However, we will focus our analysis on one of the most common progressive neurodegenerative diseases affecting dogs in late adulthood in the United States (Coates & Wininger, 2010; Ivansson et al., 2016; Neeves & Granger, 2015): Canine Degenerative myelopathy (DM). The affected dogs with this disease suffer from weak muscle and loss of coordination due to spinal cord deficiency (Cornell University, 2024). However, this disease may not show and affect dogs during their earlier stages, and it is not fatal until symptoms appear (Kobatake et al., 2021). Studies (Ivansson et al., 2016; Kobatake et al., 2021) found that this disease is controlled at a single locus with two alleles, and an individual with homozygous recessive genotype would show symptoms in late adult years. Note that allele means one of two (or more) versions of DNA sequence at each locus, and locus (loci) means a specific location at a chromosome for a gene.

Data collection

Population:

The total population we intend to examine, includes all diseased genotypes of degenerative myelopathy (DM) in domestic dogs (Canis familiaris) whose owners/families live/have their primary residence in the United States. The dogs can be mixed breed or purebred as long as they are kept as pets by owners/families who live/primarily reside in the United States.

Raw Data:

We found our original raw dataset in Dryad (Donner et al., 2019). Overall, our dataset contains information about dogs, the countries of their owners, the dog breeds and the healthy or diseased genetic composition. The diseased genetic composition is examined in relation to 152 genetic diseases (one gene has the disease and the other does not, both genes have the disease neither gene has the disease). In genetic research, it is

considered a healthy combination/genotype if neither or one genes have the disease, while it is considered a diseased combination/genotype if both genes have the diseased gene (Claussnitzer et al., 2020).

Specifically, the dataset includes samples of mixed dogs (83220) entered on the Wisdom Panel platform between February 2015 and May 2016. To focus only on mixed breeds, F1 generation hybrids were excluded. Most of these dogs originated from the USA, followed by the UK and Australia. Furthermore, the most common breeds in the mixed breeds were American Staffordshire Terrier, Labrador Retriever and German Shepherd (Donner et al, 2019).

For the comparison of mixed breeds to purebreds, a separate purebred sample (18102) of dogs sampled at various laboratories between 2005 and 2016 were analyzed. The majority also came from the USA and Finland, and dogs were classified as purebred if they were registered with recognized breeding associations. A total of 330 breeds were included in the data set. (Donner et al, 2019)

The genotype data in the data set were based on DNA samples obtained either by non-invasive cheek swabs or blood sampling at certified veterinary clinics. All owners gave their consent for the samples to be used for research purposes (Donner et al, 2019).

```
# Load necessary libraries
suppressPackageStartupMessages({
  library(readr)
  library(dplyr)
})

# Read the CSV file
dog_data <- read_csv("C:/Users/pskra/Desktop/DATA 602/project/DATA-602-PROJECT-L03-6/Full_genotype_datas

# Display the head of the data
head(dog_data)
```

```
## # A tibble: 6 x 155
##   `Sample identifier` Breed     `Region of origin` Alaskan Husky Encephalopath~1
##                 <dbl> <chr>     <chr>              <chr>
## 1                   1 Greyhound United States      A1/A1
## 2                   2 Greyhound United States      A1/A1
## 3                   3 Greyhound United States      A1/A1
## 4                   4 Greyhound United States      A1/A1
## 5                   5 Greyhound United States      A1/A1
## 6                   6 Greyhound United States      A1/A1
## # i abbreviated name: 1: `Alaskan Husky Encephalopathy (AHE)`
## # i 151 more variables: `Amelogenesis Imperfecta (AI)` <chr>,
## #   `Autosomal Recessive Severe Combined Immunodeficiency (ARSCID)` <chr>,
## #   `Bandera's Neonatal Ataxia (BNAt)` <chr>,
## #   `Benign Familial Juvenile Epilepsy or Remitting Focal Epilepsy (BFJE)` <chr>,
## #   `Bleeding disorder due to P2RY12 defect` <chr>,
## #   `Canine Cyclic Neutropenia (CN)` <chr>, ...
```

Sample:

In relation to our population, we are only interested in the sample data of the diseased genotypes of dogs in the United States. Since we were only interested in the pet dogs owned in the United States, we filtered our data and kept 85,074 samples from domestic dogs in the United States. Among the U.S. data, there were 77,048 samples for mixed-breed dogs and 7,542 purebred dogs.

Furthermore, in our data, A1 stands for the dominant allele (normal allele) and A2 for the recessive allele (defective allele). Since the A1 allele is dominant, this means that an individual with A1/A1 genotype and

A1/A2 (although this type of dog has a defective allele) will not show DM symptoms (healthy genotypes), which represents healthy dogs. The individual with the A2/A2 genotype suffers from degenerative myelopathy (DM) and the symptoms do not appear until late adulthood (diseased genotype). For our analysis, the DM disease genotypes are filtered accordingly.

Reliability and Validity:

The responsible handling of data plays a crucial role, especially when it comes to personal health data. For example, the anonymized data sets were created with the consent of those who provided the data (in this case the dog owners) to protect private data.

In addition, the question arises as to how the different populations within these data sets - especially mixed breeds and purebreds - can be appropriately generalized, compared and interpreted, as they show considerable differences.

On the one hand, the number of pure breeds vs. mixed breeds differs greatly from each other. This can lead to a possible distortion of the results, as the smaller group can have a disproportionately large influence on the analysis. In statistical analyses, especially in significance tests such as t-tests, the size of the two groups is very important. Unequal group sizes can influence the calculation of test statistics and the resulting p-values, leading to incorrect assessments of significant differences. In addition, different group sizes can lead to unfair conditions, systematically disadvantaging or favouring the larger group, which jeopardizes the validity of the results.

On the other hand, the number of breeds also varies within this subdivision. For example, diseases of popular purebred dogs such as German shepherds are more prevalent than those of rarer breeds such as the Pomeranian. This can lead to false conclusions being drawn about genetic diseases in all pure dog breeds in the United States. This includes, for example, that if canine degenerative myelopathy (DM) occurs mainly in popular dog breeds, it is assumed that it is more prevalent in all pure breeds, although Pomeranians, for example, are hardly affected.

The comparability of data collected from different platforms and over different time spans can present several challenges as well: Differences in data collection methodology (e.g. sample selection, data collection methods, timing of data collection) can lead to bias. Data collected over different time spans may reflect changes in trends, such as breeding becoming more concerned with breeding healthy purebred dogs. Additionally, different platforms may use different methodologies, technologies or algorithms to collect data, which can also lead to variations in results.

Furthermore, the genetic diseases of minority breed dogs in the USA and those that were not kept in the USA during the period are not equally represented in the dataset. Therefore, no conclusions can be drawn from the analysis about the overall difference between mixed and pure breeds worldwide.

Finally, the dataset is outdated and does not provide an up-to-date overview of the healthy/diseased genotypes of pet dogs owned by owners who either live in the USA or have their main place of residence there.

These aspects are also considered in the interpretation of the results.

```r
# Filter data only in America
dog_USA_data <- dog_data %>%
  filter(`Region of origin` == "United States")

# Clean data (delete all rows with blank, and keep the data only with "A1/A1", "A1/A2", "A2/A2")
# Using one of the most popular genetic diseases: Degenerative Myelopathy (DM) as an example
cleaned_dog_USA_data <- dog_USA_data %>%
  filter(if_all(everything(), ~ . != "" & !is.na(.)))
cleaned_dog_USA_data <- cleaned_dog_USA_data %>%
  filter(`Degenerative Myelopathy (DM)` %in% c("A1/A1", "A1/A2", "A2/A2"))
```

```r
# make two subgroup, one is mixed breed, and the other is pure breed.
mixed_breed <- cleaned_dog_USA_data %>%
  filter(Breed == "Mixed breed")
pure_breed <- cleaned_dog_USA_data %>%
  filter(Breed != "Mixed breed")
```

Two-sample proportion test and Confidence interval:

The variable & parameters:

In our sample, we focus on pet dogs (mixed and pure breeds) owned by dog owners/families in the United States and their disease degenerative myelopathy (DM) genotypes.

Two parameters are considered:

Proportion of mixed breeds: Proportion of diseased genotypes (A2/A2) of the disease Canine Degenerative Myelopathy (DM).

Proportion of pure breeds: Proportion of diseased genotypes (A2/A2) of the disease Canine Degenerative Myelopathy (DM).

Statistical analysis:

To test that mixed breed, have a higher incidence of diseased gene combinations than purebred dogs, two methodological approaches were applied:

Due to concerns about the "Reliability and Validity" of the different populations and the large sampling bias mentioned in the "Issue" section, a two-sample test for comparing two population proportions was chosen. This provides information on the frequency range in which the DM disease genotype occurs with 95% certainty in the entire population of purebred/mixed breed dogs in the USA. Further, we will analyse if the confidence interval of the two populations (mixed breed/pure breed in the USA) includes the value 0. This indicates no significant difference in the frequency range of DM disease gene combinations between the two groups. If the interval contains positive values/on the right side of zero, this could indicate that mixed-breed dogs have significantly lower diseased genotype. If the interval contains negative values, this could indicate that purebred dogs have lower diseased genotype.

Statistical Interpretation:

Null hypothesis: There are equal proportions of individuals with A2/A2 genotypes in both purebred dog population and mixed breed dog population.

Alternative hypothesis: There are more A2/A2 defective individuals in purebred dog population than mixed breed dog population.

```r
# Two-sample test for comparing two population proportions
total_rows_pure <- nrow(pure_breed)
Observed_A2_A2_pure <- sum(pure_breed$`Degenerative Myelopathy (DM)` == "A2/A2", na.rm = TRUE)
total_rows_mixed <- nrow(mixed_breed)
Observed_A2_A2_mixed <- sum(mixed_breed$`Degenerative Myelopathy (DM)` == "A2/A2", na.rm = TRUE)
prop.test(c(Observed_A2_A2_pure,Observed_A2_A2_mixed),c(total_rows_pure,total_rows_mixed),
          alternative = "greater", correct = FALSE)
```

```
##
##  2-sample test for equality of proportions without continuity correction
##
## data:  c(Observed_A2_A2_pure, Observed_A2_A2_mixed) out of c(total_rows_pure, total_rows_mixed)
## X-squared = 70.163, df = 1, p-value < 2.2e-16
## alternative hypothesis: greater
```

```
## 95 percent confidence interval:
##  0.00716104 1.00000000
## sample estimates:
##      prop 1      prop 2
## 0.018297534 0.008540131
```

Result:

p-value < 2.2e-16, which is smaller than 0.05. The 95 percent confidence interval is (0.00716,1), which does not include zero.

It is concluded that with 0.05 confidence level, we have enough evidence to reject null hypothesis, so we support the alternative hypothesis that there are more A2/A2 defective individuals in purebred dog population than mixed breed dog population.

Chi-square test:

For our second statistical test, we will conduct a goodness-of-fit test (chi-square test) to check whether these two populations (purebred or mixed breed) are in Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium means that if the individuals in a population follow random breeding without other external influences, after a certain number of generations the proportion of individuals of each genotype in the population reaches a certain equilibrium ratio (Mayo, 2008). A population that follows the Hardy-Weinberg equilibrium shows no signs of evolution. However, if a population deviates from this equilibrium, this indicates evolution due to selection caused by external influences (Winterer, 2021).

A population that follows Hardy-Weinberg equilibrium has the following genotype ratio (these would be our expected values in the chi-squared test) if we consider a single locus with two alleles:

If:

Proportion of A1 in the population = p

Proportion of A2 in the population = q (q = 1-p, because only two different alleles)

Then:

The proportion of A1/A1 individuals = p^2

The proportion of A1/A2 individuals = $2 * p * q$

The proportion of A2/A2 individuals = q^2

In this chi-square test, we will test the adaptability by comparing the observed number of dogs with the expected number of dogs for each of the three genotypes (homozygous dominant (healthy), heterozygous (healthy) and homozygous recessive (unhealthy) in two groups (purebred and mixed).

Our hypothesis is that in purebred dog populations, the observed numbers would differ from the expected numbers, so that evolution in this population occurs through crossbreeding (not random mating) due to stronger artificial selection by dog breeders. (Winterer, 2021) However, in a mixed purebred dog population, the population would be more likely in Hardy-Weinberg equilibrium due to less artificial selection.

In addition, in this specific chi-squared test to check the adaptability of Hardy-Weinberg equilibrium, we use 1 degree of freedom because the calculation starts with one of two alleles even if we have three classes of genotypes (Hosking et al., 2004; Namipashaki et al., 2015; Pappas, 2020).

Statistical Interpretation:

For purebred dog population:

Null hypothesis 1: In purebred dog population, the observed number of dogs for each three genotypes are equal to expected number of dogs for each three genotypes. Namely, purebred dog population is in Hardy-Weinberg equilibrium.

Alternative hypothesis 1: Purebred dog population is not in Hardy-Weinberg equilibrium.

```r
# Chi-square Test
# First, find each genotype counts in each subgroup
Observed_A1_A2_pure <- sum(pure_breed$`Degenerative Myelopathy (DM)` == "A1/A2", na.rm = TRUE)
Observed_A1_A1_pure <- sum(pure_breed$`Degenerative Myelopathy (DM)` == "A1/A1", na.rm = TRUE)
Observed_A1_A2_mixed <- sum(mixed_breed$`Degenerative Myelopathy (DM)` == "A1/A2", na.rm = TRUE)
Observed_A1_A1_mixed <- sum(mixed_breed$`Degenerative Myelopathy (DM)` == "A1/A1", na.rm = TRUE)
# Second, calculate proportion for each loci A1 & A2 in each subgroup
proportion_A1_pure <- (Observed_A1_A2_pure + Observed_A1_A1_pure*2)/(total_rows_pure*2)
proportion_A2_pure <- (Observed_A1_A2_pure + Observed_A2_A2_pure*2)/(total_rows_pure*2)
proportion_A1_mixed <- (Observed_A1_A2_mixed + Observed_A1_A1_mixed*2)/(total_rows_mixed*2)
proportion_A2_mixed <- (Observed_A1_A2_mixed + Observed_A2_A2_mixed*2)/(total_rows_mixed*2)
# Chi-square Test in pure subgroup
Expected_A1_A1_pure = (proportion_A1_pure^2)*total_rows_pure
Expected_A1_A2_pure = (2*proportion_A1_pure*proportion_A2_pure)*total_rows_pure
Expected_A2_A2_pure = (proportion_A2_pure^2)*total_rows_pure
x_pure=c(Observed_A1_A1_pure,Observed_A1_A2_pure,Observed_A2_A2_pure); p_pure=c(Expected_A1_A1_pure,Exp
chisq.test(x_pure, p=p_pure)
```

```
##
##  Chi-squared test for given probabilities
##
## data:  x_pure
## X-squared = 528.66, df = 2, p-value < 2.2e-16
```

Result:

Chi-square = 528.66, which is much larger than the critical value (3.84, df = 1). p-value < 2.2e-16, which is much smaller than 0.05.

Therefore, we have enough evidence to reject the null hypothesis that observed values are same as expected values. This supports the alternative hypothesis that purebred dog population is not in Hardy-Weinberg equilibrium.

For mixed breed dog population:

Null hypothesis 2: In mixed dog population, the observed number of dogs for each three genotypes are equal to expected number of dogs for each three genotypes. Namely, purebred dog population is in Hard-Weinberg equilibrium.

Alternative hypothesis 2: Mixed dog population is not in Hard-Weinberg equilibrium.

```r
# Chi-square Test in mixed subgroup
Expected_A1_A1_mixed = (proportion_A1_mixed^2)*total_rows_mixed
Expected_A1_A2_mixed = (2*proportion_A1_mixed*proportion_A2_mixed)*total_rows_mixed
Expected_A2_A2_mixed = (proportion_A2_mixed^2)*total_rows_mixed
x_mixed=c(Observed_A1_A1_mixed,Observed_A1_A2_mixed,Observed_A2_A2_mixed); p_mixed=c(Expected_A1_A1_mix
chisq.test(x_mixed, p=p_mixed)
```

```
##
##  Chi-squared test for given probabilities
##
## data:  x_mixed
## X-squared = 82.63, df = 2, p-value < 2.2e-16
```

Result:

Chi-square = 82.63, which is larger than the critical value (3.84, df = 1). p-value < 2.2e-16, which is also smaller than 0.05.

Therefore, we have enough evidence to reject the null hypothesis that observed values are same as expected values. This supports the alternative hypothesis that mixed breed dog population is not in Hardy-Weinberg equilibrium.

Conclusion

Based on two-sample proportions two-sample test, it is found that in purebred dog population, a significantly higher proportion of dogs would have homozygous genotype (A2/A2), which leads to Canine Degenerative myelopathy (DM), than mixed breed dog population. The reason for this phenomenon is mostly because inbreeding occurs more often in dog populations which are raised by dog breeders based on artificial selection than mixed breed. Therefore, it is suggested that when people choose pet dogs, if people are more concerned about the healthy condition of pet dogs rather than the external traits (such as appearance) of dogs, they should be more inclined to consider mixed breed dogs.

Based on the chi-square test, it was concluded that in both the purebred dog population and the mixed dog population, the observed number of dogs for the three different genotypes did not match the expected values, indicating that both the purebred and mixed dog populations were not in Hardy-Weinberg equilibrium. Consequently, signs of evolution could be seen in both populations.

Although both populations showed signs of evolution, the chi-square value in the purebred dog population (528.66) was much higher than that in the mixed dog population (82.63), indicating that evolution was much stronger in the purebred dog population than in the mixed dog population. The driving force of evolution in purebred dog populations is mainly due to more common inbreeding. However, the driving force of evolution in mixed-breed dog populations is still unclear, and further studies may be needed to find out why.

In summary, it can be concluded that even without strong artificial selection, evolution is taking place in mixed-breed dog populations as well. Artificial selection by dog breeders could accelerate this process, resulting in future generations of offspring having a greater chance of having homozygous defective genotypes (A2/A2).

Therefore, we recommend dog breeders to perform some relevant genetic testing on dogs before mating to prevent dogs with defective genes from mating with each other or even inbreeding. This can effectively reduce the occurrence of more homozygous individuals with diseases (A2/A2) in future offspring and ensure that the dogs complement each other optimally or produce healthier offspring that reduce the overall number and risk of dogs with canine degenerative myelopathy.

In addition, dog owners should invest in testing before purchasing a dog to avoid unexpected and costly treatments for genetic diseases in the future. Raising awareness of the higher incidence of the disease in purebreds can help both the owner and the veterinarian detect the disease earlier. Comprehensive screening of genetic diseases and possible early therapies are essential to improve the quality of life of animals and minimize the financial burden on owners. Furthermore, it could be useful to adjust dog insurance premiums to the risks of genetic diseases in order to incentivize more responsible breeding.

Nevertheless, it is important not to neglect the mixed breeds in the testing, as an evolution of the disease genes could also be determined. Again, early signs of owners and vets should not be neglected just because the probability of the mixed breeds getting the disease seems to be lower. This must also be seen in the context of the biases mentioned above, which indicate that these biases may also have led to the increased probability of disease genotypes in pure breeds.

Reference

Björnerfeldt, S., Hailer, F., Nord, M., & Vilà, C. (2008). Assortative mating and fragmentation within dog breeds. BMC Evolutionary Biology, 8, 1-11. https://doi.org/10.1186/1471-2148-8-28

Brand, C. L., , Belshaw, Z., Bryson, G. T., O'Neill, D. G., & Packer, R. M. (2024). The doodle dilemma: How the physical health of 'designer-crossbreed' cockapoo, labradoodle and Cavapoo Dogs' compares to their purebred progenitor breeds. PLOS ONE, 19(8). https://doi.org/10.1371/journal.pone.0306350

Donner, J., Anderson, H., Davison, S., Hughes, A. M., Bouirmane, J., Lindqvist, J., ... & Lohi, H. (2019). Frequency and distribution of 152 genetic disease variants in over 100,000 mixed breed and purebred dogs. PLoS genetics, 14(4), e1007361. https://doi.org/10.1371/journal.pgen.1007361

Claussnitzer, M., Cho, J. H., Collins, R., Cox, N. J., Dermitzakis, E. T., Hurles, M. E., ... & McCarthy, M. I. (2020). A brief history of human disease genetics. Nature, 577(7789), 179-189. https://doi.org/10.1038/s41586-019-1879-7

Coates, J. R., & Wininger, F. A. (2010). Canine degenerative myelopathy. Veterinary Clinics: Small Animal Practice, 40(5), 929-950. doi: 10.1016/j.cvsm.2010.05.001

Cohen, J. (1988). Statistical Power Analysis for the behavioral sciences. [Book]. L. Erlbaum Associates.

Cornell University. (2024). Degenerative myelopathy. Cornell Richard P. Riney Canine Health Center, Cornell University. https://www.vet.cornell.edu/departments-centers-and-institutes/riney-canine-health-center/canine-health-information/degenerative-myelopathy#:~:text=Overview,lateral%20sclerosis)%2C%20in%20humans

Farrell, L. L., Schoenebeck, J. J., Wiener, P., Clements, D. N., & Summers, K. M. (2015). The challenges of Pedigree Dog Health: Approaches to combating inherited disease. Canine Genetics and Epidemiology, 2(1). https://doi.org/10.1186/s40575-015-0014-9

Forsyth, K. K., McCoy, B. M., Schmid, S. M., Promislow, D. E., Snyder-Mackler, N., DAP Consortium, ... & Creevy, K. E. (2023). Lifetime prevalence of owner-reported medical conditions in the 25 most common dog breeds in the Dog Aging Project pack. Frontiers in Veterinary Science, 10, 1140417. doi: 10.3389/fvets.2023.1140417

Grand View Research. (2024). (rep.). Pet Insurance Market Size, Share & Trends Analysis Report By Coverage Type (Accident-only, Accident & Illness), By Animal Type (Dogs, Cats), By Sales Channel (Agency, Broker, Direct), By Region, And Segment Forecasts, 2024 - 2030 (pp. 1–125). San Francisco, California .

Hosking, L., Lumsden, S., Lewis, K., Yeo, A., McCarthy, L., Bansal, A., ... & Xu, C. F. (2004). Detection of genotyping errors by Hardy–Weinberg equilibrium testing. European Journal of Human Genetics, 12(5), 395-399. https://doi.org/10.1038/sj.ejhg.5201164

Ivansson, E. L., Megquier, K., Kozyrev, S. V., Murén, E., Körberg, I. B., Swofford, R., ... & Lindblad-Toh, K. (2016). Variants within the SP110 nuclear body protein modify risk of canine degenerative myelopathy. Proceedings of the National Academy of Sciences, 113(22), E3091-E3100. https://doi.org/10.1073/pnas.1600084113

Kobatake, Y., Nakata, K., Sakai, H., Sasaki, J., Yamato, O., Takashima, S., ... & Kamishina, H. (2021). The long-term clinical course of canine degenerative myelopathy and therapeutic potential of curcumin. Veterinary Sciences, 8(9), 192. https://doi.org/10.3390/vetsci8090192

Mayo, O. (2008). A century of Hardy–Weinberg equilibrium. Twin Research and Human Genetics, 11(3), 249-256. https://doi.org/10.1375/twin.11.3.249

Mellanby, R. J., Ogden, R., Clements, D. N., French, A. T., Gow, A. G., Powell, R., ... & Summers, K. M. (2013). Population structure and genetic heterogeneity in popular dog breeds in the UK. The Veterinary Journal, 196(1), 92-97. https://doi.org/10.1016/j.tvjl.2012.08.009

Jansson, M., & Laikre, L. (2014). Recent breeding history of dog breeds in S weden: modest rates of inbreeding, extensive loss of genetic diversity and lack of correlation between inbreeding and health. Journal of Animal Breeding and Genetics, 131(2), 153-162. https://doi.org/10.1111/jbg.12060

Namipashaki, A., Razaghi-Moghadam, Z., & Ansari-Pour, N. (2015). The essentiality of reporting Hardy-Weinberg equilibrium calculations in population-based genetic association studies. Cell Journal (Yakhteh), 17(2), 187. doi: 10.22074/cellj.2016.3711

Neeves, J., & Granger, N. (2015). An update on degenerative myelopathy in dogs. Companion Animal, 20(7), 408-412. https://doi.org/10.12968/coan.2015.20.7.408

Pappas, J., (2020). How to use chi-square to test for Hardy-Weinberg equilibrium. Biology Simulations LLC. How to use chi-square to test for Hardy-Weinberg equilibrium (biologysimulations.com)

Shahbandeh, M. (2024, July 1). Dog ownership in the EU by country 2023. Statista. https://www.statista.com/statistics/515475/dog-ownership-european-union-eu-by-country/

Sutter, N. B., Eberle, M. A., Parker, H. G., Pullar, B. J., Kirkness, E. F., Kruglyak, L., & Ostrander, E. A. (2004). Extensive and breed-specific linkage disequilibrium in Canis familiaris. Genome research, 14(12), 2388-2396. http://www.genome.org/cgi/doi/10.1101/gr.3147604

Winterer, J. (2001). A lab exercise explaining Hardy-Weinberg equilibrium and evolution effectively. The American Biology Teacher, 63(9), 678-687. https://doi.org/10.1662/0002-7685(2001)063%5B0678:ALEEHE%5D2.0.CO;2