

# Big Data and kennel clubs

---

IPFD workshop, 2024



---

“Data really powers everything we do”

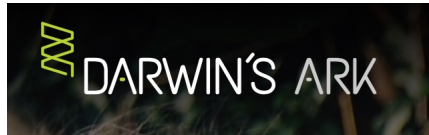
*Jeff Weiner, LinkedIn's CEO*

---

“Data really powers everything we do”

*Jeff Weiner, LinkedIn’s CEO*

**12TH INTERNATIONAL CONFERENCE OF CANINE AND FELINE GENETICS AND GENOMICS 2024**



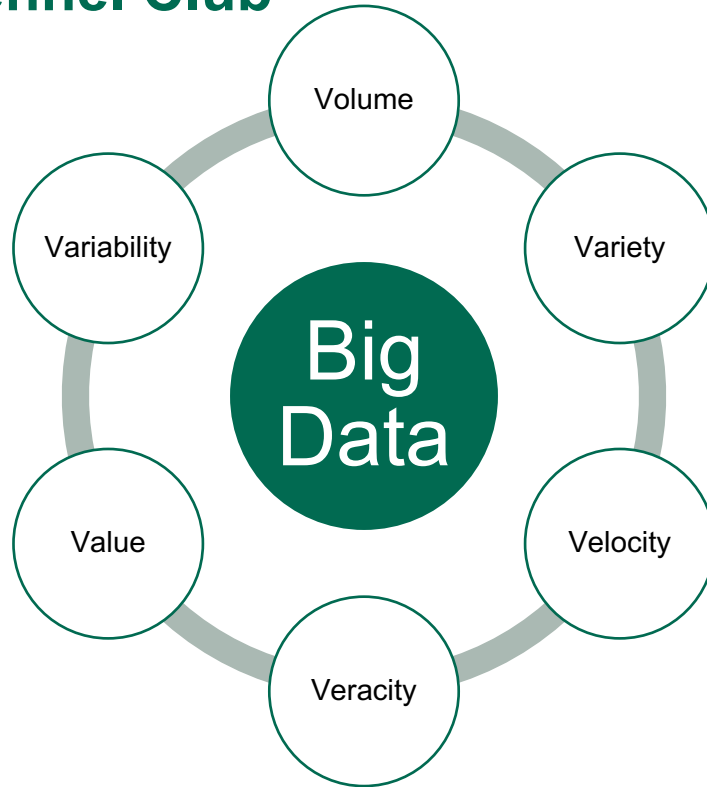
# Big Data at The Kennel Club

12M dogs

1.3M DNA test results

Close to 1M phenotypic test results

200-250K yearly registrations



# Value of Big Data at The KC

## 1. Hundreds/thousands of scientific papers:

- Genetics
- Veterinary medicine
- Epidemiology
- Behaviour
- Sociological

Copyright © 2008 by the Genetics Society of America  
DOI: 10.1534/genetics.107.084054

### Population Structure and Inbreeding From Pedigree Analysis of Purebred Dogs

Federico C. F. Calboli,<sup>\*,1</sup> Jeff Sampson,<sup>1</sup> Neale Fretwell<sup>2</sup> and David J. Balding<sup>\*</sup>

© 2010 Universities Federation for Animal Welfare  
The Old School, Brewhouse Hill, Wheathampstead,  
Hertfordshire AL4 8AN, UK

Animal Welfare 2010, 19(5): 93-98  
ISSN 0962-7286

### Optimisation of breeding strategies to reduce the prevalence of inherited disease in pedigree dogs

TW Lewis<sup>1</sup>, JA Woolliams<sup>1</sup> and SC Blott<sup>2\*</sup>

Lewis et al. Canine Genetics and Epidemiology (2015) 2:13  
DOI 10.1186/s40275-015-0027-4

Canine Genetics and Epidemiology

RESEARCH

Open Access

Trends in genetic diversity for all Kennel Club registered pedigree dog breeds

T. W. Lewis<sup>1,2\*</sup>, B.M. Abhayaratne<sup>2</sup> and S. C. Blott<sup>2</sup>

Lewis Canine Genetics and Epidemiology  
https://doi.org/10.1186/s40275-019-0014-3 02019-6-6

Canine Genetics and Epidemiology

COMMENTARY

Open Access

The impact of incorrectly recorded parentage on inferred genotypes over multiple generations

T. W. Lewis<sup>1,2\*</sup>

### Proportion of litters of purebred dogs born by caesarean section

KATY M. EVANS AND VICKI J. ADAMS

Journal of Small Animal Practice (2010)  
51, 113–118  
DOI: 10.1111/j.1748-5827.2009.00902.x

### Genetic Characterization of Dog Personality Traits

Joanna Ilska,<sup>\*</sup> Marie J. Haskell,<sup>\*</sup> Sarah C. Blott,<sup>\*</sup> Enrique Sánchez-Molano,<sup>1</sup> Zita Polgar,<sup>1</sup> Sarah E. Lofgren,<sup>1</sup> Dylan N. Clements,<sup>1</sup> and Pamela Wiener<sup>1,1</sup>

Genetics, Vol. 206, 1101-1111 June 2017

Wies et al. Canine Genetics and Epidemiology (2017) 4:8  
DOI 10.1186/s40275-017-0047-5

Canine Genetics and Epidemiology

RESEARCH

Open Access

Large-scale survey to estimate the prevalence of disorders for 192 Kennel Club registered breeds

B. M. Wies<sup>1\*</sup>, A. M. Llewellyn-Zaid<sup>2</sup>, K.M. Evans<sup>1,3</sup>, D. G. O'Neill<sup>3</sup> and T. W. Lewis<sup>1,3</sup>

PLOS ONE

RESEARCH ARTICLE

Changes in mutation frequency of eight Mendelian inherited disorders in eight pedigree dog populations following introduction of a commercial DNA test

T. W. Lewis<sup>1,2\*</sup>, C. S. Mellersh<sup>3</sup>

frontiers  
in Veterinary Science

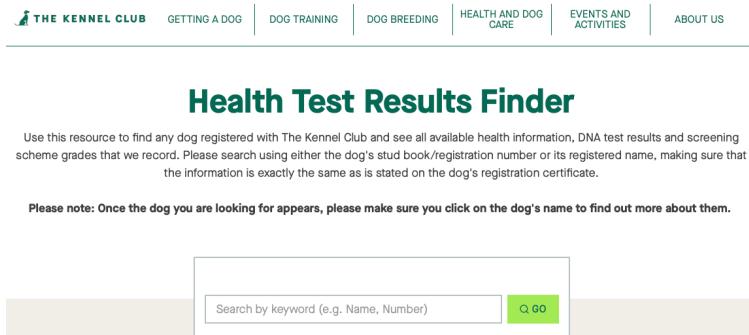
ORIGINAL RESEARCH  
published: 15 January 2020  
doi: 10.3389/fvets.2019.00490

### Effectiveness of Canine Hip Dysplasia and Elbow Dysplasia Improvement Programs in Six UK Pedigree Breeds

H. K. James<sup>1</sup>, F. McDonnell<sup>1</sup> and Thomas W. Lewis<sup>1,2\*</sup>

# Value of Big Data at The KC

1. Hundreds of scientific papers
2. Development of breeding tools:
  - Health Test Result Finder



THE KENNEL CLUB GETTING A DOG DOG TRAINING DOG BREEDING HEALTH AND DOG CARE EVENTS AND ACTIVITIES ABOUT US

## Health Test Results Finder

Use this resource to find any dog registered with The Kennel Club and see all available health information, DNA test results and screening scheme grades that we record. Please search using either the dog's stud book/registration number or its registered name, making sure that the information is exactly the same as is stated on the dog's registration certificate.

Please note: Once the dog you are looking for appears, please make sure you click on the dog's name to find out more about them.

Search by keyword (e.g. Name, Number)



RETRIEVER (LABRADOR)

**SH CH CREMINO CHRISTMAS ROSE JW**

♀ Bitch · YELLOW · Born 20 December 2005

## Health profile

The results and calculated health information below are from information received and recorded by The Kennel Club, and may not include all health screening undertaken by the dog's owners.

You can find more information on [what these results mean/breeding advice](#) and also on [what health screening is relevant to your breed](#).

## DNA tests

### DNA - prcd-PRA

Clear

Test performed on 28 June 2007; aged 1 years, 6 months

### DNA - CNM

Clear

Test performed on 09 July 2008; aged 2 years, 6 months

## Screening schemes

BVA/KC/ISDS Eye Scheme

Unaffected

Test performed on 01 October 2008; aged 2 years, 9 months

Unaffected

Test performed on 01 October 2009; aged 3 years, 9 months

### BVA/KC Hip Dysplasia

Left score: 3

Right score: 3

Total score: 6

Test performed on 16 February 2007; aged 1 years, 1 months

### BVA/KC Elbow Dysplasia

Left score: 0

Right score: 0

Total score: 0

Test performed on 16 February 2007; aged 1 years, 1 months

# Value of Big Data at The KC

1. Hundreds of scientific papers
2. Development of breeding tools:
  - Health Test Result Finder
  - MateSelect/COI calculators

## Inbreeding Coefficient (COI) lookup Results

☆ MOST DIVERSE MATCH

Inbreeding coefficient for DINASTII THE PHENOMENAL and JAYBIEM SOOKA is 7.7%.

The breed average is 4.1%  
12 generations available of which 5 are complete

Inbreeding coefficient for DINASTII THE PHENOMENAL and CLANHEIR ME AND MY SHADOW is 21.3%.

The breed average is 4.1%  
10 generations available of which 5 are complete



### This dog is not eligible to mate

DINASTII THE PHENOMENAL is unable to mate with NATASHA WITH LOVE IZ RUSSKOI DINASTII (IMP RUS) because they are related.

Janes et al. *Canine Medicine and Genetics* (2020) 7:14  
<https://doi.org/10.1186/s40575-020-00094-8>

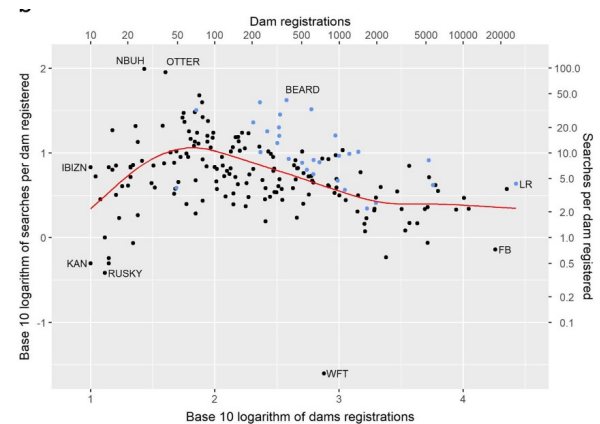
Canine Medicine and Genetics

RESEARCH

Open Access

## The usage of Mate Select, a web-based selection tool for pedigree dogs for promoting sustainable breeding

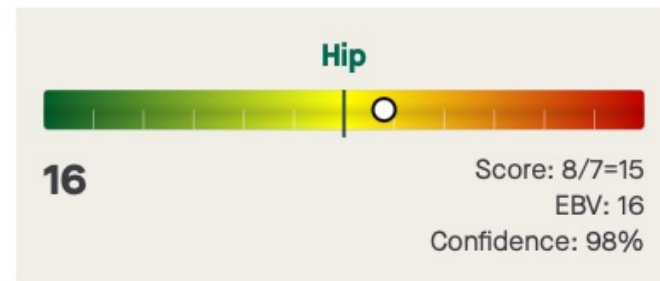
Mateja Janes<sup>1\*</sup>, Thomas W. Lewis<sup>2,3</sup>, Joanna J. Iliska<sup>1†</sup> and John A. Woodlams<sup>1†</sup>



# Value of Big Data at The KC

---

1. Hundreds of scientific papers
2. Development of breeding tools:
  - Health Test Result Finder
  - MateSelect/COI calculators
  - EBVs



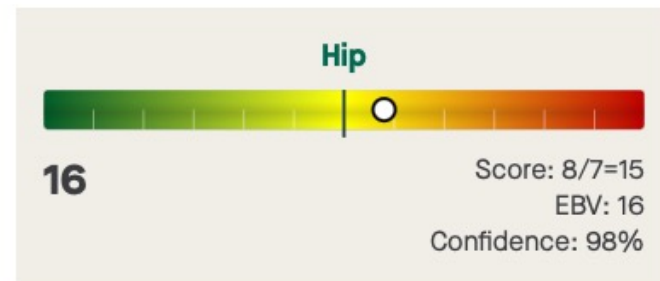
EBV results last updated 01 June 2024.

# Value of Big Data at The KC

1. Hundreds of scientific papers
2. Development of breeding tools:

- Health Test Result Finder
- MateSelect/COI calculators
- EBVs

Impossible in absence of large data quantities

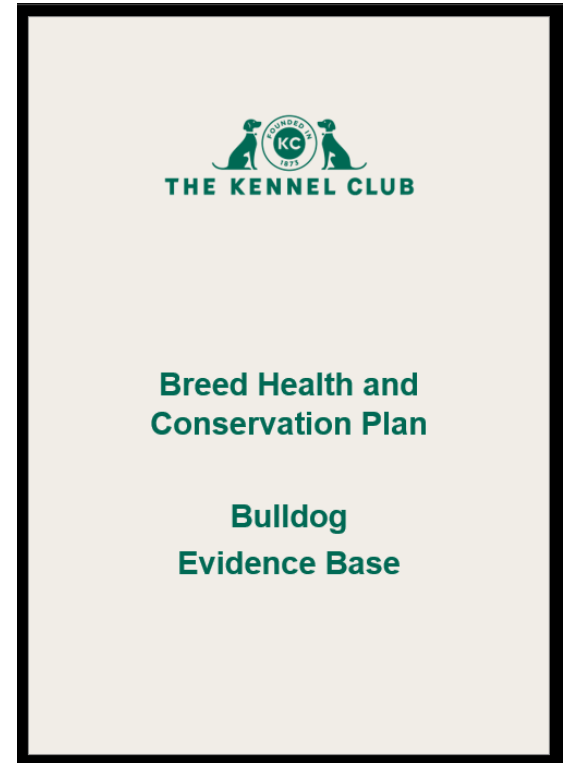


EBV results last updated 01 June 2024.

# Value of Big Data at The KC

---

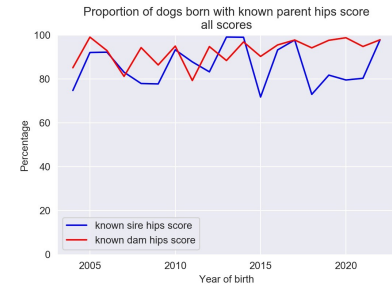
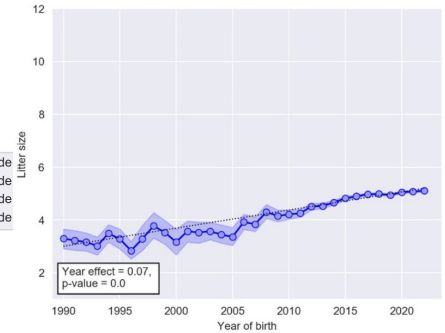
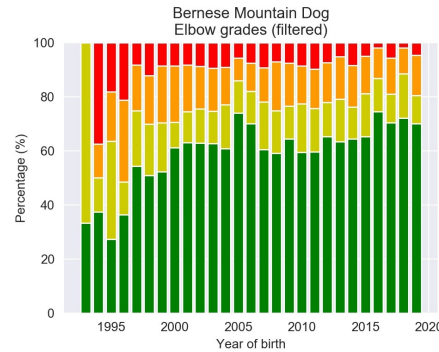
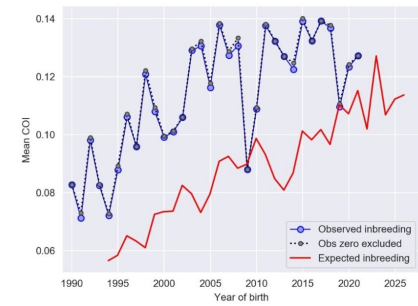
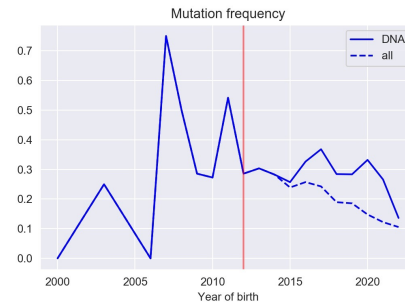
1. Hundreds of scientific papers
2. Development of breeding tools
3. **Breed-centric information**



# Breed-centric information

1. Providing individual breeders with information necessary to manage the breed population:

1. Breed average COI
2. Mutation frequencies
3. Prevalence data for complex diseases
4. Rates of inbreeding
5. Changes in population size
6. Changes in litter size



# Future opportunities

---

1. Further development of breed-specific management plans for health and genetic diversity
2. Expansion on the traits recorded – health, behaviour
3. Social aspects – who are breeders and what motivates them?
4. New methods of data analysis (AI)

# Future opportunities

---

## 1. Genomics:

1. Improved estimates of inbreeding/genetic diversity
2. Improved accuracy of (Genomic) Estimated Breeding Values
3. Retaining attractiveness in scientific collaborations
4. New tools developed by scientists are based on genomic data

# Big Data at The Kennel Club

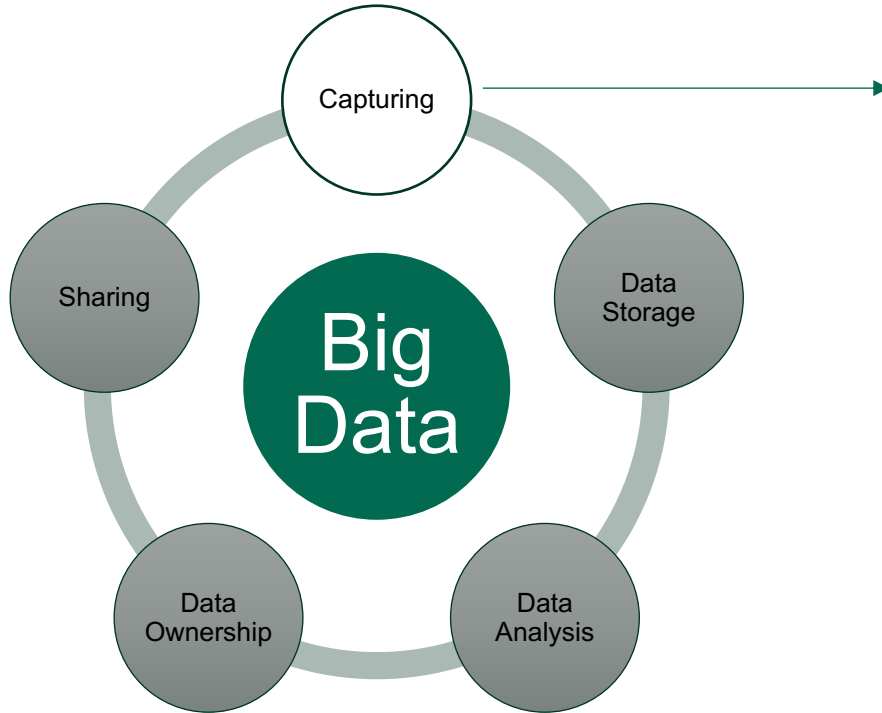
## - challenges

---



# Big Data at The Kennel Club - challenges

---



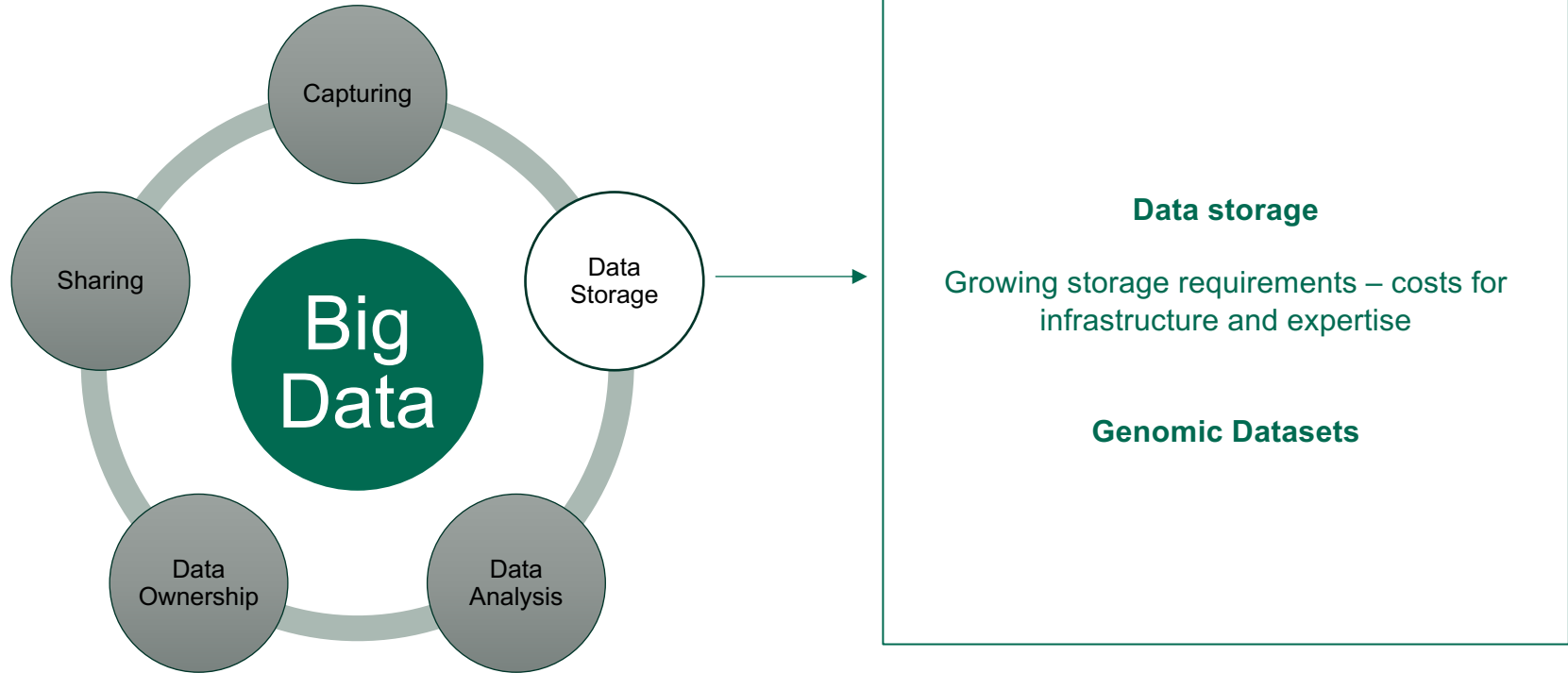
## Capturing

Methods for extracting all available information

Continued interaction with stakeholders

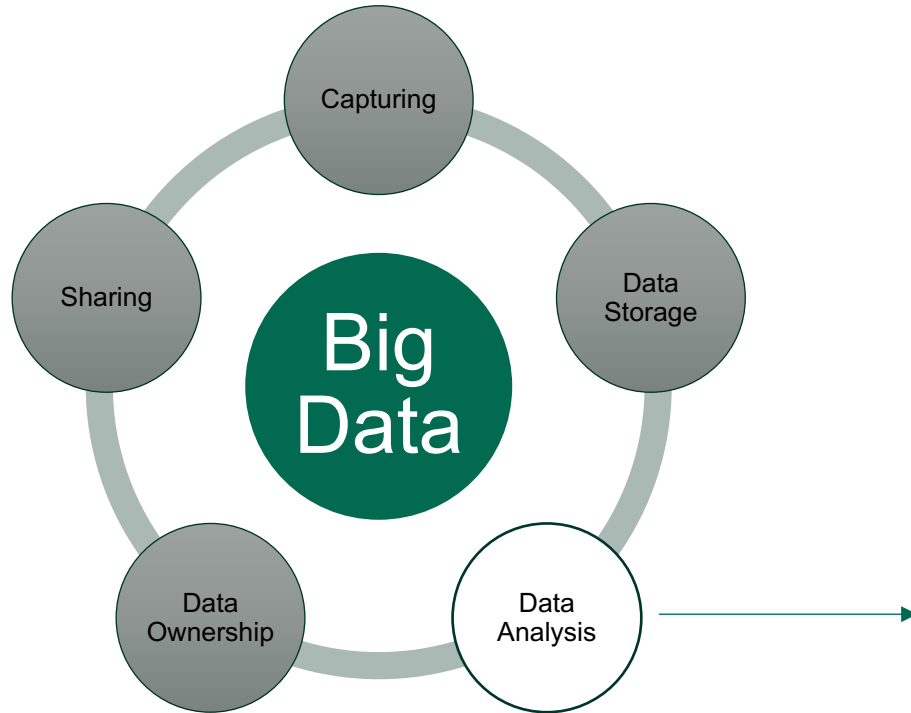
# Big Data at The Kennel Club - challenges

---



# Big Data at The Kennel Club - challenges

---



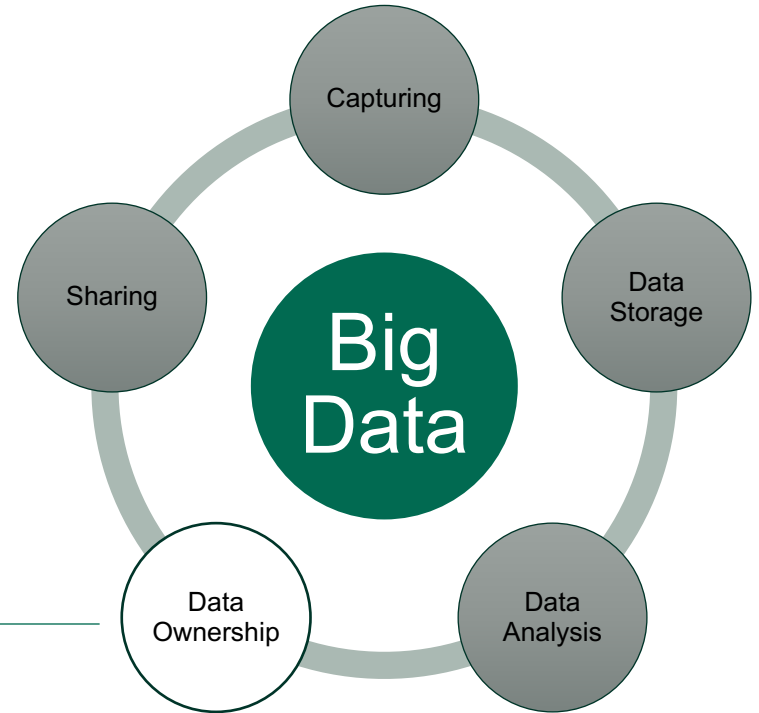
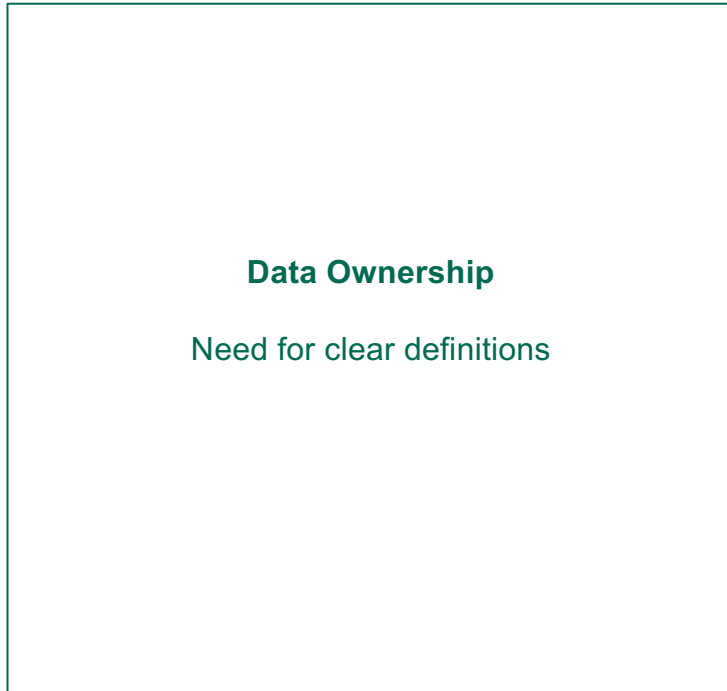
## Data Analysis

Growing costs for infrastructure and expertise  
(programming skills)

## Genomic Analyses

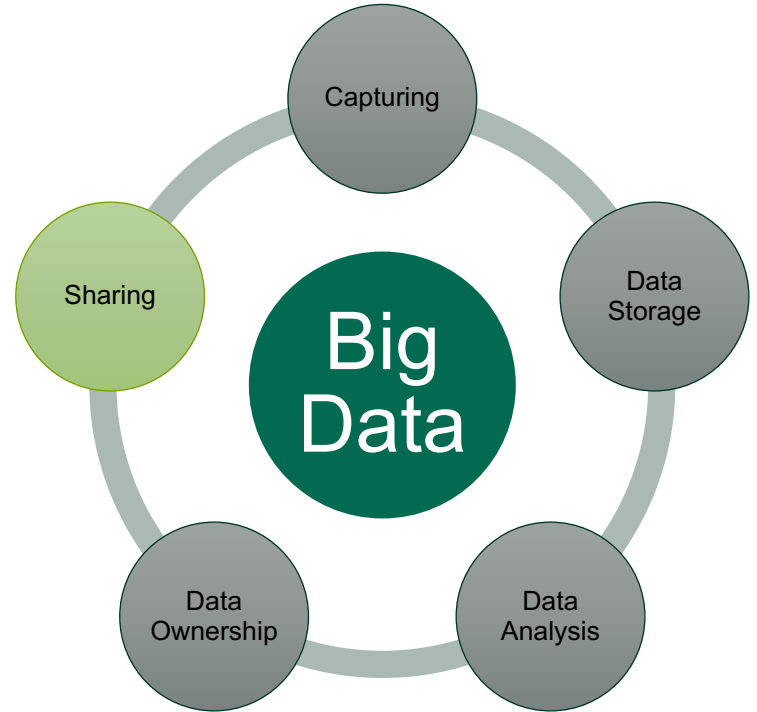
# Big Data at The Kennel Club - challenges

---



# Big Data at The Kennel Club - challenges

---



# Sharing of Big Data between kennel clubs

The Veterinary Journal 189 (2011) 189–196



Contents lists available at ScienceDirect

The Veterinary Journal

journal homepage: [www.elsevier.com/locate/tvjl](http://www.elsevier.com/locate/tvjl)



Review

## International and collaborative strategies to enhance genetic health in purebred dogs

Åke A. Hedhammar<sup>a,b,\*</sup>, Sofia Malm<sup>b,c</sup>, Brenda Bonnett<sup>d</sup>

<sup>a</sup> Department of Clinical Sciences, Swedish University of Agricultural Sciences, PO Box 7052, 750 07 Uppsala, Sweden

<sup>b</sup> Swedish Kennel Club, SE-163 85 Spånga, Sweden

<sup>c</sup> Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, PO Box 7023, 750 07 Uppsala, Sweden

<sup>d</sup> B Bonnett Consulting, RR2 Warton, Ontario, Canada, N0H 2T0

The Veterinary Journal 197 (2013) 873–875



Contents lists available at ScienceDirect

The Veterinary Journal

journal homepage: [www.elsevier.com/locate/tvjl](http://www.elsevier.com/locate/tvjl)



Short Communication

## Opportunities for international collaboration in dog breeding from the sharing of pedigree and health data

W.F. Fikse<sup>a,\*</sup>, S. Malm<sup>b</sup>, T.W. Lewis<sup>c</sup>



Mamm Genome (2012) 23:195–202

DOI 10.1007/s00335-011-9366-y

J. Anim. Breed. Genet. ISSN 0931-2668

REVIEW

ORIGINAL ARTICLE

## Merging pedigree databases to describe and compare mating practices and gene flow between pedigree dogs in France, Sweden and the UK

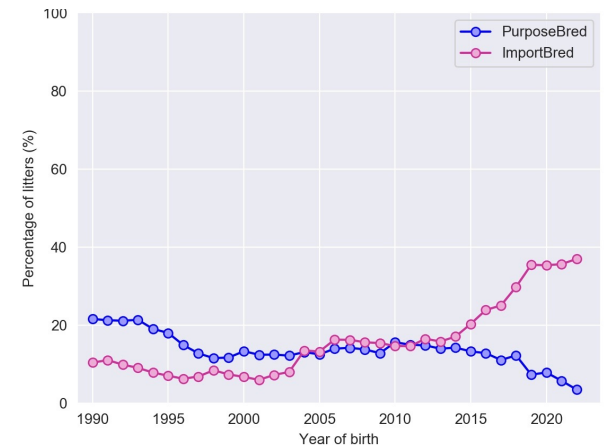
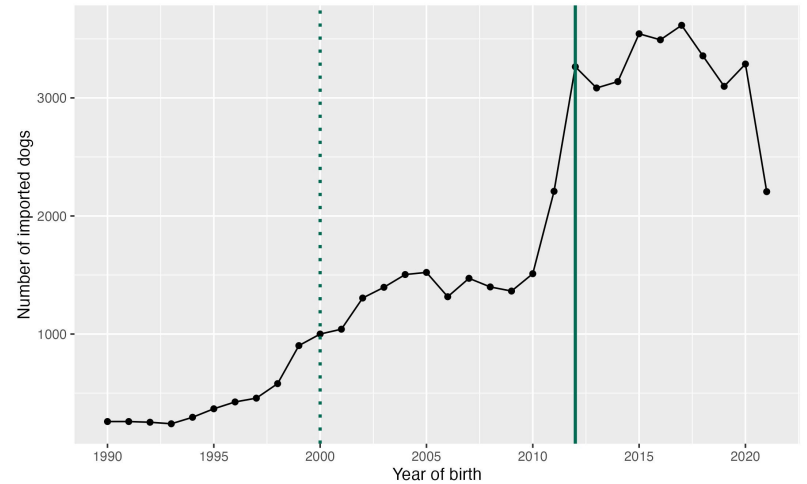
S. Wang<sup>1,2,3</sup>, G. Leroy<sup>2,3</sup>, S. Malm<sup>4</sup>, T. Lewis<sup>5,6</sup>, E. Strandberg<sup>1</sup> & W.F. Fikse<sup>1</sup>

## Empowering international canine inherited disorder management

Bethany J. Wilson · Claire M. Wade

# Sharing of Big Data between kennel clubs

1. Increasing number of imports (UK)
2. Increasing use of imports in breeding

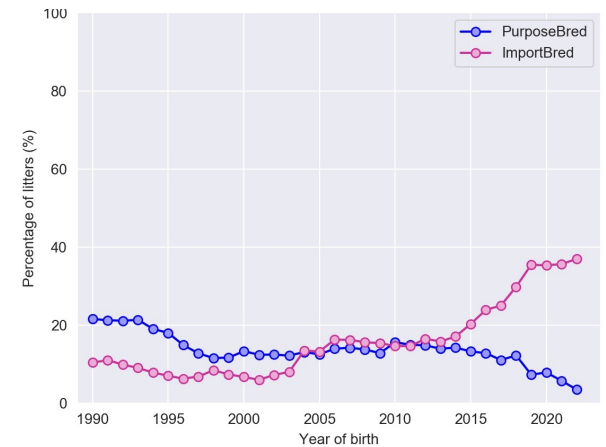
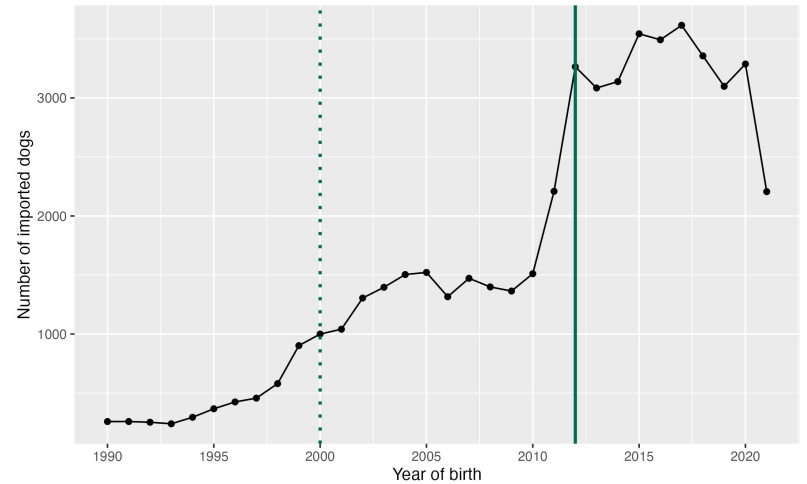


# Sharing of Big Data between kennel clubs

1. Increasing number of imports (UK)
2. Increasing use of imports in breeding



Biased estimates of inbreeding and genetic diversity



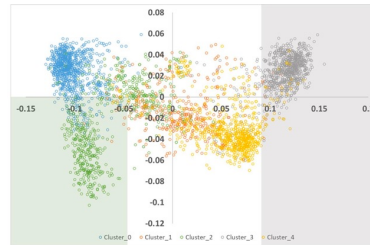
# Sharing of Big Data between kennel clubs

1. Increasing number of imports (UK)
2. Increasing use of imports in breeding

Biased estimates of inbreeding and genetic diversity

## Large breeds – smaller impact (breed level)

1. Smaller overall proportion of import-bred litters
2. Genetic differentiation between countries



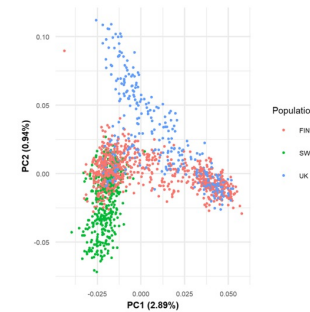
Wade et al. *Canine Medicine and Genetics* (2023) 10:7  
<https://doi.org/10.1186/s40795-023-00130-3>

Canine Medicine and Genetics

RESEARCH Open Access

Comprehensive analysis of geographic and breed-purpose influences on genetic diversity and inherited disease risk in the Doberman dog breed

Claine M. Wade<sup>1</sup>, Robin Nuttall<sup>2</sup> and Sophie Liu<sup>3</sup>



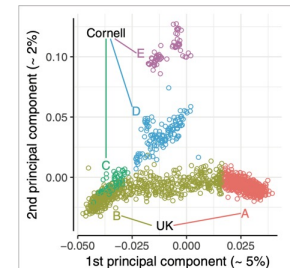
Wang et al. *BMC Genomics* (2021) 22:268  
<https://doi.org/10.1186/s12864-021-07940-z>

BMC Genomics frontiers in Genetics

RESEARCH Open Access

Genome-wide association studies for canine hip dysplasia in single and multiple populations – implications and potential novel risk loci

Shaoh Wang<sup>1</sup>, Eding Stavenberg<sup>2</sup>, Per Arvola<sup>3</sup>, Dylan N. Clements<sup>4</sup>, Pamela Werner<sup>5</sup> and Juliane Friedrich<sup>6\*</sup>



ORIGINAL RESEARCH  
published: 20 June 2024  
doi: 10.3389/gen.2024.10710

Joint Genomic Prediction of Canine Hip Dysplasia in UK and US Labrador Retrievers

Shahin M. Elmehrik<sup>1</sup>, John A. Woodhouse<sup>1</sup>, John M. Hickey<sup>1</sup>, Sarah C. Blott<sup>1</sup>, Dylan N. Clements<sup>1</sup>, Enrique Sánchez-Molano<sup>1</sup>, Rory J. Tocherin<sup>1</sup> and Pamela Werner<sup>1\*</sup>

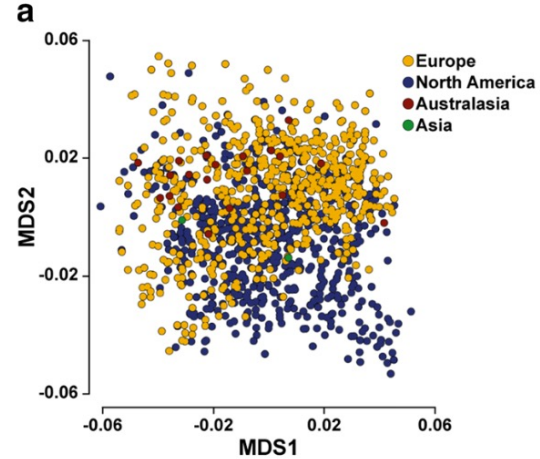
GSD

# Sharing of Big Data between kennel clubs

1. Increasing number of imports (UK)
2. Increasing use of imports in breeding



Biased estimates of inbreeding and genetic diversity



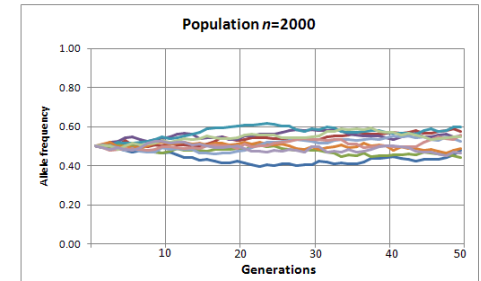
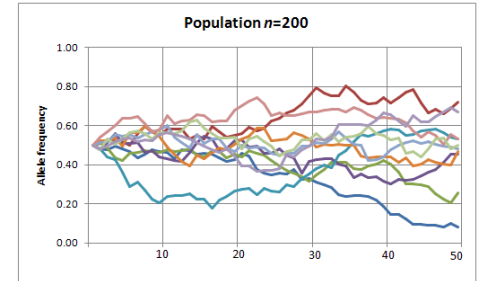
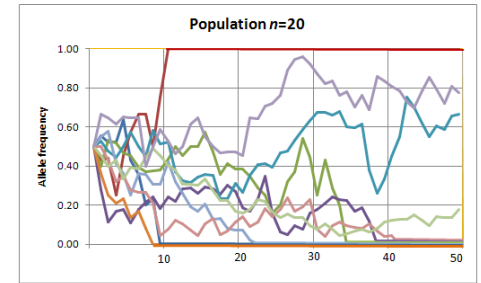
## Small breeds – large impact (breed level)

1. Significant proportion of import-bred litters
2. Small to none genetic differentiation between countries

# Small Data - challenges

## 1. Small Breeds:

1. Typically higher contributions from imports – biased estimates of inbreeding and genetic diversity
2. Small populations – more drastic changes, may need more immediate solutions



By Professor marginalia - Own work, CC BY-SA 3.0,  
<https://commons.wikimedia.org/w/index.php?curid=6052806>

# Small Data - challenges

---

## 1. Small Breeds:

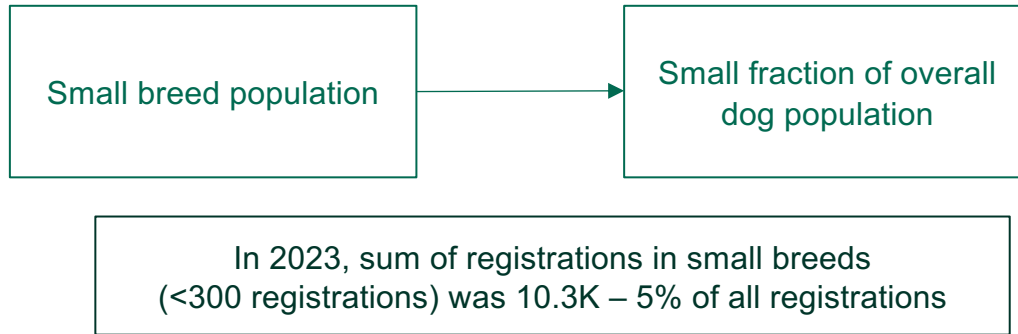
1. Typically higher contributions from imports – biased estimates of inbreeding and genetic diversity
2. Small populations – more drastic changes, may need more immediate solutions
3. Small populations – insufficient data to calculate EBVs for complex traits

# Small Data - challenges

---

## 1. Small Breeds:

1. Small populations – more drastic changes, may need more immediate solutions
2. Small populations – insufficient data to calculate EBVs for complex traits
3. Typically higher contributions from imports – biased estimates of inbreeding and genetic diversity

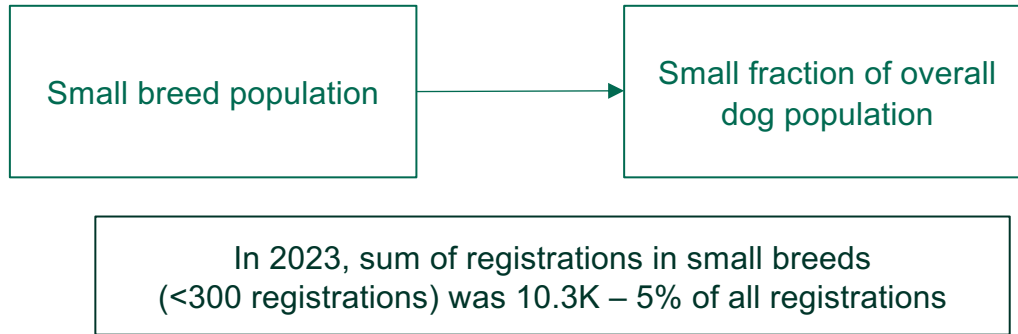


# Small Data - challenges



## 1. Small Breeds:

1. Small populations – more drastic changes, may need more immediate solutions
2. Small populations – insufficient data to calculate EBVs for complex traits
3. Typically higher contributions from imports – biased estimates of inbreeding and genetic diversity



In 2023  
144/222 (65%) breeds  
had <300 registrations

# Small Data - solutions

---

1. International registries?
2. Who would be responsible?
3. Who owns the data?
4. How is the data combined? – e.g. hip dysplasia BVA/KC, OFA or FCI?

## The Irish Wolfhound Database



[Brid Of Ballytobin](#)

11 March 1968

KCREG47281/70

6 litters 40 offspring.

Lifespan:  
10 years

Cause of death: Old age (Lived her whole life with a handicap from a broken leg whilst a puppy)

[Breeder: A. G. Ryan](#)

# Conclusions

---



# Conclusions

---

- Data held by national kennel clubs is Big Data
- Big Data offers endless opportunities
- Realisation of those opportunities requires investment in data analysis
- Collaboration between kennel clubs needed, but will be most challenging
- Small breeds most in need of collaboration



**THE KENNEL CLUB**