Chronology of Hip Dysplasia Development in a Cohort of 48 Labrador Retrievers Followed for Life

Gail K. Smith1, VMD, PhD, Dennis F. Lawler2, DVM, Darryl N. Biery1, DVM, Diplomate ACVR, Michelle Y. Powers1, DVM, Diplomate ACVS, Frances Shofer1, PhD, Thomas P. Gregor1, BS, Georgia T. Karbe1, med.vet., Mischa B. McDonald-Lynch1, DVM, Richard H. Evans3, DVM, MS, and Richard D. Kealy2, PhD

1Department of Clinical Studies, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA, 2Nestle Research Center, St. Louis, MO and 3Veterinary Pathology Services, Aliso Viejo, CA

Corresponding Author
Gail Smith, VMD, PhD, Department of Clinical Studies, School of Veterinary Medicine, University of Pennsylvania, 3900 Delancey St, Philadelphia, PA 19104-6010
E-mail: smithgk@vet.upenn.edu
Submitted June 2010
Accepted October 2010
DOI:10.1111/j.1532-950X.2011.00935.x

Objective: To determine the chronology of radiographic signs of canine hip dysplasia (CHD), specifically joint laxity and secondary osteoarthritis (OA).

Study Design: Longitudinal cohort study.

Animals: Paired littermates, 48 Labrador retrievers.

Methods: Conventional, ventrodorsal, hip-extended (HE) radiographs were evaluated yearly for CHD according to the subjective criteria of the Orthopedic Foundation for Animals (OFA). PennHIP screening was performed at 2 years of age to assess joint laxity by distraction index (DI). Histopathologic evaluation of coxofemoral joints was performed at the dogs’ natural end of life.

Results: Coxofemoral subluxation, as identified on the HE radiograph occurred by 2 years of age and not thereafter. Accuracy of OFA-criteria scoring was poor: 55% of dogs scored “normal” at 2 years of age became radiographically dysplastic by the end of life (45% negative predictive value, NPV); 92% of the dogs scored as normal at 2 years of age had histopathologic OA of CHD (8% NPV). The DI predicted all 48 dogs to be susceptible to OA of CHD and 98% had radiographic or histopathologic OA by the end of life.

Conclusion: OFA-criteria score was profoundly influenced by environmental factors, such as diet restriction and age, reducing its value as a selection criterion. DI measurements were not influenced by dietary treatment suggesting higher trait heritability.

Canine hip dysplasia (CHD) was first described by Schnelle in the mid-1930s1 and is currently recognized as the most common orthopedic disease of large and giant-breed dogs. Hip dysplasia is a complex disease reflecting a combination of genetic weaknesses and environmental stresses that can promote progressive remodeling and osteoarthritis (OA) of the hip joint.2

Age, conformational characteristics, and environmental stresses such as diet have been reported to have profound effects on the expression of OA in dogs that are susceptible to CHD.3–8 It has been shown that keeping dogs lean can delay or minimize the progression and severity of OA.4, 5, 9–11 It has also been shown that dogs with genetic susceptibility to CHD may not express the OA of CHD until late in life, or perhaps not at all, depending on the cumulative effect of environmental influences.11

Various schemes for scoring severity of radiographic changes associated with CHD have been used but all methods incorporate 2 key features: joint laxity and secondary OA, also known as degenerative joint disease. The oldest and most commonly applied diagnostic/screening schemes for CHD worldwide use the ventrodorsal, hip-extended (HE) pelvic radiograph, from which evidence of radiographic subluxation (laxity) of the coxofemoral joint or OA is scored, either subjectively or objectively. In the United States, the Orthopedic Foundation for Animals (OFA; St. Louis, MO) scores the HE radiograph of dogs ≥2 years of age, according to a 7-point subjective evaluation scheme. In much of Europe a slightly different subjective analysis is made from the same radiographic projection at or after 1 year of age depending on the breed of dog. These hip-screening systems assume that a dog with normal hips at 1 or 2 years of age will remain normal for life and is therefore considered suitable for breeding purposes. There is no requirement or recommendation for subsequent
radiographic hip evaluations later in life. The possible change in these subjective hip scores and in the PennHIP distraction index (DI) has not been tested in a fixed cohort of dogs beyond 5 and 3 years of age, respectively.12,13 Despite the worldwide efforts at CHD control, spanning more than 40 years, there has been unconvincing progress using mass selection in reducing the incidence of CHD.14–17

The poor genetic progress suggests that current hip-screening methods have diagnostic and prognostic deficiencies. These include (1) masking of inherent hip laxity by positioning the hips in extension for the standard ventrodorsal, radiographic projection, (2) variable interpretations among radiologists, and (3) late appearance of overt positive radiographic signs of CHD.4,11,18–21 One study having 5-year radiographic follow-up suggested that radiographic hip phenotype at 2 years of age accurately reflected the true phenotype (and genotype) of the dog for life.13 A more recent retrospective survey from the OFA reported a small increase in hip dysplasia in a mixed population of dog breeds from a prevalence of 18% in dogs <1 year of age to 21.2% in dogs >4 years of age.22 There are no studies beyond 5 years of age that examine early hip scoring with later development of CHD, and to date it has been undetermined if end-of-life OA represents a long-term manifestation of CHD (secondary OA) or is unrelated and exists as primary, idiopathic OA.13,23,24 In fact, no lifelong studies in the dog have specifically investigated hip OA developing later in life nor has the accuracy of the 1-or-2-years-of-age hip evaluation been compared with ultimate, end-of-life hip phenotype.

Our purpose with this lifespan study was to examine age-dependent characteristics of CHD diagnosis including the timing of subluxation and radiographic OA, as subjectively scored from the ventrodorsal, HE radiograph. Positive and negative predictive values of radiographic OFA criteria scoring and of PennHIP scoring in 2-year-old dogs as compared to the end-of-life hip phenotype, both radiographic and histopathologic, are reported.

MATERIALS AND METHODS

The cohort of 48 Labrador retrievers followed for life in this study were part of a nutritional investigation looking into the role of lean body mass on the development of OA. Several reports have been published from investigations of this pool of dogs but none focusing on the chronology of radiographic evidence of CHD or on histopathologic findings at the natural end of life. Although the protocol of the lifespan study has been presented previously, it will be briefly included to permit readers and reviewers to assess the potential for conflict or interference between research objectives. When relevant, unpublished nutritional effects on age-related hip phenotype will be included.

Forty-eight 8-week-old Labrador retriever puppies from 7 litters were divided into 2 groups of 24 dogs each.5 The 7 litters were derived from 2 sires and 7 dams. Both males and 3 of the 7 females were scored OFA “normal,” 3 females were “dysplastic,” and the hip score of 1 female was unknown. Mating these sires and dams was expected to produce a 26–51% incidence of dysplastic offspring at 2 years of age.9,25 Within each litter, dogs were paired by same gender and similar body weight to minimize genetic differences between pairings before random assignment of dietary treatment.

All dogs were fed the same dry diet. One group was control-fed (CF), and each member of the diet-restricted (DR) group was given 25% less of the same food ingested by the CF pair-mate. Diet and feeding schedules were as previously described.9 Beginning at 6 years of age the Purina Body Condition System (BCS) ranging from 1 (emaciated) to 9 (severely obese) was used, where 5 represents ideal body condition.26,27

Dogs were anesthetized and evaluated radiographically with the standard ventrodorsal, HE projection at 4, 6, 8, 12, and 18 months, and at 2, 3, and 5 years of age, then yearly thereafter for life. The ventrodorsal, HE radiographs were read by a board-certified radiologist, according to criteria established by the OFA with respect to subluxation and severity of OA. The same radiologist also performed a 2nd independent radiographic evaluation for evidence and severity of OA. At the 2nd evaluation, the radiologist read each dog’s lifelong set of radiographs sequentially. Left and right coxofemoral joints of each dog were scored based on accepted radiographic standards for OA, including sclerosis of the craniodorsal portion of acetabular subchondral bone, osteophytes on the cranial aspect of acetabular margin, osteophytes on the caudal aspects of acetabular margin, and femoral periarthritic osteophytes.4,11 Caudolateral curvilinear osteophytes (CCO) and circumferential femoral head osteophytes (CFHO) were not factors included in the diagnosis of OA of CHD. The radiographs were marked in such a way that only the statistician knew the identity of the dog’s feeding regimen.9

PennHIP radiographic evaluation consisting of 3 views (HE, compression, and distraction) was performed on each dog at 2 years of age. Hip films were evaluated at the PennHIP Analysis Center at the University of Pennsylvania (Philadelphia, PA). The evaluator was unaware of the nutritional assignment of the dogs.

The established colony protocol, including housing, exercise, monitoring of illness, therapeutic measures, and euthanasia criteria and protocols, were previously described.4,9,11,27 After euthanasia, the dogs were necropsied and hips joints were submitted for histopathology. Disarticulated joints were examined visually, preserved in 10% buffered formalin, and photographed. Both visual examination and dissecting microscope were used to examine the coxofemoral joint. The joint capsule, articular surfaces, subchondral bone, and bone and joint conformation were examined for evidence of OA. A high-speed band saw was used to cut full-thickness, 20 × 8 mm sections of joint capsule (lateral, medial, and cranial aspects) and 5-mm sagittal and transverse sections of proximal and distal articular surfaces through the level of joint capsule attachment. Sections were evaluated according to standard histopathologic

Smith et al. Chronology of Hip Dysplasia Development in Labrador Retrievers

Veterinary Surgery 41 (2012) 20–33 © Copyright 2011 by The American College of Veterinary Surgeons 21
Table 1  Histopathologic Scoring Scheme

<table>
<thead>
<tr>
<th>Loc/Tis*</th>
<th>Distribution†</th>
<th>Degree‡</th>
<th>Character§</th>
<th>Descriptors¶</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synovium</td>
<td>None to diffuse (0–5)</td>
<td>None to severe (0–5)</td>
<td>None to chronic (0–3)</td>
<td>Hyperplasia/hypertrophy (1) Fibrosoing/fibromyxoid (1) Lymphocytic (1) (0–3)</td>
<td>16</td>
</tr>
<tr>
<td>Chondroosos metaplasia</td>
<td>None to diffuse (0–5)</td>
<td>None to severe (0–5)</td>
<td>None to chronic (0–3)</td>
<td>Fibrillation (1) Eburnation, (2) Bone deformity (2) Bone reaction/necrosis (2) (0–7)</td>
<td>13</td>
</tr>
<tr>
<td>Proximal articular surface</td>
<td>None to diffuse (0–5)</td>
<td>None to severe (0–5)</td>
<td>None to chronic (0–3)</td>
<td>Fibrillation (1) Eburnation, (2) Bone deformity (2) Bone reaction/necrosis (2) (0–7)</td>
<td>20</td>
</tr>
<tr>
<td>Distal articular surface</td>
<td>None to diffuse (0–5)</td>
<td>None to severe (0–5)</td>
<td>None to chronic (0–3)</td>
<td>Fibrillation (1) Eburnation, (2) Bone deformity (2) Bone reaction/necrosis (2) (0–7)</td>
<td>20</td>
</tr>
</tbody>
</table>

*Location/Tissue.
†Distribution: none (0), focal (1), multifocal (2), focally disseminated (3), coalescing (4), diffuse (5).
‡Degree: none (0), mild (0), mild-moderate (0), moderate (0), moderate-severe (0), severe (0).
§Character: none (0), acute (1), chronic-active (2), chronic (3).
¶Additional descriptors: hyperplasia/hypertrophy (1), fibrosing/fibromyxoid (1), lymphocytic (1), fibrillation (1), eburnation, (2), bone deformity (2), bone reaction/necrosis (2).

criteria (Table 1). The pathologist was unaware of the dogs’ feeding regimen.

Statistical Analyses

The Kolmogorov–Smirnov test for normality was used to determine whether subject weight distributions at 2 years of age for the DR group, the CF group, and both groups pooled were distributed normally. Wilcoxon signed rank test was used to test for differences in the median OFA-criteria scores between treatment groups. A signed rank test was performed separately for each year of age (1–14 years). The McNemar test was used to test for differences in the frequency of dysplasia/non-dysplasia diagnoses between treatment groups. A McNemar test was performed separately for each year of age (1–14 years). A paired sample t-test was used to test for differences in mean DI between dietary groups. Wilcoxon and Mc Nemar tests were also used to test for differences in severity and the presence/absence of histopathologic OA, respectively, between dietary groups. Correlation analysis was used to determine the degree of linear association ($r^2$) between the age and the cumulative prevalence of the appearance of radiographic evidence of OA. Kendall’s Tau-b correlation was used to test the association between DI groups and the presence of subluxation. The Kaplan–Meier product limit method was used to (1) compare differences in the OA-free interval (difference between groups in time to first diagnosis of radiographic evidence of OA) between OFA-criteria score groups over the lifetime of the dogs, (2) compare differences in the OA-free interval between groups with and without subluxation, and (3) compare differences in the OA-free interval among DI groups (2 years of age). Significance for the Kaplan–Meier analyses was tested by the log-rank test, pair-wise comparisons were performed by the Holm–Sidak method. For all analyses, statistical significance was set at $P < .05$). All analyses were performed with statistical software (SPSS, version 12, SPSS Inc., Chicago, IL; SAS, version 8, SAS Institute Inc., Cary, NC).

RESULTS

Median lifespan among the CF dogs was 11.2 years, compared with 13.0 years among the DR dogs. Mean body weight of dogs in the 2 groups closely paralleled differences in caloric intake. The DR dogs weighed ∼25% less than the CF dogs from 1 year of age onward. Mean (±SD) BCS of CF dogs were significantly ($P < .01$) higher (6.7 ± 0.19) than for dogs in the DR group (4.6 ± 0.19). Distributions of body weight at 2 years of age are given in Fig 1. In the pooled sample of 48 dogs (ie, both feed groups combined), body weight was normally distributed and not bimodal ($P = .873$) providing justification to consider the sample population as a single cohort with respect to age-related changes in hip phenotype.

OFA-Criteria Scores

There were distinct differences in CHD frequency based on OFA-criteria score associated with diet throughout the study (Fig 2). For example, at 2 years of age CHD (mild, moderate, or severe) was diagnosed in 5 dogs (21%) in the DR group, compared with 14 dogs (58%) in the CF group. Similar, diet-related differences in CHD frequency were recorded throughout the life of this cohort (Fig 2).
Grade (severity) of subjective hip score by age was also significantly affected by diet. That is, there was a significant difference in median hip scores between the 2 diet groups at all ages; Fig 3A shows the diet-related difference at 2 years of age (Wilcoxon signed rank test, \( P < 0.001 \)). Consistent with this scoring distribution, the median OFA-criteria score in the CF population was significantly larger (ie, higher dysplastic scores) than in the DR group at all ages. Therefore, OFA-criteria scores of DR group dogs were consistently better than those of the dogs in the CF group, and the DR dogs overall had a lower disease frequency, later disease onset, and slower disease progression than their matched CF littermates (Fig 2).

In the pooled sample of 48 dogs, OFA-criteria scoring identified 19 of the 48 dogs in the study to be dysplastic (mild, moderate, or severe) at 2 years of age and 29 dogs were scored as normal (ie, excellent, good, and fair). The 19 dysplastic dogs remained dysplastic for life; however, 35 dogs were scored dysplastic by the end of life, meaning 16 of the 19 dogs with “normal hips” at 2 years of age developed hip dysplasia by the end of life (Fig 4). It was observed in chronological review of hip films that OFA-criteria scores gradually increased in severity with age for most of the dogs, whether dysplastic or not. In summary OFA-criteria scorings at 2 years of age showed no false-positive diagnoses (except one dog that scored dysplastic at
3 years of age, became normal at 5 and 6, and dysplastic again from 7 onward, Fig 4), however a considerable number of false-negative diagnoses were observed. Twenty-nine dogs were scored normal at 24 months of age, of which 16 (55%) developed hip dysplasia (radiographically) by the end of life (Table 2, part [a]; Fig 4).

Components of OFA-criteria Scores
A previous publication from this sample of 48 Labrador retrievers reported the profound effect of diet on radiographic OA frequency.\textsuperscript{11} Pooling the sample of 48 dogs (Fig 1) showed the age-dependent prevalence of radiographic OA to increase linearly over the lives of the dogs from a
The cumulative prevalence of hip dysplasia throughout life for dogs was influenced drastically by dietary group (control-fed versus diet-restricted) as shown in Figure 2. But in addition, as shown (Fig 4), the pooled sample of dogs showed a linear increase in the prevalence of hip dysplasia with age, irrespective of dietary group. Note the flattening of the curve from 3 to 7 years of age, explained by the time-dependent onset of subluxation (see the text). Despite this flattening the plot from 1 to 14 years of age satisfies statistical criteria for linearity \( (P < .001) \). Such non-genetic or environmental influences, ie, age and diet, on a target phenotype diminish the value of that phenotype when used as an estimate of the genotype for breeder selection. That is, younger and leaner dogs have better hips phenotypically but not necessarily genotypically.

Table 2  
(a) Contingency Table: 2-year OFA-criteria Score versus the End-of-Life OFA-criteria Score; (b) Contingency Table: 2-year OFA-criteria Score versus the End-of-Life Prevalence of Radiographic OA; (c) Contingency Table: 2-year OFA-criteria Score versus the End-of-Life Histopathologic OA

<table>
<thead>
<tr>
<th>Pooled Groups</th>
<th>OFA-Criteria Score (Last Year of Life)</th>
<th>Radiographic OA (Last Year of Life)</th>
<th>Histopath OA (Last Year of Life)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CHD+</td>
<td>CHD−</td>
<td>OA+</td>
</tr>
<tr>
<td>OFA-criteria score (2 years of age)</td>
<td>19</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>13</td>
<td>14</td>
</tr>
</tbody>
</table>

prevalence of 7 (15%) at 2 years of age, to 19 (40%) at 8 years of age, to 32 (67%) at the end of life.\(^{11}\) In the current study, parallel findings were noted in OFA-criteria scoring (Fig 4). Subjective radiographic diagnosis of CHD increased linearly over the course of the study \( (r^2 = 0.90, P < .001) \). Approximately twice as many dogs were dysplastic according to OFA-criteria score at the end of life compared to the 2-year scoring (35 versus 19; Table 2, part [a]). There was a notable flattening of the CHD prevalence curve between 3 and 7 years of age (Fig 4), despite the curve’s satisfying statistical conditions of linearity over the lifespan of the dogs.

Subjective OFA-criteria scores, based on subluxation and OA, assigned at 2 years of age were evaluated for the ability to predict end-of-life radiographic hip OA. Of the 19 dogs scored dysplastic at 2 years of age, 18 dogs developed definitive radiographic signs of hip OA by the end of life, representing only 1 false-positive diagnosis. Of the 29 dogs receiving normal OFA-criteria scores at 2 years of age, 14 (48%) showed obvious radiographic changes of OA by the end of life (Table 2, part [b]) revealing a high number of false-negative diagnoses. The onset of first radiographic evidence of OA was correlated to the OFA-criteria scores (7 point) assigned at 2 years of age using Kaplan–Meier plots (Fig 5). Better OFA-criteria scores were associated with later onset of OA; however, OA was seen to manifest in dogs receiving normal OFA-criteria scores (Table 2, part [b]; Fig 5). There were no discernible differences in the onset and rate of OA expression between dogs graded OFA “good” and those graded OFA “fair” (Fig 5). There were no OFA-criteria “excellent” hips in this cohort; therefore, conclusions about timing, frequency, or severity of OA in this scoring category could not be made from this study sample.

Hip subluxation occurred up to 2 years of age and no new expression of subluxation was observed thereafter.
Dogs not showing OA or subluxation by 2 years of age had normal OFA-criteria scores at 2 years of age. All dogs having subluxation went on to develop radiographic OA (first dog with OA at 1 year of age; Fig 6). Of the dogs with subluxation, 43% (9/21) did not develop radiographic OA until after 5 years of age and 33% (7/21) after 6 years of age.

Diet affected the expression of subluxation. Only 25% (6/24) of the DR dogs had subluxation whereas 63% (15/24) of the CF dogs had subluxation (not depicted graphically).

Relationship of OFA-criteria Scores to Histopathologic OA

Forty-five of the 48 dogs were examined for gross and histopathologic evidence of OA at the end of their natural lives. Three dogs died before the age of 7, which was before the decision to submit joints for full histopathologic evaluation and were therefore not assessed histopathologically. Of the 45 dogs that were evaluated histopathologically, 43 (96%) had changes in the hip joint consistent with OA (Table 2, part [c]). Of the 2 dogs without histopathologic signs of OA, 1 had obvious signs of radiographic OA, presumably the lesion was missed during specimen sectioning. The other dog died at 10.7 years and was negative for OA both histopathologically and radiographically. Combining histologic and radiographic evidence showed that 44 of the 45 dogs (98%) developed characteristic signs of OA, either histopathologically or radiographically, by the end of life.

At the end of life, no statistical differences in severity of any of the individual or combined measures of histopathologic OA were noted between the 2 dietary groups. Hence, dietary groups could not be differentiated at the end of life based on histopathologic OA severity scores. Accordingly, the histopathologic score was evaluated as a binary (present or absent).

Of the 19 dogs with OFA dysplastic designation at 2 years of age, all (100%) showed histopathologic evidence
Kaplan–Meier curves of pooled dogs showing radiographic osteoarthritis (OA) free interval as a function of subluxation (+/−) and age. Of the dogs without subluxation, the first dog to show OA did not do so until 6 years of age, whereas dogs with subluxation started showing OA at 1 year of age. Of the dogs showing subluxation (and therefore scored dysplastic), 43% did not show radiographic OA until after 5 years of age. Of the dogs without subluxation 48% (13/27) expressed radiographic OA but after 5 years of age and 52% showed no radiographic OA. Interestingly, radiographic subluxation manifested by 2 years of age in this pooled sample of retrievers and not thereafter.

of OA by the end of life, meaning there were no false-positive diagnoses at 2 years of age. Of the 29 dogs with OFA-criteria normal scores at 2 years of age, 26 were evaluated histopathologically and 24 of these dogs (92%) had histopathologic evidence of OA characteristic of CHD, equating to an 8% negative predictive value (Table 2, part [c]).

Distraction Index (DI)

PennHIP hip screening was performed at 2 years of age. Diet (body weight) had no effect on the DI, i.e., the frequency distribution of hip laxity as measured by DI was not different whether the dogs were restricted-fed or CF (Fig 3B). Mean (± SD) DI of dogs in the CF group (0.57 ± 0.12) was not significantly different from those of dogs in the DR group (0.53 ± 0.11).

None of the 48 dogs in this cohort had hip laxity below a DI of 0.35. Therefore all dogs were classified as being “susceptible” to the OA of hip dysplasia. Accordingly, no conclusions could be drawn regarding true-negative or false-negative diagnoses. However, of the 48 dogs with DI’s >0.35, 32 (67%) developed radiographic signs of OA by the end of life. The onset of radiographic OA was closely correlated with DI: tighter hips developed radiographic OA later (Table 3; Fig 7). With each 0.1 increase in DI, the odds for having OA by the end of life increased by 2.2 times ($P < .001$).

Of the 45 dogs having histopathologic examination, 43 (96%) developed definitive histopathologic evidence of OA by the time of necropsy.

The percent of dogs with subluxation was correlated with DI (Kendall’s Tau-b correlation = −0.531; $P < .0001$).
For the dogs having the greatest hip laxity (DI scores > 0.60) 91% had subluxation. For dogs with DI scores between 0.51 and 0.60, 44% had subluxation, for dogs with DI scores between 0.41 and 0.50, 21% had subluxation, and for dogs with DI scores <0.40 but >0.35, none had subluxation (Table 3).

DISCUSSION

Hip dysplasia is a disease of complex inheritance. It is well recognized that environmental (non-genetic) factors can influence the expression of diseases of complex inheritance. In this lifespan investigation, food consumption, body weight and condition, and age were found to be potent environmental (non-genetic) factors affecting the onset and severity of OA. In terms of OFA criteria score, restricted food intake had a profound, beneficial effect on radiographic hip joint phenotype of Labrador retrievers. Diet-restricted dogs had significantly lower frequency and severity of CHD compared to their CF littermates. Likewise, age at the time of evaluation significantly influenced OFA-criteria scoring; the likelihood of CHD diagnosis increasing linearly after 1 year of age.

The capacity of environmental factors such as age and diet, to have such profound influence on the OFA-criteria phenotype in otherwise genetically similar, paired littermates is not a desirable feature of a selection test for determining breeding desirability. Such environmental (non-genetic) influences on test score lower the value and utility of a phenotypic test in making selection decisions. Dogs carrying the genes for disease may be made to appear phenotypically normal, depending on their environment, and thus may be approved for breeding while harboring the alleles for hip dysplasia. A better phenotypic test would be one strongly associated with the CHD genes but not influenced or masked by non-genetic (environmental) factors. This study showed that DI was unaffected by environmental factors such as diet and weight, and a previous study showed DI to be unchanged with age.

In the pooled sample, only 1 of 45 dogs was found to be free of both radiographic and histopathologic OA at the end of life, indicating an astonishingly high frequency of hip dysplasia in this Labrador retriever cohort. One could argue...
that this sample of Labrador retrievers was intentionally bred to have a high prevalence (26–51%) of hip dysplasia by 2 years of age\(^9\) (in fact, achieved 40% prevalence by 2 years of age), and therefore that such data do not apply to the population of Labrador retrievers at large. However, of interest is that the average DI of the sample of 48 dogs was 0.54 compared to the current average Labrador retriever DI in the PennHIP database of 0.49 (N = 20,500), which indicates that the OA susceptibility of the study sample is not far removed from the OA susceptibility of Labrador retrievers at large (Unpublished data, PennHIP database, 2010).

This study provides convincing evidence that Labrador retrievers can develop CHD long after 2 years of age, a finding that conflicts sharply with a 5-year study of Viszlas and German Shepherd dogs published in 1972 (Fig 8).\(^{13}\) This evidence also conflicts with the results of a recent report from the OFA which showed only a modest increase in prevalence of hip dysplasia in dogs <1 year of age (18%) compared with dogs >4 years of age (21.2%).\(^{22}\) The OFA study, however, did not prospectively follow a fixed cohort of a specific dog breed longitudinally and, as a retrospective survey, was likely confounded by the well-recognized and indeterminate selection bias.\(^{22, 38}\) In a more recent publication referencing OFA data, it was conceded that the existence of selection bias associated with the practice of voluntary film submission precluded making accurate estimates of CHD frequencies or changes therein.\(^{17}\)

Radiographic hip dysplasia prevalence figures from the Labrador retrievers in the present study were 31% at 1 year of age, 42% at 3 years, and 73% at the end of life (Figs 4 and 9). This age dependency of CHD as indicated by the linear increase in the OA component of CHD over the life of the dogs\(^{11}\) and the parallel increase in incidence and severity of CHD (OFA-criteria) are sobering statistics that warrant scrutiny of all current hip scoring schemes that rely solely on the ventrodorsal, HE radiograph. In the United Kingdom, Australia, Sweden, Europe, and much of the world, a similar subjective scoring system applied to the HE radiograph is performed in dogs even younger, at 1 year of age; a situation that magnifies this concern.

Longitudinal studies in a fixed cohort of dogs are costly but essential for understanding the biology of a complex disease such as CHD. To our knowledge, no similar studies on CHD have been published. Clearly, a screening test for hip dysplasia that is not influenced by non-genetic factors, such as diet, weight, and age, is important for more rapid progress in controlling CHD. The optimum screening test should be reliable early in life and should key on a

![Figure 8](image_url)  
**Figure 8** Rate of subjectively scored canine hip dysplasia (CHD) diagnosis (%) as a function of age. The rate (slope) of CHD diagnosis asymptotically approaches a limit (i.e., 0 slope) as dogs approach 5 years of age. After 2 years of age very few Viszlas or German Shepherd dogs (~5%) were diagnosed with CHD and accordingly 2 years of age or older was chosen as a practical age for hip certification by the OFA.\(^{13}\) Note that hip screening at 1 year of age (as is the norm in most of the other common hip-screening systems used worldwide) has approximately a 30% false-negative rate based on this data. (Reproduced with permission from Dr. Jessen, 1972.)

![Figure 9](image_url)  
**Figure 9** Combined plots of cumulative prevalence of Orthopedic Foundation for Animals (OFA) criteria scoring of hip dysplasia, radiographic osteoarthritis (OA), and theoretical cumulative prevalence of histopathologic OA in the context of ideal hip-screening behavior. The screening performance of the OFA-criteria method is an attempt to interpret radiographic hip laxity (subluxation) as an indicator of susceptibility to OA. Note that an ideal hip-screening system would predict at a very early age, the susceptibility to developing histopathologic evidence of OA later in life. The dashed line labeled “Hypothetical Histopath OA” is included to give an idea of how radiographic OA may lag behind histopathologic OA, a commonly held belief. The cumulative prevalence of histopathologic OA was estimated to be linear based on the actual observed linear cumulative prevalence of radiographic OA. The hip-screening performance of the PennHIP distraction index in this study approximated that of ideal hip-screening performance. See the text for complete explanation.
quantifiable phenotype with high heritability and high association with well-characterized signs of the disease. Higher heritability indicates that the phenotype is a truer measure of the genotype, less affected by environmental influences, and therefore a better selection criterion. The results of previous studies combined with those of this investigation indicate that the DI satisfies each of these conditions.\textsuperscript{2, 39, 40}

The only prior longitudinal study assessing CHD prevalence based on a subjective score was published in 1972 (Fig 8).\textsuperscript{13} This study found that 92–96% of the Viszla and German Shepherd dogs, respectively, that ultimately developed CHD by 5 years of age, were accurately diagnosed radiographically at 2 years of age.\textsuperscript{13} From these data, it was decided that radiographic evaluations beyond 5 years of age were unnecessary and so the 2 year score was adopted as the practical standard in the United States. A similar plot drawn from data of the present study was constructed (Fig 9). Before 5 years of age, the plot of OFA-criteria score in the 48 dogs shows a remarkably similar pattern to the report of 1972 (Fig 8). After 5 years of age however, the data from the present study showed that new cases of CHD or OA continued to develop (Figs 4 and 9). Therefore contrary to the conclusions from the earlier study, CHD and OA can occur anytime throughout the life of the dog and at a constant rate (constant slope).\textsuperscript{11} Many radiologists and orthopedists believe empirically that OA occurring late in life is idiopathic or “old-age” OA and unrelated to secondary OA of CHD. The linearity of OA shown in a previous study of this sampling of dogs and the linearity of CHD prevalence shown in this report suggest that all hip OA develops secondary to CHD and represents convincing evidence refuting the commonly held belief of primary hip OA in the canine hip.\textsuperscript{11, 41} Additional evidence is derived from the observation that 43% of the dogs with subluxation at 2 years of age, and therefore scored dysplastic, did not develop the OA of CHD until after 5 years of age (Fig 6). This evidence, that the OA of CHD-positive dogs can occur after 5 years of age, argues that late onset OA is secondary to CHD and, in fact, not primary OA. Dogs without subluxation and considered normal at 2 years of age began developing radiographic OA at 6 years of age and ultimately all but one dog had histopathologic OA by the end of life.

Before 5 years of age the frequency of radiographic CHD conventionally defined far exceeds the frequency of radiographic OA but after 6 years of age, both plotted frequencies continue to rise in parallel (Fig 9). After 6 years of age every new case of hip dysplasia is a new case of radiographic OA whereas no new cases of subluxation were found after 2 years of age (Table 3; Fig 9). Before 5 years of age, the diagnosis of CHD is reliant on radiographic evidence of subluxation and/or radiographic OA. For any given age, the vertical distance between OA and CHD prevalence (Fig 9) represents dogs whose CHD diagnosis was based solely on subluxation. At 2 years of age, radiographic subluxation is responsible for ~63% (12/19) of the CHD diagnoses, while at the end of life it corresponds to only 9% (3/32) of the CHD diagnoses. A novel finding in our study is that no new expression of subluxation on the HE radiograph was observed after 2 years of age. This explains the flattening of the CHD versus age curves (Fig 9) between 3 and 7 years of age. It is an artificial phenomenon based on the conventional definition of hip dysplasia, which includes either subluxation or OA. This study did not evaluate factors such as hind-limb positioning, breed of dog, form of anesthesia, or radiologist interpretation of subluxation. These and other factors may have influenced outcome.

It is generally accepted that radiographic evidence of joint OA lags behind histopathologic evidence. The results of our study were consistent with this impression: at the end of life, histopathologic OA was demonstrated in 96% of dogs compared with 67% of dogs showing characteristic radiographic evidence of OA.

Other newer methods of hip joint assessment have been applied to evaluate the hip with the goal of finding an early visual or molecular marker that will signal susceptibility to the OA of CHD. These include ultrasound,\textsuperscript{29–31} 2D and 3D CT,\textsuperscript{32–35} arthroscopy,\textsuperscript{36} and MRI.\textsuperscript{37} Few studies, however, compare the results of imaging with gross or histopathologic evaluation of the hip. One arthroscopic study found gross lesions of the cartilage in the absence of radiographic changes; results consistent with the present investigation.\textsuperscript{36} Another study looked at the relationships of various radiographic and 2D and 3D CT measurements of hip images when compared with histopathologic scoring of cartilage damage in 30-month-old mixed-breed hounds. The study showed the PennHIP OA score and the PennHIP DI to have good correlations with cartilage damage, 0.89 and 0.65, respectively, at 30 months of age.\textsuperscript{35} However, as shown in this study, the incidence and severity of cartilage damage will worsen with age, so it is the later outcome that is of utmost importance in deriving correlations or sensitivity and specificity estimates (Fig 9).\textsuperscript{11} For example, it was previously reported in this sample of 48 Labrador retrievers that 15% (7/48) of dogs had radiographic OA at 2 years of age; however, in truth, 67% (32/48) developed radiographic OA by the end of life.\textsuperscript{3} The correlations, therefore, between early hip score and later OA could vary greatly depending on the age the outcome data are collected. Our study suggests that the PennHIP DI, in contrast to OFA-type score and subluxation, had the best association with ultimate cartilage damage and this association improved with the age of the dogs studied. For any age-dependent chronic degenerative diseases like OA, older dogs give a truer estimate of disease prevalence, and therefore serve as a better outcome measure when assessing the sensitivity and specificity of an earlier predictive test. Histopathologic testing late in life, as performed in the present investigation, is a more reliable standard for making comparisons or correlations. This is not to say that histopathologic evidence of OA at the end of life should be considered the gold standard for CHD because it is now well understood that some of the undesirable genes for a quantitative disease, like CHD or cancer, can be present without causing expression of the disease in the lifetime of the individual. Obviously lifespan studies entail great expense and are therefore rare. One would hope that well-designed, multi-center, long-term clinical
trials using client-owned pets could address the accuracy of new and promising tests and treatments without incurring such prohibitive costs.

It is a useful exercise to extrapolate to the ideal hip-screening test. Such a test could use radiography, ultrasound, CT, MRI, or molecular methods to mention a few modalities. In terms of the present data, ideal hip-screening performance can be represented as a line starting at or around birth and extending vertically, showing the proportion of dogs that are CHD susceptible long before they manifest overt radiographic or histopathologic signs of CHD (see ideal hip-screening behavior, Fig 9). The hypothetical cumulative prevalence of histopathologic OA was represented as linear based on the actual cumulative prevalence of radiographic OA, which was found to be strongly linear (Fig 9). In this study, all but 1 of 45 dogs had histopathologic or radiographic OA by the end of life, so ideal hip-screening performance would be represented as a theoretical line extending vertically to 98% (assuming early susceptibility detection would be possible) and the line would then continue horizontally with aging to the 98% point. In the context of this “ideal hip-screening behavior,” the performance of OFA-criteria scoring gains perspective. It can be appreciated that the “bump” or apparent non-linearity in the plot of OFA-criteria diagnoses before 5 years of age (Fig 9), results from the diagnostic acceptance of hip subluxation as a predictor of those dogs susceptible to the OA of CHD. That is, dogs showing hip subluxation were scored dysplastic even before showing radiographic signs of OA, because it has long been accepted (but not tested) that such dogs inevitably will show signs of OA. By convention, and as confirmed for the first time by this data, subluxation is indeed predictive of eventual radiographic or histologic OA. All dogs with hip subluxation did indeed go on to develop hip OA. Unfortunately, in contrast to the model of ideal hip screening, the appearance of subluxation alone did not identify all of the OA-susceptible dogs at 2 years of age. Results of this study showed that 48% (13/27) of the dogs without subluxation at 2 years of age, and presumed normal, expressed radiographic OA beyond 2 years of age (Fig 6). This group of dogs therefore constituted a pool of false-negative diagnoses and arguably, such dogs should not have been certified to be free of susceptibility. This proportion of dogs occupies the vertical interval between the “Subjective Score” line at 2 years of age and the horizontal “Ideal Hip-Screening” line (Fig 9). Again, an ideal screening method would be capable of predicting with accuracy those dogs having susceptibility for the definitive radiographic or histopathologic signs of CHD long before the signs manifest. OFA-criteria scoring at 2 years of age was capable of identifying only 40% (19/47) of the dogs that would ultimately express signs of OA. However, it needs to be emphasized that the inclusion in the mid-20th century of radiographic subluxation as an “indicator” of eventual hip OA was appropriate because although it was not perfect in its prediction of OA, its accuracy was much better than the use of radiographic OA alone. Similarly the PennHIP DI is a further refinement and improvement in the ability to determine susceptibility to ultimate hip OA, ie, a better target phenotype. However, it is essential to understand that PennHIP in no way recommends that dogs must have hip laxity, DI < 0.3, (therefore have no susceptibility to hip OA), to be suitable for breeding. Such a practice would eliminate from breeding a large proportion of the most popular dog breeds, not to mention, it would create detrimental genetic bottlenecks. In a recent publication by Bell et al., from the OFA, this mistaken claim was made, that “more than 89%” of some breeds of dogs, such as the Labrador retriever, would be removed from breeding following PennHIP breeding recommendations. In fact, PennHIP breeding recommendations are patterned after time-tested principles of quantitative genetics. Within any given generation or birth year, the dogs (of a given breed) having hips in the better (tighter) 50th percentile are suitable candidates to enter the gene pool. This is a well-documented and proven breeding strategy that is used worldwide particularly in food animals. It allows for selection among at least half the population of individuals within any given generation or birth year, thereby avoiding genetic bottlenecks and permitting with each generation a gradual but methodical improvement in hip quality.

Our study draws critical attention to the previously unrecognized change in subjective hip score with aging. Currently there is no requirement or recommendation for dogs that are OFA-certified at 2 years of age to undergo repeat evaluations to validate the 2-year score. In fact, no radiographic follow-up is recommended by any of the hip scoring systems commonly used worldwide. Minimally, these new data warrant a strong recommendation for hip radiography well beyond 2 years of age. Particularly for breeding dogs, hips should be evaluated at regular intervals for the life of the dog to confirm validity of early radiographic screening and the ultimate fate of the dog’s hips even beyond breeding age. This practice will ultimately remove the false-negative diagnoses from the gene pool at the time that CHD is expressed. Unfortunately, until the time of definitive CHD expression, the gene pool remains replete with these unwanted genes. It is recognized that regular follow-up hip radiographs add considerable expense to dog breeding, but this expense pales in comparison to the expense and discomfort associated with rampant hip OA and associated pain-management therapies, particularly in dogs over 7 years of age. It is the pet and the ultimate end user, the pet owner, who benefits both financially and emotionally from having a dog with sound hips and therefore it is the pet owner who must ultimately shoulder the cost of improving canine hips. The PennHIP DI showed that all 48 dogs in the study were susceptible to the OA of CHD and indeed 98% of the dogs went on to develop either radiographic or histopathologic OA by the end of life. This hip-screening behavior approximated the ideal (Fig 9). A considerable limitation of our study is that none of the Labrador retrievers had DI scores below the disease threshold (< 0.3) and therefore no conclusions could be drawn about the negative predictive value of DI. However, from the Kaplan–Meier plot of OA-onset as a function of DI (Fig 7), it can be seen that tighter hips
developed OA later in life. In fact dogs having DI < 0.40 did not begin to show OA until 12 years of age and did not show subluxation ever (Table 3). The data suggest that had there been Labrador retrievers with even tighter hips (say DI < 0.30), no radiographic OA would have been observed throughout life. Such has been the experience with performance Borzoi and racing Greyhounds (GH), breeds with DIs consistently in this range (<0.30). Further support of this contention is derived from a report showing that both aged GH (mean age, 9.59 years having mean DI 0.235) as well as older non-GH with tight hips (mean age, 7.55 years and mean DI 0.286) have markedly lower frequencies of radiographic OA when compared with the high frequency of OA observed in age-matched dogs with greater hip laxity.4,41 This suggests that tight hips (DI < 0.3) confer to dogs of any breed the freedom from susceptibility (risk) for hip OA, and accordingly should be the preferred target phenotype.

This study exposed diagnostic deficiencies in the most popular hip scoring system used in the United States. Similar concerns likely apply to all hip scoring systems worldwide that rely on the ventrodorsal, HE pelvic radiographic assessment of dogs at 1 or 2 years of age for selecting breeding candidates. The diagnostic deficiencies identified are sufficient to explain why significant reduction in the incidence of CHD has not been made despite decades of application of the subjectively scored, HE radiographic method to control this pervasive disease.15,17 It is doubtful that further genetic improvement can be made using this screening method (coupled with mass selection) to select breeding dogs, at the very least, in Labrador retrievers. Some have suggested that a possible solution to this dilemma would be to use estimated breeding values as a better means to select candidate breeding dogs. Incorporating EBVs, however, would perhaps make for faster progress, but as a recent publication showed, the endpoint or target phenotype will not change.40,45 Even if one were to apply the maximum selection pressure by breeding only dogs receiving OFA “excellent” hip scores, the best achievable outcome, assuming all offspring would have OFA-criteria “excellent” hips would be that 52% of the offspring would be OA susceptible. For Labrador retrievers, 80% would be OA susceptible.40,45

These findings44–48 have profound clinical significance for the health and welfare of dogs. Making progress to reduce the high prevalence (and associated pain) of CHD requires the use of a hip-screening test having solid scientific validation.

REFERENCES


44. Ferguson R: Hip arthritis in older Greyhounds. Proceedings, Pan Pacific Veterinary Conference, Brisbane, Australia, May 2010


